International Journal of Interactive Multimedia and Artificial Intelligence

September 2024, Vol. VIII, Number 7
ISSN: 1989-1660

unir LA UNIVERSIDAD EN INTERNET

"We have a deep need and desire to connect. Everything in the history of communication technology suggests we will take advantage of every opportunity to connect more richly and deeply. I see no evidence for a reversal of that trend."
Peter Morville

Special Issue on Deep Learning Techniques for Semantic Web in Web of Things (WoT) and Internet of Everything (IoE)

INTERNATIONAL JOURNAL OF INTERACTIVE MULTIMEDIA AND ARTIFICIAL INTELLIGENCE

ISSN: 1989-1660 –VOL. 8, NUMBER 7

EDITORIAL TEAM

Editor's Note

With the rise of global economy and Electronic Commerce (EC), efficient inter-organizational planning and deployment for value chain processes have become important. Radio-frequency Identification (RFID), Near Field Communication (NFC), and related wireless technologies are evaluated to be some of the most significant technological innovations in the twenty-first century. In the past few years, wireless and context-awareness technology have led to much hope and optimism. The mainstream press hails these innovations as the avant-garde in technology and business. The Internet of Everything (IoE) goal is the intelligent connection of people, process, data, and things. The IoE describes a world where billions of objects have sensors to detect, measure, and assess their status, all connected over public or private networks using standard and proprietary protocols. Hence, this special issue investigates the state-of-art AI and deep learning approaches for successful systems or applications in the IoE environment. In addition, this special issue also wants to understand the direct and indirect effects of using these smart technologies to build language information processing based on the Web of Things (WoT) around the smart cities and societies.

The first article entitled "What Drives IoT-Based Smart Pet Appliances Usage Intention? The Perspective of the Unified Theory of Acceptance and Use of Technology Model", by Chen and Lin, investigate the key factors for pet owner adoption of "smart" pet appliances. The Unified Theory of Acceptance and Use of Technology (UTAUT) model is used as the main research framework. The statistical analysis finds out that performance expectancy, effort expectancy and facilitating condition all have a positive impact on the intention to use.

The second article entitled "Energy-Aware Path Planning by Autonomous Underwater Vehicle in Underwater Wireless Sensor Networks for Safer Maritime Transportation", by Acarer, solves the energy-aware path planning problem with autonomous underwater vehicles by using five popular bio-inspired algorithms including Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO), Teaching Learning-based Optimization (TLBO), Genetic Algorithm (GA), and Grey Wolf Optimizer (GWO). The experimental results indicate that the GA approach achieves better than others for the system convergence and energy consumption.

The third article entitled "Evaluating the Impact of Pumping on Groundwater Level Prediction in the Chuoshui River Alluvial Fan Using Artificial Intelligence Techniques", by Su et al., applies the multiple linear regression (MLR), support vector regression (SVR), and Long Short-Term Memory Networks (LSTM) to predict the groundwater level and collection data from the Chuoshui River Alluvial Fan (CRAF) area in Taiwan. The experimental results indicate that the LSTM model can achieve the better stability, strong generalization capabilities, and high prediction accuracy.

The fourth article entitled "Semi-Supervised Machine Learning Approaches for Thyroid Disease Prediction and its Integration With the Internet of Everything", by Agraz, implements a hybrid model of semi-supervised learning methods, namely FixMatch, Co-training, and self-training, in conjunction with other supervised learning methods, including Naive Bayes and logistic regression. This research finds the potential contributions of semi-supervised learning techniques and presents a good reference for the IoE in healthcare advancement.

The fifth article entitled "Enhancing Tennis Serve Scoring Efficiency: An AI Deep Learning Approach", by Liu, applies the deep learning model to automatically classify the serving position, landing position, and use of tennis techniques. The proposed model can achieve a 98.27% accuracy in the automatic classification of serve scores.

The sixth article entitled "The Human Motion Behavior Recognition by Deep Learning Approach and the Internet of Things", by Li et al., integrates the deep learning model - Convolutional Neural Networks (CNN) with the Internet of Things (IoT) technology for human motion behavior recognition. Finally, the recognition approach for human motion behavior can achieve an average accuracy of 94.41% and acts as a good reference for the intelligent surveillance and health management.

The seventh article entitled "Predicting Consumer Electronics E-Commerce: Technology Acceptance Model and Logistics Service Quality", by Wu et al., provides a structural model and machine learning algorithm with SHapley Additive exPlanations (SHAP) for a comprehensive analysis of Technology Acceptance Model (TAM) in conjunction with logistics service quality. The findings show that attitude, perceived usefulness, and informativeness are the most critical factors affecting the consumers' purchase intention.

The eighth article entitled "Constructing the Public Opinion Crisis Prediction Model Using CNN and LSTM Techniques based on Social Network Mining", by Lou et al., adopts the Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) techniques to build a predictive model for anticipating public opinion crises in social network mining. As a result, the hybrid CNN-LSTM model obtains a high accuracy rate of 92.19% with low loss value of 0.4075.

The ninth article entitled "Design of Traffic Electronic Information Signal Acquisition System Based on Internet of Things Technology and Artificial Intelligence", by Wang, designs and implements a real-time signal acquisition system within the IoT environment. It can promptly gather, analyze, and process collected signals. Finally, the experimental results prove that it can outperform the accuracy of existing equipment, making it more appropriate for traffic signal acquisition applications.

Mu-Yen Chen[1]
Patrick C. K. Hung[2]

[1] National Cheng Kung University (Taiwan)
[2] Ontario Tech University (Canada)

Acknowledgments

The guest editors are thankful for the best efforts of our authors and reviewers without whom the current issue and its publication in the journal itself would not have been possible. We also thank the Editor-in-Chief, Professor Elena Verdú and Professor Rubén González Crespo with their gentle assistance and guide in the whole process of preparation of this issue.

DOI: 10.9781/ijimai.2020.08.006

TABLE OF CONTENTS

OPEN ACCESS JOURNAL

ISSN: 1989-1660

The International Journal of Interactive Multimedia and Artificial Intelligence is covered in Clarivate Analytics services and products. Specifically, this publication is indexed and abstracted in: *Science Citation Index Expanded, Journal Citation Reports/ Science Edition, Current Contents®/Engineering Computing and Technology.*

What Drives IoT-Based Smart Pet Appliances Usage Intention? The Perspective of the Unified Theory of Acceptance and Use of Technology Model

Chia-Chen Chen*, Chia-Pei Lin

National Chung Hsing University, Taichung City 402 (Taiwan)

* Corresponding author: emily@nchu.edu.tw

Received 25 April 2023 | Accepted 6 October 2023 | Early Access 7 March 2024

Abstract

The advancement of IOT (Internet of Things) has facilitated the development of smart pet appliances, and the market for these products has growing rapidly, this study seeks to identify key factors for pet owner adoption of "smart" pet appliances. The Unified Theory of Acceptance and Use of Technology (UTAUT) a well-established model in the field of IOT research is used as the main framework, integrating brand trust, perceived value and perceived enjoyment as the basis for hypothesis formulation and testing based on data collected through questionnaires distributed through online social platforms. Reliability analysis, validity analysis and structural equation model analysis were carried out through confirmatory factor analysis to test the variables and research hypotheses. Results for the UTAUT indicate that effort expectancy has a direct impact on performance expectancy, while performance expectancy, effort expectancy and facilitating condition all have a positive impact on intention. While social influence does not directly or significantly affect use intention, it can indirectly affect intention through perceived value and perceived enjoyment. Brand trust does not have a significant impact on use intention, but can indirectly affect use intention through perceived value. This study further compares user age and number of smart pet home appliances owned to better understand the impact of demographic factors. Findings indicate that, for users under the age of 30, effort expectancy has no significant impact on use intention, while brand trust has no significant impact on perceived value among users over 30. Among the research results based on age as a basis, the impact of hardships in the ethnic group in the age of 30 is not significant, nor do facilitating conditions or perceived value have significant impact on use intention. For users with one smart pet device at home, neither favorable conditions not perceived value have significant impact on use intention, while for users with two smart pet devices, perceived enjoyment does not significantly impact use intention. These finding have potential reference value for future related research in the IOT or smart pet home appliance research field.

Keywords

Brand Trust, Internet Of Things (IoT), Perceived Value, Perceived Enjoyment, Smart Pet Appliances, Unified Theory Of Acceptance and Use Of Technology (UTAUT), Use Intention.

DOI: 10.9781/ijimai.2024.03.001

I. Introduction

The rapid development of wireless networks in recent years has driven the popularization of new technologies in the home. Such "Smart Family" technologies use wireless communications to integrate and coordinate smart appliances and devices (e.g., cameras, locks, kitchen appliances, speakers, etc.) to be monitored and controlled in real time using mobile phones or tablets. A report by the OMDIA research agency found that, from 2020 to 2025, the global smart home industry has a compound annual growth rate of 24.1%. In 2020, the global smart home market was valued at US$60.8 billion, and was projected to reach US$178 billion by 2025 [1], reflecting the rapid growth of this sector.

According to statistics from the Taiwan Council of Agriculture, the market value of the pet industry in Taiwan is projected to reach NT60 billion in 2022. Data from the Ministry of Finance indicates that the pet-related industry in Taiwan has been experiencing continuous growth in recent years. The total number of businesses in this sector has increased from 6,486 in 2018 to 8,335 in 2022, representing a growth rate of approximately 28.5%. Moreover, the sales revenue has risen from NT26.58 billion to NT38.7 billion, marking a significant growth rate of 45.7% [2]. With the rapid development of the economy and the rise in people's living standards, more and more individuals are choosing to keep pets for emotional companionship [3]. A 2020 report by the Market Intelligence & Consulting Institute (MIC) found that medical care and

Please cite this article in press as: C. C. Chen, C. P. Lin. What Drives IoT-Based Smart Pet Appliances Usage Intention? The Perspective of the Unified Theory of Acceptance and Use of Technology Model, International Journal of Interactive Multimedia and Artificial Intelligence, vol. 8, no. 7, pp. 5-14, 2024, http://dx.doi.org/10.9781/ijimai.2024.03.001

basic physiological needs (food, snacks and other daily necessities) accounted for the majority of pet owner spending, averaging more than NT$8,000 annually per pet owner, while the top three non-essential expense categories were grooming and accommodation (40.6%), home supplies (38.8%), entertainment and toys (37.6%). Thus, pet owners spend significantly on non-essentials [4].

This expansion of the pet supply market has also driven the development of new pet-oriented "smart" home appliances purporting to help pet owners enhance care convenience and safety. Nearly 40% of pet owners in Taiwan report having used some form of smart pet home appliance (e.g., water dispensers, litter boxes, interactive cameras, etc.). Online retailer PCHOME 24H found that pet owners are more willing to buy products they believe will help improve pet health and quality of life. In addition to smart appliances related to daily feeding, sales of other types of appliances are also growing rapidly, with overall sales jumping 30% in June, and sales for "black technologies" increasing 75% in the same period [5]. These developments reflect the growing opportunities in the pet supply market, particularly in terms of pet-oriented smart appliances.

Many studies in the literature on the Internet of Things (IoT) use the Unified Theory of Acceptance and Use of Technology (UTAUT) as the main architecture [6]-[9], with many such studies applying this model to integrate Perceived Value [10]-[13]. However, the existing literature on technology acceptance does not address the recent development of new smart pet home products. The authors have many pets at home and they and their friends use many types of smart pet home appliances and similar products, including automatic feeders, interactive pet cameras, and automatic litter boxes. However, while smart pet appliances have been proven to alleviate the burden on pet owners [14], most research literature focuses on the development of related technologies, with less emphasis on user intention [15]-[17]. Based on this experience and the related literature, this study uses UTAUT to integrate brand trust, perceived value, and perceived enjoyment to explore the factors that affect the use of smart pet home appliances. The study seeks to make the following contributions:

- Determine whether the UTAUT (Performance Expectancy, Effort Expectancy, Social Influence and Facilitating Conditions) can explain the use intention of smart pet home appliances.
- Determine whether brand trust affects use intention.
- Determine whether perceived value and perceived enjoyment affect use intention.

II. Literature Review

A. Pet Technology

Aashish (2022)[18] refers to smart pet home appliances as the use of technology products to improve the basic physiological care of pets, to monitor their health and safety, and to improve their overall quality of life. Such technologies can include the use of robotics, big data analysis, artificial intelligence, and others such as:

- Pet training equipment: Tools and equipment used to improve and modify pet behavior.
- Automatic feeding systems: Using sensors and remote monitoring to feed pets without direct human intervention.
- Pet monitoring equipment: Helping owners determine pet health and activity by remote.
- Pet toys: Entertainment and stimulation aids, or modes of remote interaction with owners.
- Wearable tracking devices: Small wearable devices to track pet health, activity and location.

Smart home appliances, integrated with IoT technology, have gained popularity among many people. By simply connecting to the internet, they can effectively automate various household activities [19]-[21]. Similarly, smart pet appliances assist pet owners in caring for their pets more efficiently. Apart from providing precise feeding, they also record the pet's behavior and analyze the data, enabling early detection and prevention of potential health issues [3], [22]. However, there is limited existing research specifically focusing on smart pet appliances. Therefore, this study will center on the topic of smart pet appliances.

B. Unified Theory of Acceptance and Use of Technology

The Unified Theory of Acceptance and Use of Technology (UTAUT) combines eight research models related to the Technology Acceptance Model (Combined TAM and TPB, C-TAM-TPB) and Social Cognitive Theory (SCT), including the components Theory of Reasoned Action (TRA) and Theory of Planning Behavior (TPB) by Venkatesh et al. (2003) [23], along with Technology Acceptance Model (TAM), Innovation Diffusion Theory (IDT), Motivational Model (MM), Model of PC Utilization (MPCU), and Planned Behavior Theory. Individually, these theories provide explanatory power in different fields, which Venkatesh et al. (2003) [23] integrated into four main facets:

- Performance Expectancy: the degree to which users believe that using new information technology can improve work performance.
- Effort Expectancy: the degree to which user believe an information system is easy to use.
- Social Influence: the degree to which users are aware of how others view their use of a new information technology.
- Facilitating Condition: the degree to which users believe that existing organizational or technical infrastructure supports the use of new information technologies.

In previous research, the Unified Theory of Acceptance and Use of Technology (UTAUT) has been applied to examine the usage behavior of smart home devices, which utilize IoT technology, such as health care systems, home appliances, security systems, and more. [24]-[26]. These studies found that factors like performance expectancy, effort expectancy, social influence, and facilitating conditions significantly influence users' intention to use and actual usage behavior [27], leading to improved personal well-being [28]. If a technology can improve individual performance or reduce inconvenience, people will have a higher willingness to use it [29]. However, few studies have extended this line of research to smart home appliances for pets, so this study applies UTAUT theory to examine user intention to use such pet-oriented devices.

C. Brand Trust

"Trust" has long been regarded as a catalyst for transactions between two parties. When a consumer does not have particular insight into a product, trust in the seller can reduce their purchasing uncertainty [30]-[32]. In the field of information technology, trust indicates the degree to which a user's expectations are met [33], and has an important impact on consumer behavior. Gefen (2000) [34] confirmed that trust helps consumers accept Internet technologies. Luor et al. (2015) [35] found a positive relationship between the extent of a user's trust in smart home appliances and their service attitudes towards such devices. Mashal & Shuhaiber (2019) [36] found that "trust", as a personal factor, affects the user's purchasing intention for smart home equipment, along with such factors as personalization and cost. Shuhaiber & Mashal (2019) [21] and Shomakers, Beirmann and Ziefle (2021) [37] also found that trust is an important factor in determining user intention to use smart home appliances. Therefore, this study uses the dimension of brand trust to explore degree of user trust in particular suppliers of smart pet home appliances.

D. Perceived Value

Perceived value means that consumers' decisions involve cost/benefit calculations [38]-[40] where such costs and benefits extend beyond monetary considerations. From a non-monetary perspective, the value dimension can be divided into the following five categories: functional value [41], [42], social value [42], cognitive value [41], affective value [41], [42] and conditional value [43]. In research related to information technology, Pitchayadejanant (2011) [11] found that performance expectancy, effort expectancy, and social influence do not directly affect usage intention, but indirectly affect usage intention through perceived value. Alwahaishi and Snasel (2013) [12] found that perceived value will affect consumer intention to use specific communication technologies. Xie et al. (2021) [13] also found that perceived value will positively enhance user intention to use financial technology platforms. Therefore, this study explores the impact of perceived value on use intention for pet-oriented smart home appliances.

E. Perceived Enjoyment

According to the third-generation Technology Acceptance Model (TAM) proposed by Mashal & Shuhaiber (2019) [36] and Venkatesh & Bala (2008) [44], perceived enjoyment is defined as system interaction stimulating feelings of interest, imaginativeness, meaning and creativity on the part of the user. Previous studies have found that perceived enjoyment can motivate users to adopt new information technologies. For example, Park et al. (2016) [45] found that perceived enjoyment has a significant impact on the willingness to use paid LTE services, while Mashal, Shuhaiber & Daoud (2020) [46] and Shuhaiber & Mashal (2019) [21] found that perceived enjoyment will positively affect users' attitudes towards the use of smart home appliances, which then affect usage intention. Mashal & Shuhaiber (2019) [36] found that perceived enjoyment will positively increase users' willingness to purchase smart home appliances, while Al Amri & Almaiah (2021) [47] found that perceived enjoyment is directly and positively correlated to learners' intention to use smart educational technologies. This study explores the impact of perceived enjoyment on use intention for pet-oriented smart home appliances.

F. Operational Variable Definitions

The research dimensions examined in this study include performance expectancy, effort expectancy, social influence, facilitating conditions, brand trust, perceived value, perceived enjoyment and sue intention, assessed using a 5-point Likert scale from 1 (strongly disagree) to 5 (strongly agree). These variables are defined as follows:

- Performance Expectancy: The degree to which consumers believe that the use of smart pet appliances can improve quality of life.
- Effort Expectancy: The degree of ease consumers experience using smart pet appliances.
- Social Influence: The degree to which consumers are influenced by the opinions of others when using smart pet appliances.
- Facilitating Conditions: The extent to which the technical knowledge or infrastructure required by consumers to use smart pet appliances is available.
- Perceived Enjoyment: The degree of pleasure consumers feel when using smart pet appliances.
- Perceived Value: The subjective value consumers feel through the use of smart pet appliances.
- Brand Trust: Consumers' trust in smart pet appliance suppliers.
- Intention to use: The possibility of consumers using smart pet appliances in the future.

III. Research Method

A. Research Constructs

This research explores people's willingness to use smart pet appliances, mainly through the Unified Theory of Acceptance and Use of Technology (UTAUT), integrating brand trust, perceived value, and perceived enjoyment, because trust can reduce consumers' uncertainty about using products [30]-[32], while perceived value allows consumers to assess the benefits of their decision-making behavior (i.e., performance expectancy) and decision-making effort (i.e., effort expectancy) of the decision-making result [38], and perceived enjoyment can motivate users to adopt new information technologies [45]. Based on the above, the research framework is as shown in Fig. 1.

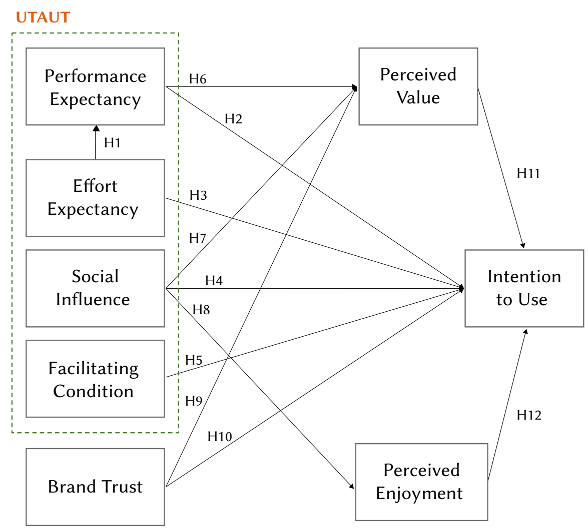


Fig. 1. Research Model.

B. Research Hypotheses

This study will employ a questionnaire survey method to collect data and validate the research hypotheses. Additionally, the samples will be divided into groups based on age and the number of different types of smart pet appliances used, and further analysis will be conducted on each group individually [9], [23], [48], [49]. Based on the literature review and research framework, this study develops 12 research hypotheses as follows:

1. Relationship Between Performance Expectancy and Effort Expectancy

In previous research on the Internet of Things and smart homes, perceived usefulness is defined as the degree to which users think that using smart homes will improve their quality of life, and perceived ease of use is defined as the degree to which users think that using smart homes does not require physical or mental labor. Perceived ease of use is also positive correlated with perceived usefulness [21], [50]-[52]. Based on the technology acceptance model and related research, we propose the following hypothesis:

H1: Effort expectancy will positively affect consumers' performance expectancy.

2. Relationship Between UTAUT and Usage Intention

In the IoT context, perceived usefulness is the degree to which the consumer perceives the technology as improving his or her overall performance in everyday situations [53]. Perceived ease of use is defined as the degree to which a person perceives that using a system requires no mental effort [54]. Venkatesh et al. (2003) [23] define social influence as other people's perceptions of whether consumers should use new technologies. Facilitating condition is defined as the consumers' perception of resources and support available to perform behaviors [23]. Regarding the use of digital technology or Internet of Things (IoT) technologies, previous research has confirmed that Performance expectancy, Effort expectancy, Social influence, and Facilitating conditions are all important factors influencing users' intention to use [27], [55]-[59]. Based on the above, this study suggests that consumers' intention to use smart pet home appliances will be affected by UTAUT, and proposes the following hypotheses:

H2: Performance expectancy will positively impact consumers' use intention.

H3: Effort expectancy will positively impact consumers' use intention.

H4: Social influence will positively impact consumers' use intention.

H5: Facilitating conditions will positively impact consumers' use intention.

3. Effect of Performance Expectancy on Perceived Value and Use Intention

According to Venkatesh et al. (2012) [60], performance expectancy describes the degree to which individuals believe that they will secure benefits from using new technologies. For Xie et al. (2021) [13], performance expectancy refers to the degree to which individuals believe that they benefit from using online wealth management platforms, and these benefits are perceived value. Performance expectancy, which reflects the actions individuals take based on their desire for extrinsic rewards [23], are related to the "received" part of perceived value, so it is argued that performance expectancy affects perceived value in the same way. Based on the above, the following hypothesis is proposed:

H6: Performance expectancy will positively impact consumers' perceived value.

4. Effect of Social Influence on Perceived Value, Perceived Enjoyment and Use Intention

Social influence refers to the degree to which consumers are influenced by the opinions of others to change their original attitudes or behavioral intentions [23]. In our daily lives, receiving positive feedback about the use of a particular product from those whose opinions we value will positively impact our opinion towards the product. Such people indicating they feel that the product presents good value-for-money or can provide enjoyment, will also positively affect our feelings towards the product [61]. Based on the above, we propose the following hypotheses:

H7: Social influence will positively impact consumers' perceived value.

H8: Social influence will positively affect consumers' perceived enjoyment

5. Effect of Brand Trust on Perceived Value and Use Intention

Studying the use of IoT products in the agricultural sector, Ha & Stoel (2009) [62] found that trust plays a more critical role in IoT-related IT services than in brick-and-mortar industries, due to the inherent intangibility of IoT services and the lack of face-to-face interaction between technology suppliers and users. In a study of global B2B services, Doney et al. (2007) [63] found a positive relationship between perceived value and trust. Based on the above, we propose the following hypotheses:

H9: Brand trust will positively impact consumers' perceived value.

H10: Brand trust will positively impact consumers' use intention.

6. Effect of Brand Trust on Perceived Value and Use Intention

Studying the use of IoT products in the agricultural sector, Ha & Stoel (2009) [62] found that trust plays a more critical role in IoT-related IT services than in brick-and-mortar industries, due to the inherent intangibility of IoT services and the lack of face-to-face interaction between technology suppliers and users. In a study of global B2B services, Doney et al. (2007) [63] found a positive relationship between perceived value and trust. Based on the above, we propose the following hypotheses:

H11: Perceived value will positively impact consumers' use intention.

H12: Perceived enjoyment will positively impact consumers' use intention.

IV. Results

A. Descriptive Statistical Analysis

The questionnaire was constructed and hosted using Surveycake. According to data statistics, the main age group that primarily uses smart home appliances, internet technology, or digital technology falls within the 18-34 age range [64]-[66]. The resulting data set was analyzed using IBM SPSS 25 to obtain respondents' descriptive statistics, including gender, age, and number of smart pet appliances used. The results are presented in Table I.

TABLE I. Descriptive Statistics

Item	Response	n	%
Gender	Male	46	16.5%
	Female	233	83.5%
Age	< 30	146	52.3%
	>= 30	133	47.7%
Number of Smart Pet Appliances Used	One	140	50.2%
	Two and more	139	49.8%

B. Reliability

Reliability analysis mainly seeks to determine whether the items in each research construct of the questionnaire are consistent, stable and reliable, wherein a Cronbach's α threshold value of 0.7 indicates adequate reliability [67]. The Cronbach's α values for each facet of our questionnaire are shown in Table II.

TABLE II. Reliability Analysis Results

Research Construct	Cronbach's α
Performance Expectancy	0.898
Effort Expectancy	0.878
Social Influence	0.866
Facilitating Conditions	0.786
Brand Trust	0.950
Perceived Value	0.865
Perceived Enjoyment	0.854
Intention To Use	0.893

C. Convergent Validity

Convergent validity analysis uses factor loading, composite reliability (CR), and average variation extracted (AVE) as three indicators to measure whether the degree of correlation between items in the same facet converges sufficiently. Fornell & Larcker (1981) [68] suggest a minimum factor loading of 0.5, a minimum composite reliability of 0.7, and a minimum AVE of 0.5. As shown in Table III, the results for the research constructs in the present study all present good convergent validity.

TABLE III. Convergent Validity Results

Research Construct		Factor Loading	CR	AVE
Performance Expectancy (PE)	PE1	0.824	0.8471	0.6487
	PE2	0.794		
	PE3	0.798		
Effort Expectancy (EE)	EE1	0.837	0.8777	0.7051
	EE2	0.853		
	EE3	0.829		
Social Influence (SI)	SI1	0.811	0.8422	0.6418
	SI2	0.873		
	SI3	0.711		
Facilitating Conditions (FC)	FC1	0.837	0.8164	0.5986
	FC2	0.691		
	FC3	0.786		
Brand Trust (BT)	BT1	0.867	0.9358	0.7848
	BT2	0.908		
	BT3	0.889		
	BT4	0.879		
Perceived Value (PV)	PV1	0.730	0.7071	0.5470
	PV2	0.749		
Perceived Enjoyment (PENJ)	PENJ1	0.753	0.7587	0.5203
	PENJ2	0.845		
	PENJ3	0.529		
Intention To Use (ITU)	ITU1	0.704	0.7847	0.5488
	ITU2	0.768		
	ITU3	0.749		

D. Discriminant Validity

Discriminant validity analysis tests whether correlations can be distinguished between the constructs, seeking a higher correlation between questionnaire items within a single construct, and a lower correlation between items in different constructs. Fornell & Larcker (1981) [68] suggest the square root of the AVE for each construct should exceed the Pearson Correlation Coefficient of each construct. The results in Table IV show that the questionnaire used in this study has good discriminant validity.

TABLE IV. Correlation Coefficient Matrix for Each Construct

	BT	PE	EE	SI	ITU	FC	PENJ	PV
BT	0.886							
PE	0.388	0.805						
EE	0.343	0.448	0.840					
SI	0.377	0.481	0.395	0.801				
ITU	0.358	0.649	0.487	0.534	0.741			
FC	0.404	0.403	0.399	0.384	0.463	0.774		
PENJ	0.464	0.470	0.367	0.576	0.637	0.393	0.721	
PV	0.428	0.620	0.436	0.501	0.677	0.401	0.676	0.740

Note: The diagonal line is the square root of the AVE for each construct.

E. Model Fit

Structural equation modeling is a statistical method that uses factor and path analysis to verify research hypotheses. It explores the causal relationship and degree of influence among variables, and uses model fitness to evaluate the fit between the research framework model and sample data. Following Hair et al. (1998) [69], we use Absolute Fit Measures, Incremental Fit Measures and Parsimonious Fit Measures, with the model satisfying 11 measurement indicators, indicating that the research model presents adequate model fitness.

F. Hypothesis Validation

Following verification of reliability, validity and model fitness, this study uses AMOS 24 for structural analysis. To explore the relationship between the various model constructs, a structural model was established to test the various hypotheses, with the standardized path coefficients and hypothesis validation results shown in Fig. 2. Positive and significant impacts are found for effort expectancy on performance expectancy ($\beta = 0.490$, $p < 0.001$; H1), performance expectancy on use intention ($\beta = 0.345$, $p < 0.001$; H2), effort expectancy on use intention ($\beta = 0.125$, $p < 0.001$; H3), facilitating conditions on use intention ($\beta = 0.114$, $p < 0.001$; H5), performance expectancy on perceived value ($\beta = 0.470$, $p < 0.001$; H6), social influence on perceived value ($\beta = 0.271$, $p < 0.001$; H7), social influence on perceived hedonic ($\beta = 0.654$, $p < 0.001$; H8), brand trust on perceived value ($\beta = 0.186$, $p < 0.001$; H9), perceived value on use intention ($\beta = 0.226$, $p < 0.001$; H11) and perceived enjoyment on use intention ($\beta = 0.359$, $p < 0.001$; H12). However, the two hypotheses regarding the impact of social influence on use intention (H4) and brand trust on use intention (H10) are not supported.

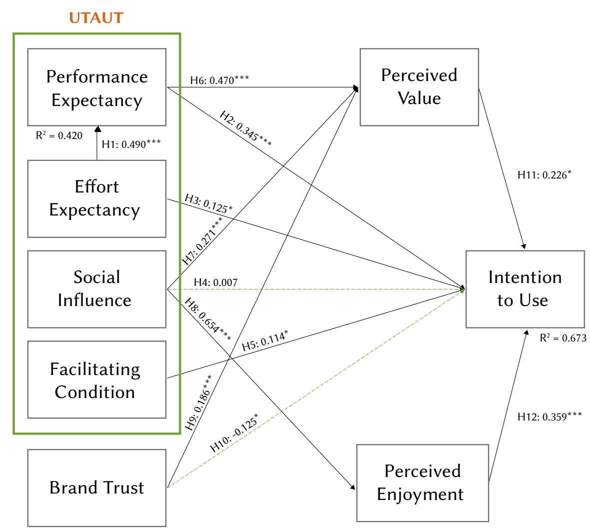


Fig. 2. Path Analysis Results.

G. Comparative Results for Age

Previous studies on the adoption of IoT services found that the impact of consumer age on behavioral attitude and response is subject to a variety of [6], [9], [70]. Chua (2004) [71] noted that most East Asian pop culture consumption occurs among people under the age of 30. This study uses 30 as the cutoff age for group analysis, and the results presented in Fig. 3 and Fig. 4 show the following differences in the use intention for smart pet appliances among different age groups:

1. Among users under 30, performance expectancy is the most important factor influencing use intention, which means that these users are mainly concerned with the usefulness of smart pet appliances. The secondary factor affecting usage intention is perceived value, and the main factor affecting perceived value is social influence, indicating that these users are sensitive to peer attitudes towards smart pet appliances, which in turn will affect perceived value, and finally willingness to use.
2. Among user over 30, perceived enjoyment is the most important factor affecting use intention, and the main factor affecting perceived enjoyment is social influence, which means that this group is affected by peer attitudes regarding hedonic emotions, which then affects use intention. The secondary reason affecting use intention is the implications of effort expectancy on perceived ease of use. Previous studies found that older people tend to

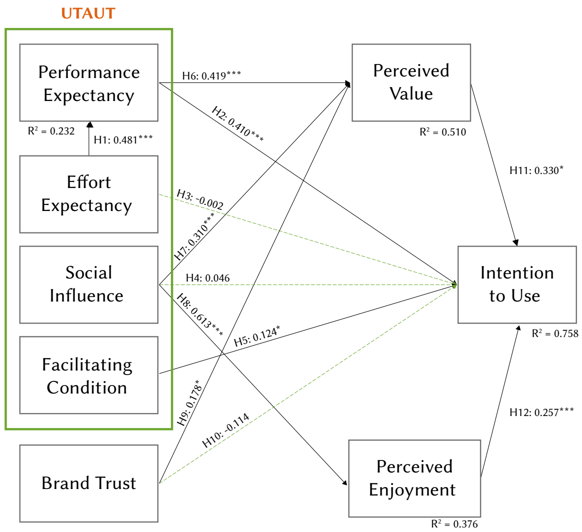


Fig. 3. Path analysis results (under 30 years old).

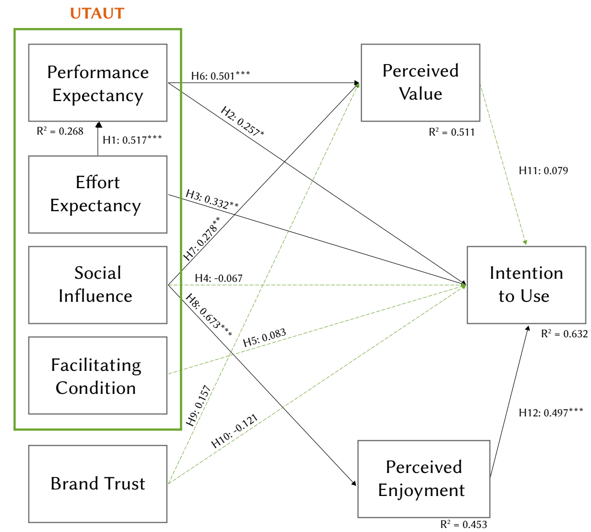


Fig. 4. Path analysis results (over 30 years old).

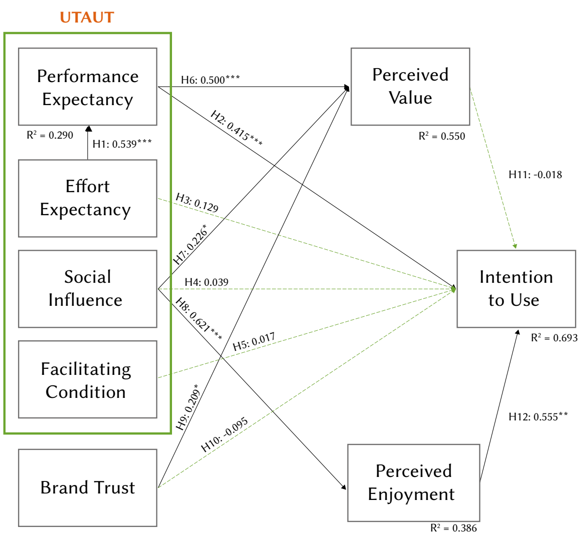


Fig. 5. Path analysis results (one appliance).

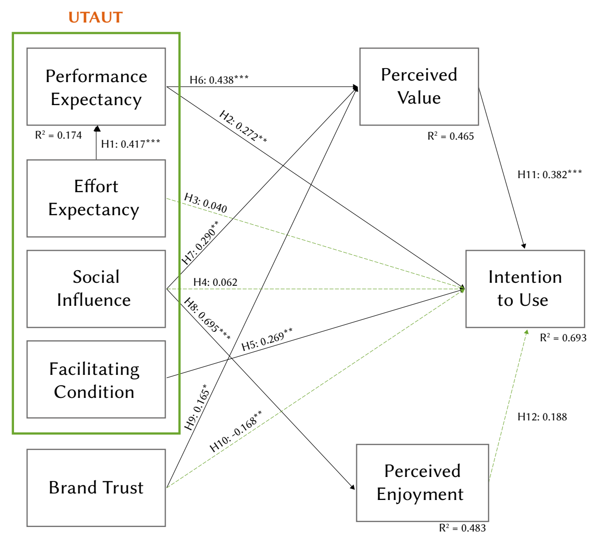


Fig. 6. Path analysis results (two or more appliance).

experience relatively higher levels of anxiety when learning to use computers [23], therefore, effort expectancy plays a very important role in the use intention of this group.

3. Social influence affects performance expectancy among users under 30, and performance expectancy is the most important factor affecting use intention; in users over 30, social influence affects perceived enjoyment, which in turn is the key impact factor for use intention While social influence does not directly affect the use intention, peer opinions are an important component for other use factors.

H. Number of Appliances in Use

According to the third-generation technology acceptance model (Venkatesh & Bala, 2008) and the Unified Theory of Acceptance and Use of Technology (UTAUT) [23], use experience will impact use intention. Previous research on Internet of Things applications suggest that user experience will affect users' willingness to adopt [9], [48], [49], thus this study conducted grouping and hypothesis verification analysis on users in terms of numbers of smart pet appliances used, with results for various classifications shown in Fig. 5 and Fig. 6. In terms of users with different numbers of smart pet appliances (1 vs. 2 or more), a significant difference was found in terms of use intention:

1. For owners of a single smart pet appliance, perceived enjoyment has a strong direct impact on use intention, and the main factor affecting perceived enjoyment is social influence. This shows that such users are sensitive to peer opinion, which affects their enjoyment usage motivation, which in turn affects their use intention. The second factor that affects use intention performance expectancy, thus in addition to the enjoyment they derive from using smart pet appliances, these users place significant value on increasing their convenience in caring for their pets and enhancing their quality of life.
2. Among users with two or more smart pet appliances at home, the most important factor affecting use intention is performance expectancy, which means that their main goal in using multiple smart pet appliances is to improve the efficiency of pet care and pet quality of life. The second factor affecting use intention is perceived value, which in turn is affected by performance expectancy. Thus, the product usefulness is the key factor impacting use intention among this group of users.
3. While there are differences in the impact factors between these two groups of users, overall performance expectancy and effort expectancy still play a major role in adoption for both groups.

V. Conclusions And Recommendations

A. Conclusions

This study uses the Unified Theory of Acceptance and Use of Technology (UTAUT) proposed by Venkatesh et al. (2003) [23], combining three variables (brand trust, perceived value and perceived enjoyment) to explore factors impacting consumers' intentions to use smart pet appliances.

Previous research on the Internet of Things found that effort expectancy has a significant impact on performance expectancy [21], [50]. In the present study, overall sample analysis or clustering results both indicate that effort expectancy positively impacts consumers' performance expectancy. If users feel new technologies will be easy to use, and can increase productivity without excessive additional effort, they will perceive such technologies as having high perceived usefulness.

In terms of the relationship between UTAUT and use intention, the results of this study show that performance expectancy impacts the use intention of smart pet appliances, indicating that product usefulness is an important consideration in determining product usage. These findings are consistent with those of Gao & Bai (2014) [72], who found that usefulness is an important driver of IoT technology adoption, and Kowatsch & Maass (2012) [53] who found that users adopt IoT technologies perceived as being conducive to promoting work productivity. In addition, effort expectancy is found to be a factor affecting use intention for smart pet home appliances, where ease of learning and use correlates with increased willingness to adopt. Yong et al., (2011) [57] found that ease of use is an important determinant for the acceptance of new technologies, particularly among older users, which is consistent with the present finding that ease of use is a particular consideration for users over the age of 30. Facilitating conditions are found to have a positive impact on use intention for smart pet appliances, consistent with the findings of Abushkra et al., (2019) [8] that adoption is effectively promoted by access to sufficient technical knowledge or infrastructure support [8]. However, this result does not hold for users under the age of 30 or those who only have one smart pet home appliance. This is possibly explained because single-device use does not require technical acumen or support including integration, device cross-compatibility, WiFi networking, or integration with mobile apps. Overall sample analysis and clustering results do not support the influence of social influence on usage intention. Venkatesh et al. (2003) [23] note the impact of social influence is weaker in regards to new technologies in early development stages. Rogers (2003) [73] also notes that, given the limited number of users of emerging technologies, the impact of social influence remains low until the technology matures.

On the other hand, performance expectancy is found to impact perceived value in both the overall sample and group analysis results. Consumers who use smart pet appliances feel they are useful and can effectively help them overcome challenges. This is consistent with Kim & Chan (2007) [74] who found that performance expectancy has a positive impact on perceived value. The results also show that social influence impacts perceived value and perceived enjoyment, consistent with Li (2011) [75] and Pitchayadejanant (2011) [11] who found that social influence indirectly enhances use intention through perceived value and perceived enjoyment. That is, social influence affects the degree of perceived value and perceived enjoyment. One's peers expressing positive attitudes towards the value or enjoyment derived from the use of smart pet appliances will positive impact one's use intention.

The influence of brand trust on use intention is not established in the overall sample or group analysis results, possibly because smart pet appliances are a relatively recent product category, and are not in widespread use, thus consumers have not had time or exposure needed to acquire relevant information or develop brand loyalty. Therefore, brand trust does not constitute an influencing factor for use intention.

Perceived value has a positive impact on use intention among users over the age of 30 and among users with two or more such appliances. However, perceived enjoyment was found to have a significant impact on use intention for users both above and below the age of 30, along with users who only have one smart pet appliance. The group analysis presented in the previous chapter shows that the factor with the greatest effect on the impact of perceived value and perceived enjoyment on use intention is social influence, which means that consumers' use intention is influenced by peer attitudes.

B. Research Contributions

The increasing ubiquity of advanced technologies has driven increasing research on topics related to the Internet of Things and smart home appliances. A recent addition to this product category is smart pet home appliances, for which little research has been conducted. Motivated by the authors' own experience and that of their peer group, this study examines factors that may influence consumers when adopting smart pet home appliances. Data collection and analysis potentially provide a better understanding of the future of consumer adoption patterns in this product segment, and the study makes the following contributions:

1. Expanding the application scope of the Unified Theory of Acceptance and Use of Technology (UTAUT) to the domains of IoT, smart home appliances, and smart pet appliances. Additionally, this study includes Brand trust, Perceived Value, and Perceived Enjoyment as additional factors within the research model for investigation.
2. Identifying factors that affect perceived value and perceived enjoyment.
3. Establishing the importance of brand trust on perceived value.
4. The research findings can serve as a valuable reference for developers or manufacturers in the smart home appliance and IoT technology industry, enabling them to develop products that better meet the expectations and needs of the researchers.

C. Limitations and Future Work

This study seeks to identify factors that affect the use intention of smart pet appliances, but such factors operate on multiple levels. Despite the rigor of the research design, the contribution of many factors remains ambiguous, thus the research results should be interpreted with caution and future work should seek further clarification by:

1. Increasing sample diversity: Survey respondents were largely between 20 and 33 years old, with a significant gender imbalance, which could potentially result in insufficient sample representativeness. Therefore, future work should seek to broaden the age range of respondents and normalize the gender distribution, thereby avoiding excessive data concentration in the resulting clusters.
2. Evaluate and select influencing factors: Consumer use intention is subject to a wide range of influencing factors. However, due to time constraints and other considerations, the present study selected only eight factors in the final model. Future work should seek to incorporate other theoretical perspectives to increase the comprehensiveness of research findings.
3. Enrich data analysis clustering criteria: Regarding the clustering analysis of data, this study solely focuses on conducting comparative analysis of data concerning different age groups and the quantity of smart pet appliances used. In future research,

it would be valuable to expand the analysis by incorporating variables such as usage duration or gender, as these factors may also impact the user experience and yield different research outcomes.

Appendix

Research Questionnaire

Construct	Item	Reference
Performance Expectancy (PE)	*(PE1)* Using smart pet appliances can improve my efficiency. *(PE2)* Using smart pet appliances will improve my quality of life. *(PE3)* Using smart pet appliances saves time in pet care.	Davis et al. (1989) [54], Moore & Benbasat (1991) [76]
Effort Expectancy (EE)	*(EE1)* Smart pet appliances are easy to learn to use. *(EE2)* Smart pet appliances are easy to operate. *(EE3)* I don't need help to use smart pet appliances.	Davis et al. (1989) [54], Moore & Benbasat (1991) [76]
Social Influence (SI)	*(SI1)* There are many online recommendations for smart pet appliances. *(SI2)* My peers recommend the use of smart pet appliances. *(SI3)* My peers support my use of smart pet appliances	Venkatesh et al. (2003) [23], Venkatesh & Zhang (2010) [77], Foon & Fah (2011) [78]
Facilitating Condition (FC)	*(FC1)* I have the required network environment to use smart pet appliances. *(FC2)* I have the knowledge needed to use smart pet appliances. *(FC3)* Smart pet appliances are compatible with my other devices (e.g., mobile apps).	Ajzen (1991) [79], Taylor & Todd (1995) [80]
Perceived Enjoyment (PE)	*(PENJ1)* I enjoy interacting with my pet using smart pet appliances. *(PENJ2)* I enjoy using smart pet appliances. *(PENJ3)* Using smart pet appliances is worth the required time.	Venkatesh et al. (2012) [60]
Perceived Value (PV)	*(PV1)* Using smart pet appliances is worth the required effort. *(PV2)* Using smart pet appliances is worth the required time.	Sweeney & Soutar(2001) [81]
Brand Trust (BT)	*(BT1)* I trust the brand of smart pet appliances I use. *(BT2)* I think smart pet appliance brands are reliable. *(BT3)* The smart pet appliance brand I use is trustworthy. *(BT4)* I have confidence in the brand of smart pet appliance I use.	Hsu et al. (2014) [82], Delgado-Ballester (2004) [83]
Intention to Use (ITU)	*(ITU1)* I am willing to use smart pet appliances. *(ITU2)* I will probably use smart pet appliances in the future. *(ITU3)* I plan to use smart pet appliances.	Davis et al.(1989) [54], Venkatesh &Zhang (2010) [77]

Acknowledgment

The authors wish to thank the National Science and Technology Council of the Republic of China for financially supporting this research under Contract Grants No. 111-2410-H-005-022-MY3 & 111-2410-H-005-023-.

References

[1] OMDIA. "Omdia report finds purpose-driven smart homes will lead to a market size of $178bn in 2025." Accessed: Oct. 1, 2022. [Online]. Available: https://omdia.tech.informa.com/pr/2021-sep/omdia-report-finds-purposedriven-smart-homes-will-lead-to-a-market-size-of-178bn-in-2025.

[2] LnData-Taiwan. "[Data Feature] Pet Economy Takes Off!" Accessed: Jul. 1, 2023. [Online]. Available: https://lndata-taiwan.medium.com/pet-trends-report-2023-13c721720cb8.

[3] Q. Su, Y. Liu, and H. Wu, "Smart Pet Feeding Device Based on Single Chip Microcomputer," *Journal of Physics: Conference Series,* vol. 1885, pp. 052032, 2021, doi: 10.1088/1742-6596/1885/5/052032.

[4] M. I. C. Institute (MIC). "Pet consumer survey 1: 51-55 year olds are the key target market for pet industry, while 26-35 year olds are the biggest users of pet technology products" is a commonly used translation in Taiwan." Accessed: Jul. 1, 2023. [Online]. Available: https://mic.iii.org.tw/news.aspx?id=574&List=1.

[5] PChome Online. "Smart pet care" is becoming a trend! PChome 24h Shopping's smart pet appliances have seen a whopping 75% surge in performance, and the pet industry is growing against the trend, presenting a great business opportunity for pet lovers." Accessed: Jul. 1, 2023. [Online]. Available: https://corp.pchome.tw/

[6] W. B. Arfi, I. B. Nasr, T. Khvatova, and Y. B. Zaied, "Understanding acceptance of eHealthcare by IoT natives and IoT immigrants: An integrated model of UTAUT, perceived risk, and financial cost," *Technological Forecasting and Social Change,* vol. 163, pp. 120437, 2021, doi: 10.1016/j.techfore.2020.120437.

[7] W. Lee and S. Shin, "An empirical study of consumer adoption of Internet of Things services," *International Journal of Engineering and Technology Innovation,* vol. 9, no. 1, pp. 1-11, 2019.

[8] A. Abushakra and D. Nikbin, "Extending the UTAUT2 model to understand the entrepreneur acceptance and adopting internet of things (IoT)," in *Knowledge Management in Organizations: 14th International Conference,* Zamora, Spain, pp. 339-347, 2019.

[9] K. Nikolopoulou, V. Gialamas, and K. Lavidas, "Habit, hedonic motivation, performance expectancy and technological pedagogical knowledge affect teachers' intention to use mobile internet," *Computers and Education Open,* vol. 2, pp. 100041, 2021, doi: 10.1016/j.caeo.2021.100041.

[10] N. Shaw and K. Sergueeva, "The non-monetary benefits of mobile commerce: Extending UTAUT2 with perceived value," *International journal of information management,* vol. 45, pp. 44-55, 2019, doi: 10.1016/j.ijinfomgt.2018.10.024.

[11] K. Pitchayadejanant, "Intention to use of smart phone in Bangkok extended UTAUT model by perceived value," in Proceedings of the *International conference on management,* pp. 160-172, 2011.

[12] S. Alwahaishi and V. Snásel, "Acceptance and use of information and communications technology: a UTAUT and flow based theoretical model," *Journal of technology management & innovation,* vol. 8, no. 2, pp. 61-73, 2013, doi: 10.4067/S0718-27242013000200005.

[13] J. Xie, L. Ye, W. Huang, and M. Ye, "Understanding FinTech platform adoption: impacts of perceived value and perceived risk," *Journal of Theoretical and Applied Electronic Commerce Research,* vol. 16, no. 5, pp. 1893-1911, 2021, doi: 10.3390/jtaer16050106.

[14] S. Kim, "Smart pet care system using internet of things," *International Journal of Smart Home,* vol. 10, no. 3, pp. 211-218, 2016, doi: 10.14257/ijsh.2016.10.3.21.

[15] K. Jadhav, G. Vaidya, A. Mali, V. Bankar, M. Mhetre, and J. Gaikwad, "IoT based automated fish feeder," in *2020 International Conference on Industry 4.0 Technology,* Pune, India, pp. 90-93, 2020, doi: 10.1109/I4Tech48345.2020.9102682.

[16] M. A. Hidayat and S. Jayakrista, "Smart Pet Feeder on Cat Food Portions

Using Mamdani's Fuzzy Logic Inference System Method," *Journal of Computer Engineering, Electronics and Information Technology,* vol. 2, no. 1, pp. 9-20, 2023.

[17] Y. Quiñonez, C. Lizarraga, R. Aguayo, and D. Arredondo, "Communication architecture based on IoT technology to control and monitor pets feeding," *Journal of Universal Computer Science,* vol. 27, no. 2, pp. 190-207, 2021, doi: 10.3897/jucs.65094.

[18] P. Aashish. " What Is Pet Tech? - Use Cases, Examples, & Future.Stay Ahead with Feedough." Accessed: Oct. 1, 2022. [Online]. https://www.feedough.com/pet-tech-definition-use-cases-examples-future/.

[19] W. Gu, P. Bao, W. Hao, and J. Kim, "Empirical examination of intention to continue to use smart home services," *Sustainability*, vol. 11, no. 19, pp. 5213, 2019, doi: 10.3390/su11195213.

[20] M. Kim and B. R. Choi, "The impact of privacy control on users' intention to use smart home internet of things (IoT) services," *Asia Marketing Journal,* vol. 24, no. 1, pp. 29-38, 2022, doi: 10.53728/2765-6500.1586.

[21] A. Shuhaiber and I. Mashal, "Understanding users' acceptance of smart homes," *Technology in Society,* vol. 58, pp. 101110, 2019.

[22] Y. Chen and M. Elshakankiri, "Implementation of an IoT based pet care system," in *2020 Fifth International Conference on Fog and Mobile Edge Computing,* Paris, France, 2020, pp. 256-262, doi: 10.1109/FMEC49853.2020.9144910.

[23] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: Toward a unified view," *MIS quarterly,* vol. 27, no. 3, pp. 425-478, 2003, doi: 10.1016/j.techsoc.2019.01.003.

[24] L. Ferreira, T. Oliveira, and C. Neves, "Consumer's intention to use and recommend smart home technologies: The role of environmental awareness," *Energy,* vol. 263, pp. 125814, 2023.

[25] S. F. Hussin, M. F. Abdollah, and I. B. Ahmad, "Acceptance of IoT Technology for Smart Homes: A Systematic Literature Review," in *International Conference on Information Systems and Intelligent Applications*, pp. 187-202, 2023, doi: 10.1007/978-3-031-16865-9_16.

[26] W. B. Arfi, I. B. Nasr, G. Kondrateva, and L. Hikkerova, "The role of trust in intention to use the IoT in eHealth: Application of the modified UTAUT in a consumer context," *Technological Forecasting and Social Change,* vol. 167, pp. 120688, 2021, doi: 10.1016/j.techfore.2021.120688.

[27] H. J. Kang, J. Han, and G. H. Kwon, "The Acceptance Behavior of Smart Home Health Care Services in South Korea: An Integrated Model of UTAUT and TTF," *International Journal of Environmental Research and Public Health,* vol. 19, no. 20, pp. 13279, 2022, doi: 10.3390/ijerph192013279.

[28] H. Sequeiros, T. Oliveira, and M. A. Thomas, "The impact of IoT smart home services on psychological well-being," *Information Systems Frontiers,* vol. 24, pp. 1009-1026, 2022, doi: 10.1007/s10796-021-10118-8.

[29] B. A. Akinnuwesi et al., "A modified UTAUT model for the acceptance and use of digital technology for tackling COVID-19," *Sustainable Operations and Computers,* vol. 3, pp. 118-135, 2022, doi: 10.1016/j.susoc.2021.12.001.

[30] P. S. Ring and A. H. Van de Ven, "Developmental processes of cooperative interorganizational relationships," *Academy of management review,* vol. 19, no. 1, pp. 90-118, 1994, doi: 10.2307/258836.

[31] P. M. Doney and J. P. Cannon, "An examination of the nature of trust in buyer–seller relationships," *Journal of marketing,* vol. 61, no. 2, pp. 35-51, 1997, doi: /10.2307/1251829.

[32] M. S. Featherman and P. A. Pavlou, "Predicting e-services adoption: a perceived risk facets perspective," *International journal of human-computer studies,* vol. 59, no. 4, pp. 451-474, 2003, doi: 10.1016/S1071-5819(03)00111-3.

[33] S. Cannizzaro, R. Procter, S. Ma, and C. Maple, "Trust in the smart home: Findings from a nationally representative survey in the UK," *Plos one,* vol. 15, no. 5, pp. e0231615, 2020, doi: 10.1371/journal.pone.0231615.

[34] D. Gefen, "E-commerce: the role of familiarity and trust," *Omega,* vol. 28, no. 6, pp. 725-737, 2000, doi: 10.1016/S0305-0483(00)00021-9.

[35] T. T. Luor, H. P. Lu, H. Yu, and Y. Lu, "Exploring the critical quality attributes and models of smart homes," *Maturitas,* vol. 82, no. 4, pp. 377-386, 2015, doi: 10.1016/j.maturitas.2015.07.025.

[36] I. Mashal and A. Shuhaiber, "What makes Jordanian residents buy smart home devices? A factorial investigation using PLS-SEM," *Kybernetes,* vol. 48, no. 8, pp. 1681-1698, 2018, doi: 10.1108/K-01-2018-0008.

[37] E. M. Schomakers, H. Biermann, and M. Ziefle, "Users' preferences for smart home automation–investigating aspects of privacy and trust," *Telematics and Informatics,* vol. 64, pp. 101689, 2021, doi: 10.1016/j.tele.2021.101689.

[38] E. J. Johnson and J. W. Payne, "Effort and accuracy in choice," *Management science,* vol. 31, no. 4, pp. 395-414, 1985.

[39] V. A. Zeithaml, "Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence," *Journal of marketing,* vol. 52, no. 3, pp. 2-22, 1988, doi: 10.2307/1251446.

[40] K. B. Monroe, *Pricing: Making profitable decisions*, 2nd ed., New York, USA: McGraw-Hill College, 1990.

[41] T. F. Stafford, "Consumption values and the choice of marketing electives: Treating students like customers," *Journal of Marketing Education,* vol. 16, no. 2, pp. 26-33, 1994, doi: 10.1177/027347539401600204.

[42] G. LeBlanc and N. Nguyen, "Listening to the customer's voice: examining perceived service value among business college students," *International Journal of Educational Management,* vol. 13, no. 4, pp. 187-198, 1999, doi: 10.1108/09513549910278106.

[43] R. M. Unni, "Value perceptions and retention of textbooks among marketing and other business majors," *Marketing Education Review,* vol. 15, no. 2, pp. 71-79, 2005, doi: 10.1080/10528008.2005.11488909.

[44] V. Venkatesh and H. Bala, "Technology acceptance model 3 and a research agenda on interventions," *Decision sciences,* vol. 39, no. 2, pp. 273-315, 2008, doi: 10.1111/j.1540-5915.2008.00192.x.

[45] S. T. Park, H. Im, and K.-S. Noh, "A study on factors affecting the adoption of LTE mobile communication service: The case of South Korea," *Wireless Personal Communications,* vol. 86, pp. 217-237, 2016 doi: 10.1007/s11277-015-2802-7.

[46] I. Mashal, A. Shuhaiber, and M. Daoud, "Factors influencing the acceptance of smart homes in Jordan," *International Journal of Electronic Marketing and Retailing,* vol. 11, no. 2, pp. 113-142, 2020, doi: 10.1504/IJEMR.2020.106842.

[47] M. Al Amri and M. A. Almaiah, "Sustainability Model for Predicting Smart Education Technology Adoption Based on Student Perspectives," *International Journal of Advances in Soft Computing & Its Applications,* vol. 13, no. 2, pp. 60-77, 2021.

[48] X. Dong, Y. Chang, Y. Wang, and J. Yan, "Understanding usage of Internet of Things (IOT) systems in China: Cognitive experience and affect experience as moderator," *Information Technology & People,* vol. 30, no. 1, pp. 117-138, 2017, doi: 10.1108/ITP-11-2015-0272.

[49] D.-H. Shin, "Conceptualizing and measuring quality of experience of the internet of things: Exploring how quality is perceived by users," *Information & Management,* vol. 54, no. 8, pp. 998-1011, 2017, doi: 10.1016/j.im.2017.02.006.

[50] A. Shuhaiber, I. Mashal, and O. Alsaryrah, "Smart homes as an IoT application: predicting attitudes and behaviours," in 2019 IEEE/ACS 16th International Conference on Computer Systems and Applications, pp. 1-7, 2019, doi: 10.1109/AICCSA47632.2019.9035295.

[51] S. Chatterjee, "Factors impacting behavioral intention of users to adopt IoT in India: from security and privacy perspective," *International Journal of Information Security and Privacy,* vol. 14, no. 4, pp. 92-112, 2020, doi: 10.4018/IJISP.2020100106.

[52] A. Al-Husamiyah and M. Al-Bashayreh, "A comprehensive acceptance model for smart home services," *International Journal of Data and Network Science,* vol. 6, no. 1, pp. 45-58, 2022, doi: 10.5267/j.ijdns.2021.10.005.

[53] T. Kowatsch and W. Maass, "Critical privacy factors of internet of things services: An empirical investigation with domain experts," *in Knowledge and Technologies in Innovative Information Systems: 7th Mediterranean Conference on Information Systems,* Guimaraes, Portugal, pp. 200-211, 2012.

[54] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw, "User acceptance of computer technology: A comparison of two theoretical models," *Management science,* vol. 35, no. 8, pp. 982-1003, 1989.

[55] X. Wang, C. F. Lee, J. Jiang, G. Zhang, and Z. Wei, "Research on the Factors Affecting the Adoption of Smart Aged-Care Products by the Aged in China: Extension Based on UTAUT Model," Behavioral Sciences, vol. 13, no. 3, p. 277, 2023, doi: 10.3390/bs13030277.

[56] T. Coughlan, M. Brown, R. Mortier, R. J. Houghton, M. Goulden, and G. Lawson, "Exploring Acceptance and Consequences of the Internet of Things in the Home," in 2012 IEEE international conference on green computing and communications, IEEE, pp. 148-155, 2012.

[57] Y. W. Sek, S. H. Lau, K. K. Teoh, C. Y. Law, and S. B. Parumo, "Prediction of user acceptance and adoption of smart phone for learning with

technology acceptance model," *Journal of Applied Sciences (Faisalabad),* vol. 10, no. 20, pp. 2395-2402, 2010, doi: 10.3923/jas.2010.2395.2402.

[58] T. A. Sykes, V. Venkatesh, and S. Gosain, "Model of acceptance with peer support: A social network perspective to understand employees' system use," *MIS quarterly,* vol. 33, no. 2, pp. 371-393, 2009, doi: 10.2307/20650296.

[59] R. Gatautis and A. Medziausiene, "Factors affecting social commerce acceptance in Lithuania," *Procedia-Social and Behavioral Sciences,* vol. 110, pp. 1235-1242, 2014, doi: 10.1016/j.sbspro.2013.12.970.

[60] T. F. Stafford, "Consumption values and the choice of marketing electives: Treating students like customers," *Journal of Marketing Education,* vol. 16, no. 2, pp. 26-33, 1994, doi: 10.1177/027347539401600204.

[61] G. LeBlanc and N. Nguyen, "Listening to the customer's voice: examining perceived service value among business college students," International *Journal of Educational Management,* vol. 13, no. 4, pp. 187-198, 1999, doi: 10.1108/09513549910278106.

[62] R. M. Unni, "Value perceptions and retention of textbooks among marketing and other business majors," *Marketing Education Review,* vol. 15, no. 2, pp. 71-79, 2005, doi: 10.1080/10528008.2005.11488909.

[63] V. Venkatesh and H. Bala, "Technology acceptance model 3 and a research agenda on interventions," *Decision sciences,* vol. 39, no. 2, pp. 273-315, 2008, doi: 10.1111/j.1540-5915.2008.00192.x.

[64] Statista. "Distribution of internet users worldwide as of 2021, by age group." Accessed: Jul. 1, 2023. [Online]. Available: https://www.statista.com/statistics/272365/age-distribution-of-internet-users-worldwide/.

[65] Statista (2023). "Awareness of smart or connected home devices among adults in the United States as of July 2020, by age group." Accessed: Jul. 1, 2023. [Online]. Available:https://www.statista.com/statistics/1201173/united-states-smart-home-device-awareness-by-age/.

[66] DATAREPORTAL (2022). "DIGITAL 2022: GLOBAL OVERVIEW REPORT." Accessed: Jul. 1, 2023. [Online]. Available: https://datareportal.com/reports/digital-2022-global-overview-report.

[67] J. C. Nunnally, *Psychometric Theory,* 2nd ed., New York, USA: Mcgraw hill book company, 1978.

[68] C. Fornell and D. F. Larcker, "Evaluating structural equation models with unobservable variables and measurement error," *Journal of marketing research,* vol. 18, no. 1, pp. 39-50, 1981, doi: 10.2307/3151312.

[69] J. Hair, R. E. Anderson, R. L. Tatham, and W. C. Black, *Multivariate data analysis with readings,* Englewood Cliffs, NJ, USA: Prentice Hall, 1998.

[70] J. Shin, Y. Park, and D. Lee, "Who will be smart home users? An analysis of adoption and diffusion of smart homes," Technological Forecasting and Social Change, vol. 134, pp. 246-253, 2018, doi: 10.1016/j.techfore.2018.06.029.

[71] C. B. Huat, "Conceptualizing an East Asian popular culture," *Inter-Asia Cultural Studies,* vol. 5, no. 2, pp. 200-221, 2004, doi: 10.1080/1464937042000236711.

[72] L. Gao and X. Bai, "A unified perspective on the factors influencing consumer acceptance of internet of things technology," *Asia Pacific Journal of Marketing and Logistics,* vol. 26, no. 2, pp. 211-231, 2014, doi: 10.1108/APJML-06-2013-0061.

[73] E. M. Rogers, "A prospective and retrospective look at the diffusion model," *Journal of health communication,* vol. 9, no. S1, pp. 13-19, 2004, doi: 10.1080/10810730490271449.

[74] H. W. Kim, H. C. Chan, and S. Gupta, "Value-based adoption of mobile internet: an empirical investigation," *Decision support systems,* vol. 43, no. 1, pp. 111-126, 2007, doi: 10.1016/j.dss.2005.05.009.

[75] D. C. Li, "Online social network acceptance: a social perspective," *Internet research,* vol. 21, no. 5, pp. 562-580, 2011, doi: 10.1108/10662241111176371.

[76] G. C. Moore and I. Benbasat, "Development of an instrument to measure the perceptions of adopting an information technology innovation," *Information systems research,* vol. 2, no. 3, pp. 192-222, 1991, doi: 10.1287/isre.2.3.192.

[77] V. Venkatesh and X. Zhang, "Unified theory of acceptance and use of technology: US vs. China," *Journal of global information technology management,* vol. 13, no. 1, pp. 5-27, 2010, doi: 10.1080/1097198X.2010.10856507.

[78] Y. S. Foon and B. C. Y. Fah, "Internet banking adoption in Kuala Lumpur: an application of UTAUT model," *International Journal of Business and Management,* vol. 6, no. 4, p. 161, 2011, doi: 10.5539/ijbm.v6n4p161 .

[79] I. Ajzen, "The theory of planned behavior," *Organizational behavior and human decision processes,* vol. 50, no. 2, pp. 179-211, 1991, doi: 10.1016/0749-5978(91)90020-T.

[80] S. Taylor and P. Todd, "Decomposition and crossover effects in the theory of planned behavior: A study of consumer adoption intentions," *International journal of research in marketing,* vol. 12, no. 2, pp. 137-155, 1995, doi: 10.1016/0167-8116(94)00019-K.

[81] J. C. Sweeney and G. N. Soutar, "Consumer perceived value: The development of a multiple item scale," *Journal of retailing,* vol. 77, no. 2, pp. 203-220, 2001, doi: 10.1016/S0022-4359(01)00041-0 .

[82] M. H. Hsu, C. M. Chang, K. K. Chu, and Y. J. Lee, "Determinants of repurchase intention in online group-buying: The perspectives of DeLone & McLean IS success model and trust," *Computers in Human Behavior,* vol. 36, pp. 234-245, 2014, doi: 10.1016/j.chb.2014.03.065.

[83] E. Delgado-Ballester, "Applicability of a brand trust scale across product categories: A multigroup invariance analysis," *European journal of Marketing,* vol. 38, no. 5/6, pp. 573-592, 2004, doi: 10.1108/03090560410529222.

Chia-Chen Chen

Dr. Chen is a Distinguished Professor of Management Information Systems at National Chung Hsing University, Taiwan. Her current research interests focus on information management, educational technology, deep learning, human behavior, and social interactions. Dr. Chen's has been named to the prestigious World's Top 2% Scientists list published by Stanford University in 2021, 2022, & 2023.

Chia-Pei Lin

Miss Lin received the M.S. degree with the Department of management information systems at National Chung Hsing University, Taiwan. She is interesting on information management, human behavior, and social interactions.

Energy-Aware Path Planning by Autonomous Underwater Vehicle in Underwater Wireless Sensor Networks for Safer Maritime Transportation

Tayfun Acarer

Maritime Transportation and Management Vocational School of Higher Education, Piri Reis University, Istanbul (Turkey)

* Corresponding author: tacarer@hotmail.com

Received 21 April 2024 | Accepted 2 August 2024 | Early Access 20 August 2024

Abstract

Throughout history, maritime transportation has been preferred for international and intercontinental trade thanks to its lower cost than other transportation ways, which have a risk of ship accidents. To avoid these risks, underwater wireless sensor networks can be used as a robust and safe solution by monitoring maritime environment where energy resources are critical. Energy constraints must be solved to enable continuous data collection and communication for environmental monitoring and surveillance so they can last. Their energy limitations and battery replacement difficulties can be addressed with a path planning approach.This paper considers the energy-aware path planning problem with autonomous underwater vehicles by five commonly used approaches, namely, Ant Colony Optimization-based Approach, Particle Swarm Optimization-based Approach, Teaching Learning-based Optimization-based Approach, Genetic Algorithm-based Approach, Grey Wolf Optimizer-based Approach. Simulations show that the system converges faster and performs better with genetic algorithm than the others. This paper also considers the case where direct traveling paths between some node pairs should be avoided due to several reasons including underwater flows, too narrow places for travel, and other risks like changing temperature and pressure. To tackle this case, we propose a modified genetic algorithm, the Safety-Aware Genetic Algorithm-based Approach, that blocks the direct paths between those nodes. In this scenario, the Safety-Aware Genetic Algorithm-based approach provides just a 3% longer path than the Genetic Algorithm-based approach which is the best approach among all these approaches. This shows that the Safety-Aware Genetic Algorithm-based approach performs very robustly. With our proposed robust and energy-efficient path-planning algorithms, the data gathered by sensors can be collected very quickly with much less energy, which enables the monitoring system to respond faster for ship accidents. It also reduces total energy consumption of sensors by communicating them closely and so extends the network lifetime.

Keywords

Atificial Intelligence, Autonomous Underwater Vehicle, Energy-aware Path Planning, Maritime Commerce, Maritime Industry, Maritime Operations, Optimization Algorithm, Ship Management Systems, Safe Sailing Planning, Underwater Wireless Sensor Networks, Water Monitoring.

DOI: 10.9781/ijimai.2024.08.003

I. Introduction

There has been a close relationship between maritime and trade throughout history. Therefore, the most important reason for this is that the majority of international trade and especially intercontinental transportation is carried out using maritime transportation [1]. The most important factor in this is that sea transport is 3.5 times cheaper than railways, 7 times cheaper than road transport, and 22 times cheaper than air transport. This cost advantage causes the importance and volume of maritime transportation to increase day by day [2]. According to data from the International Chamber of Shipping (ICS), 90% of world trade is carried out by sea today [3]. For this reason, the report of the United Nations Conference on Trade and Development predicts that world maritime trade will grow at an annual growth rate of 3.8% between 2018 and 2023 [4]. The increase in world maritime trade causes intense maritime traffic and the inevitable result of this is the increase in the risk of maritime accidents. Historical data shows that these accidents generally occur on the busiest routes [5].

Increasing both the volume and value of the cargo transported over time further magnifies the damage caused by accidents in maritime transportation. It is not possible to define the cost of loss of life occurring during these accidents in monetary terms [6]. For example, the ship accident and the transportation blockage in the Suez Canal [7]–[9] caused severe economic consequences in the global supply chains

Please cite this article as:
T. Acarer. Energy-Aware Path Planning by Autonomous Underwater Vehicle in Underwater Wireless Sensor Networks for Safer Maritime Transportation, International Journal of Interactive Multimedia and Artificial Intelligence, vol. 8, no. 7, pp. 15-27, 2024, http://dx.doi.org/10.9781/ijimai.2024.08.003

like its impacts on the transportation costs of the Chinese fleet in the shipping network [10]. A case study [11] provides a scenario analysis explanation for observed outcomes in a retrospective analysis using constrained Suez Canal case material. The findings can be utilized to diagnose backward risk sources for accident investigation and estimate forward risk for restricted waterway accident prevention to prevent similar incidents like the Suez Canal blockage. If such risks can be observed and detected very quickly with low energy consumption and little system maintenance, then precautions can be taken earlier to avoid such accidents which may have devastating consequences.

The developments experienced in the industry in recent years are evident in many sectoral areas. These developments lead to major changes in conventional structures today, and in particular, they greatly change the way work is done and the functions of employees [12]. These changes significantly affect the activities of maritime enterprises and the management of ships, as in many sectors. As a result of these developments, it has become possible to obtain a lot of data that is of great importance for the management and safe navigation of ships traveling in international waters in recent years. In particular, data regarding sea conditions are among the primary elements of safe navigation. Marine data are not only related to the sea surface; today, seabed data are as important as above-water information. For this purpose, instantaneous collection of seabed movements and data on the sea floor and timely transmission to relevant maritime organizations is of great importance for safe sea navigation. In addition, these data are necessary for safe route planning of maritime businesses and navigational routes of ships. It is also extremely important to take precautions.

In recent years, the great opportunities provided in data transmission and communication systems, in addition to ship systems, have reached dimensions that easily allow monitoring of data on the seabed as well as the sea surface from long distances and intervening when necessary. Since the target of maritime communication is the communication between the units needed by the maritime vehicles, it can be defined as communication between ships and other ships, land units, and aircraft [13].

A. Motivation

The importance of underwater wireless sensor networks (UWSNs) [14] is evident in ocean data collection, resource exploration, and navigation due to rapid development. The concept of intelligent ocean underwater Internet of Things (IoT) has been proposed recently [15], with numerous applications. Various underwater sensor nodes feed environmental data to a data processing center. In harsh marine settings, these battery-operated nodes require expensive and complicated battery replacement. Energy efficiency is essential for improving UWSN performance and reliability due to limited energy and short lifetime [16].

By proposing a robust, energy-efficient AI-based metaheuristic algorithm for path planning in UWSN, the data gathered by sensors can be collected very quickly by consuming much less energy, which enables faster response of the monitoring system in case of any risks of ship accidents. It also reduces the total energy consumption of sensors by communicating them at a closer point. It so extends the network lifetime of UWSN, which monitors the underwater environment to avoid ship accidents.

In this paper, we consider a 3D energy-aware path planning problem with autonomous underwater vehicle that visits multiple sensor nodes. This paper also considers a case (broader than obstacle avoidance) where direct traveling paths between some node pairs should be avoided due to several reasons including obstacles, underwater flows, too narrow places for travel, and other risks like changing temperature and pressure.

B. Our Contributions

Our main contributions can be summarized as follows:

- This work provides a comparative study of the five commonly used metaheuristic-based approaches (Ant Colony Optimization-based Approach, Particle Swarm Optimization-based Approach, Teaching Learning Based Optimization-based Algorithm, Genetic Algorithm-based Approach, Grey Wolf Optimizer-based Approach) for 3D path planning problem by an AUV for data collection problems in UWSN.
- We also consider the traveling limitations between some of the sensor pairs such as obstacles between sensors, pressure, water flows, and changing temperature.
- We propose a modified version of the genetic algorithm, Safety-Aware Genetic Algorithm (SAGA)-based Approach, under the traveling limitations through the links between some sensor pairs by modifying the distance-based cost matrix for the path planning problem.

C. Organization

The rest of this paper is organized as follows. Section II provides related literature. Section III provides the system model and defines the problem. Section IV tackles the 3D path planning problem as a traveling salesman problem and proposes several algorithms: Ant Colony Optimization-based Approach, Particle Swarm Optimization-based Approach, Teaching Learning Based Optimization-based Algorithm, Genetic Algorithm-based Approach, Grey Wolf Optimizer-based Approach. In Section V, we propose a novel approach, Safety-Aware Genetic Algorithm-based Approach, by considering the problem with some limitations between some of the sensor pairs. In Section VI, we evaluate performances of the proposed algorithms. Section VII concludes the paper and provides future directions.

II. Related Work

This section considers the related literature for the path planning problem in UWSN.

There have been numerous attempts to resolve this problem. First, a significant portion of the energy used by UWSNs is usually attributed to data transmission. The collected sensor data is aggregated and reduced using data compression and optimization algorithms to reduce transmission data and energy consumption [17]. Second, the energy efficiency of UWSNs can also be increased with the use of smart node placement and routing strategies. Based on the uneven distance and energy expenditure between sensor nodes, optimal deployment and routing techniques can decrease energy consumption and increase network lifetime.

However, even with these methods, changing the battery is still necessary when it runs out. Thus, to charge underwater sensors, energy transfer technologies accomplish long-term monitoring and data transmission while avoiding the difficulty of regular battery replacement. DeMauro et al. [18] created a rechargeable lithium-ion battery module specifically for underwater applications to combat high water pressure and short circuit. Autonomous underwater vehicles (AUVs) are required to help with charging due to the limited transmission distance of energy, and route planning for AUVs is required.

An AUV is a self-propelled submersible that can be used for moderate tasks without human intervention. Underwater resource research, underwater environmental monitoring, and marine safety have all made extensive use of the AUV, which is regarded as an affordable and secure tool for seabed inquiry, search, identification, and rescue [19], [20]. The difficulty of losing data from its successive nodes arises from the AUV's limited power carrying capacity, which also limits its charging area. This makes it challenging to guarantee the AUV's practicality in cases for larger detecting region, particularly in marine conditions.

A plan to create magnetically charged cars for wireless rechargeable sensor networks (WRSNs) was put forth [21]. UWSNs are often used in a three-dimensional framework, in contrast to ground-based wireless rechargeable sensor networks, and their transmission power increases significantly with underwater distance.

Communication protocol design can conserve energy due to battery limits. Lee et al. compared energy-efficient UWSN MAC techniques depending on network topology [22]. A work [23] examines energy-efficient and dependable UWSN MAC and routing techniques. Another work [24] creates a packet-sending mechanism to eliminate redundancy and increase channel quality. A hybrid-coding-aware routing system for underwater acoustic sensor networks (UASNs) by Su et al. [25] reduces transmission overhead and ensures reliability.

Underwater sensor networks benefit from clustering's energy efficiency, data aggregation, resource management, and lifespan [26]. It divides the network into clusters, each with a cluster head (CH) that aggregates and relays information from individual nodes, eliminating redundant transfers [27]. This saves energy and bandwidth in underwater areas with restricted communication resources [28].

Sun et al. [29] developed a clustering-based communication protocol that lowered sensor node energy usage. A topology management approach for underwater sonar detection networks (USDNs) by Jin et al. can improve coverage performance and extend network lifetime with guaranteed coverage and connection [30]. A work [31] developed a virtual force-based distributed node deployment strategy to improve UWSN network coverage. Another work [32] builds a network topology control model including underwater aspects like robustness, energy consumption balance, and topology to extend UWSN lifetime and optimize data delivery.

Data collection, charging, and more are conceivable with autonomous underwater vehicles. AUVs gather data. AUVs with sensors can gather data on geology underwater, water conditions, marine life. A work [33] tested AUV-assisted communication, where the AUV collects energy-saving data as a mobile node. Another work [34] suggested using AUVs to collect data and plan pathways with K-means [35].

Underwater networking and communication require AUVs. Stationary or mobile sensors can provide data to a central station or other AUVs. Smooth communication and real-time underwater operation monitoring and control are possible. A field-deployable three-phase wireless charging system by Kan et al. [36] charges AUVs quickly, efficiently, and conveniently. To speed up AUV battery life, Ramos et al. [37] used dynamic system theory for navigation in 0–100 m ocean depths.

Autonomous docking and battery charging AUVs are being developed. This lets them run for long periods without human assistance. AUV batteries and sensor nodes charge when docked. Avoiding retrieval and recharge makes them more independent and efficient.

Energy efficiency is improved via AUV path design. To save electricity and increase network lifetime, Cheng et al. worldwide design the AUV's path, avoid underwater obstacles, and analyze its energy consumption model using kinematic and dynamical models [38]. Kumar et al. [39] propose a hybrid subsea AUV exploration method that greatly reduces their range. The work [40] sectors the exploring region into numerous smaller sections with data-receiving points. Path planning saves AUV energy while collecting data. A rechargeable UWSN path planning method [41] increases network lifetime.

To solve UWSN energy shortages and battery replacement difficulties, the work [42] proposes a path planning and energy-saving technique for charging underwater sensor nodes using AUVs. A genetic algorithm determines the optimal AUV path, while many AUVs charging the sensor network nodes maximize network size and transmission reliability. The outcomes of the simulation demonstrate that the AUV path planning scheme converges more quickly than conventional algorithms and increases the lifetime of UWSNs while energy balancing following node density and network size. In high-density networks, the proposed path planning technique lowers the energy consumption of exploratory AUVs by 15% per AUV.

A work [43] considers a path planning problem of unmanned aerial vehicle (UAV) from one point to another point by avoiding obstacles between them and it presents a comparative study of genetic algorithm, simulated annealing, grey wolf optimizer, and an improved version of grey wolf optimizer algorithm. However, it tackles a problem like the shortest path problem while our paper considers a problem like a traveling salesman problem (TSP).

Another research [44] tackles a path planning problem of AUV from one point to another point by considering ocean currents. It tackles the problem for both cases without obstacles between them and with obstacles between them. It presents a comparative study of A* [45], rapidly exploring random tree (RRT) [46], [47], genetic algorithm, particle swarm optimization, and an improved version of particle swarm optimization algorithm. However, it tackles a problem like the shortest path problem in an ocean environment with ocean currents while our paper considers a problem like a TSP.

A study [48] considers a motion planning problem of an autonomous ground vehicle from one point to another point by avoiding obstacles between them and it presents a comparative study of the probabilistic roadmap (PRM) [49], RRT, and the proposed algorithm, Optimistic Motion Planning using Recursive Sub-Sampling. The investigated problem is a 2D motion planning problem which differs from our TSP-type path planning problem.

Another work [50] proposes a new optimal path planning method for long-term autonomous underwater vehicle operations in areas where ocean currents change over time. These currents may surpass the AUV's top speed and momentarily reveal obstructions. Paths require both geographical and temporal characterisation, in contrast to the majority of other path design methodologies. This method allows for a trade-off between mission duration and energy requirements by utilising ocean currents to limit energy usage and achieve mission objectives. By using a parallel swarm search, the proposed method reduces the susceptibility to large local minima on the complex cost surface. The efficiency of the optimisation strategies is evaluated computationally and empirically using the Starbug AUV on a validated ocean model of Brisbane's Moreton Bay.

In another research [51], the genetic algorithm, grey wolf optimizer algorithm and nearest neighbour algorithm have been applied to solve this problem. It is shown that the nearest neighbour algorithm shows much quicker (nearly 30 times quicker) performance than the genetic algorithm and grey wolf optimizer algorithmn. On the other hand, the genetic algorithm exhibits better performance than the nearest neighbour algorithm while grey wolf optimizer algorithm demonstrates the worst performance among all them.

This present paper considers a 3D energy-aware path planning problem with autonomous underwater vehicle that visits multiple nodes. Then, it applies the five most commonly used metaheuristic-based approaches, namely, Ant Colony Optimization-based Approach, Particle Swarm Optimization-based Approach, Teaching Learning-based Optimization-based Approach, Genetic Algorithm-based Approach, Grey Wolf Optimizer-based Approach. This paper also considers a case (broader than obstacle avoidance) where direct traveling paths between some node pairs should be avoided due to several reasons including obstacles, underwater flows, too narrow places for the travel, and other risks like changing temperature and pressure.

Table I provides a brief comparison of the energy-aware path planning approaches in the closely related literature.

TABLE I. Brief Comparison of the Energy-aware Path Planning Approaches

Included Features	Underwater	3D	Multiple Point visit	Obstacle Avoidance
Ding *et al.* [43]	no	no	no	yes
Zeng *et al.* [44]	yes	no	no	yes
Kenye *et al.* [48]	no	no	no	yes
Witt *et al.* [50]	no	no	no	no
Gul *et al.* [51]	yes	yes	yes	no
This work	yes	yes	yes	yes

III. System Model and Problem Definition

This paper considers the energy-aware path planning problem with an AUV to visit sensors underwater. This section presents a motivating scenario and formulates the problem based on this motivation. First, we consider the system model of the UWSN. Then, we define the energy-aware path planning problem more precisely.

A. System Model

The network model is shown in Fig. 1. Every sensor node is connected to an underwater acoustics link, which transmits data to SINK nodes. Starting at a charge station (CS), the magnetic resonance coupling AUV charges each sensor node before making its way back to the CS for a charge and rest. It functions as a mobile sink to gather data.

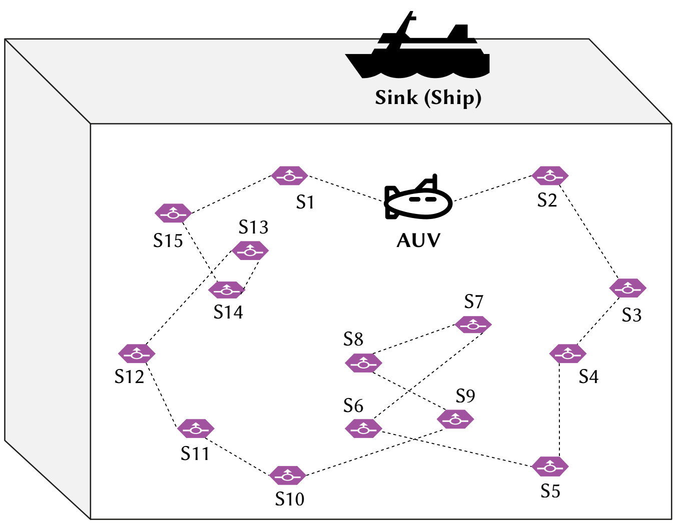


Fig. 1. The system model of the UWSN where the AUV collects data from all the fifteen sensors (N1, N2, ..., N15) which monitor the sea for anomaly/risk detection to avoid ship accidents. After aggregating all the gathered data, this is sent to the data sink which collects all the data to evaluate them.

One important problem in UWSNs is the energy consumption balance of underwater sensors. In several studies [34, 35, 36], AUVs have been used to collect underwater data to solve the problem of unequal energy use. The AUV moves and visits each sensor node by a predefined plan to balance each node's energy consumption.

B. Problem Definition

The energy-aware path planning problem via AUV can be categorized as a traveling salesman problem (TSP) [52]–[54]. The two main methods for resolving the TSP are the intelligent evolutionary algorithm and the classical search algorithm. Examples of the former include the greedy algorithm, the artificial potential field technique, and the quick progress algorithm. The latter includes methods like the ant colony algorithm, particle swarm optimization approach, teaching learning-based optimization algorithm, grey wolf optimizer algorithm, and genetic algorithm.

The most prominent NP-hard optimization problem is the TSP [53], [54]. TSP finds a route for a salesman that starts from home, visits a collection of locations, and returns to the original place with the minimum trip distance with each city visited once [56].

In a TSP problem with m sensor nodes, let cij denote the node distance from node i to node j. Let xij denote a binary variable that takes the value of 1 if node j is visited just after node i. Otherwise, it takes the value of 0. In this case, the energy-aware path planning problem can be considered as an NP-hard TSP as follows [55]:

Problem 1. *Minimizing the following cost function:*

$$\min_{x_{ij}} \sum_{j=1}^{m} \sum_{i=1}^{m} c_{ij}\, x_{ij} \tag{1}$$

where

$$\begin{aligned} \sum_{i=1}^{m} x_{ij} &= 1, j = 1, \ldots, m \\ \sum_{j=1}^{m} x_{ij} &= 1, i = 1, \ldots, m \\ \sum_{i \in K} \sum_{j \in K} x_{ij} &\leq |K| - 1, \forall K \subset \{1, \ldots, m\} \end{aligned}$$

IV. Proposed Energy-Aware Path Planning (EAPP) Approaches

In this section, we tackle the energy-aware path planning problem of AUV which includes the distance between each pair of sensor nodes.

We present the following algorithms by tackling the EAPP problem as a TSP problem: Ant Colony Optimization-based Approach, Particle Swarm Optimization-based Approach, Teaching Learning based Optimization-based Algorithm, Genetic Algorithm-based Approach, Grey Wolf Optimizer-based Approach.

A. Ant Colony Optimization (ACO)-Based Approach

ACO has many inherent limitations, despite its strong performance in discrete problem solutions. Despite having great stability, it has several disadvantages when working with large amounts of data in terms of convergence speed and results in correctness [59].

We tackle the EAPP problem as a TSP and propose a 3D path planning solution based on the ACO [57], [58].

B. Particle Swarm Optimization (PSO) Algorithm-Based Approach

PSO has been popular with researchers for its ability to hybridize, specialize, and exhibit novel emergent behaviors in various application areas. PSO's main benefit is tweaking fewer parameters. PSO finds the optimal particle interaction solution but converges slowly to the global optimum in high-dimensional search space. It performs poorly on large, complex datasets. PSO rarely finds the global optimum solution in multidimensional situations. Local optima traps and particle velocity variations confine trials to a sub-plain of the search hyper-plain [62], [63].

We tackle the EAPP problem as a TSP and propose a 3D path planning solution based on the PSO [60], [61].

C. Teaching Learning Based Optimization (TLBO)-Based Algorithm

TLBO solves large global optimum optimization problems with a sophisticated metaheuristic method. Several TLBO variations have been proposed to improve local optima avoidance and convergence speed [65].

We tackle the EAPP problem as a TSP and propose a 3D path planning solution based on the TLBO-based Algorithm [64].

D. Grey Wolf Optimizer (GWO)-Based Approach

GWO is easier to implement than PSO and GA. On the other hand, it has the drawbacks of poor convergence speed, low solution precision, and local optimum tendency.

We tackle the EAPP problem as a TSP and propose a 3D path planning solution based on the GWO algorithm [66].

E. Genetic Algorithm (GA)-Based Approach

We tackle the EAPP problem as a TSP and propose a 3D path planning solution based on the Genetic Algorithm [67],[68].

The basic principle of genetic algorithms is to solve complex optimization problems by imitating biological evolution. The first steps in applying a genetic algorithm to tackle TSP issues are identifying the individuals of the TSP solution and initializing the population. Every member of the population is rated according to a fitness function, and the most fit individuals are selected for genetic processes including selection, crossover, and mutation. The genetic algorithm's termination criterion is the maximum number of iterations selected. Furthermore, the individual fitness for this work is the total route size or the total AUV energy consumption. By summing up the distances of all the sensing nodes, equation (2) may be utilized to determine the fitness of each individual in this circumstance.

$$Fitness = \sum_{l=1}^{N-1} \frac{1}{\sqrt{(x_l - x_{l-1})^2 + (y_l - y_{l-1})^2 + (z_l - z_{l-1})^2}} \tag{2}$$

where N denotes the number of nodes; (x_l, y_l, z_l) denotes the 3D position of node l.

V. Safety-Aware Genetic Algorithm (SAGA)-Based Approach

In this section, we consider the energy-aware path planning problem by also considering the limitations that emerge between some of the sensor pairs. The obstacles between sensors can block direct traveling from one sensor to the other. Changing pressure, water flows and changing temperature can be other reasons for the AUV not prefer to travel from one sensor to the other sensor directly. In this case, the AUV will visit some other sensor/s between those two sensors.

We propose a modified version of the genetic algorithm, the Safety-Aware Genetic Algorithm (SAGA)-based Approach, for the 3D path planning problem with small obstacles that emerge between some of the sensor pairs.

In the SAGA approach, we do not modify the standard GA itself; however, we transform the distance cost matrix by replacing the cost of the unavailable path between some nodes with a very large number of M to avoid preferring those paths during path planning.

Fig. 2 shows the flow diagram of SAGA, which exhibits its difference from GA.

Before applying the genetic algorithm, the distance cost matrix obtained for n nodes can be written as

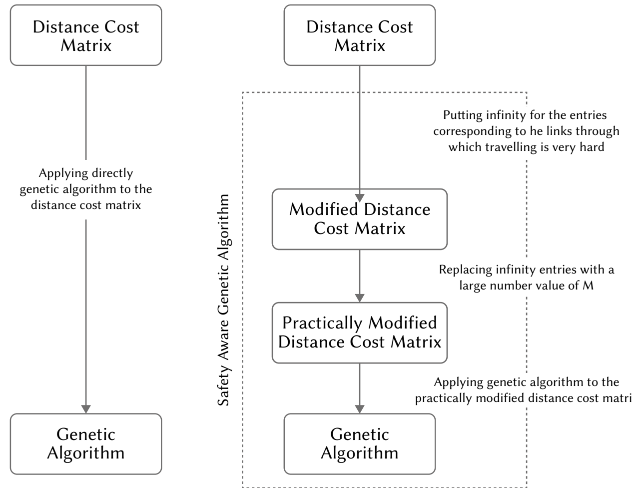


Fig. 2. The flow charts of GA (left one) and SAGA (right one), which exhibits the difference of SAGA from GA.

$$D = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1(n-1)} & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2(n-1)} & d_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ d_{(n-1)1} & d_{(n-1)2} & \dots & d_{(n-1)(n-1)} & d_{(n-1)n} \\ d_{n1} & d_{n2} & \dots & d_{n(n-1)} & d_{nn} \end{bmatrix} \tag{3}$$

Due to several reasons including underwater flows, too narrow places for the UAV's travel, and other risks like changing temperature and pressure, direct traveling paths between some node pairs should be avoided in some cases. In such cases, if we consider blockage between node $(n-1)$ and node n such that traveling from node $(n-1)$ to node n is not possible, then their distance can be modified as $d_{(n-1)n} = \infty$.

By replacing all $d_{(n-1)n}$ entries $(d_{12}, d_{23}, \dots, d_{(n-1)n})$ with ∞ such that $d_{(n-1)n} = \infty$, the modified distance cost matrix D_{mod} obtained before applying the genetic algorithm for n nodes can be written as

$$D_{mod} = \begin{bmatrix} d_{11} & \infty & \dots & d_{1(n-1)} & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2(n-1)} & d_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ d_{(n-1)1} & d_{(n-1)2} & \dots & d_{(n-1)(n-1)} & \infty \\ d_{n1} & d_{n2} & \dots & d_{n(n-1)} & d_{nn} \end{bmatrix} \tag{4}$$

Giving a very large number M instead of ∞ can be more practical for the implementation. M can be chosen as the square of the maximum distance between two nodes in the matrix.

By replacing all ∞ with M such that $d_{(n-1)n} = M$, the practically modified distance cost matrix D_{mod}^{prac} obtained before applying genetic algorithm for n nodes can be written as

$$D_{mod}^{prac} = \begin{bmatrix} d_{11} & \mathbf{M} & \dots & d_{1(n-1)} & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2(n-1)} & d_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ d_{(n-1)1} & d_{(n-1)2} & \dots & d_{(n-1)(n-1)} & \mathbf{M} \\ d_{n1} & d_{n2} & \dots & d_{n(n-1)} & d_{nn} \end{bmatrix} \tag{5}$$

By considering the practically modified distance cost matrix D_{mod}^{prac} instead of the distance cost matrix D, we apply the 3D genetic algorithm, which brings safety-awareness about each link between node i and node $i+1$. Hence, we propose the SAGA-based Approach, for the 3D energy-aware path planning problem.

To sum up the implementation of SAGA, we applied the GA in the 3D path planning problem with the difference that we modified the distance cost matrix before applying the GA, which brings safety awareness to GA and converts it into SAGA.

VI. Numerical Results

In this section, we evaluate the performance of the algorithms for the 3D energy-aware path planning problem of AUV which includes the distance between each pair of sensor nodes. For the simulations, we formed a 500 m × 500 m × 500 m space by locating sensor nodes randomly (The related works choose similar range of dimension length and distances).

In the first subsection, we consider a scenario with an AUV and 50 nodes by considering no limits in links that block traveling directly between some node pairs. In the second subsection, we consider a scenario with an AUV and 100 nodes by considering no limits in links that block traveling directly between some node pairs. In the last subsection, we consider two separate scenarios with an AUV and 50 nodes and with an AUV and 100 nodes by considering limits in some links that block traveling directly between some node pairs.

A. 50-Node Scenario

In this subsection, we will consider a scenario with 50 nodes and a single AUV. Fig. 3 illustrates the locations of the 50 nodes in the 500 m × 500 m × 500 m space.

Locations of the 50 nodes are given as { (440, 20, 472), (500, 57, 325), (283, 289, 65), (292, 309, 106), (55, 147, 388), (452, 20, 45), (236, 381, 178), (423, 9, 16), (157, 461, 36), (75, 28, 450), (141, 72, 153), (290, 354, 346), (2, 213, 111), (146, 175, 125), (156, 207, 283), (447,261, 360), (106, 139, 228),(448, 482, 170),(426, 110, 221), (487, 279, 21), (257,240, 184), (323, 160, 225), (139, 20, 384), (294, 106, 400), (412, 353, 409), (437, 60, 12), (239, 5, 27), (256, 270, 219), (317, 388, 242), (207, 297, 388), (167,154, 441), (148, 463, 94), (185, 103, 278), (270, 445, 26), (346, 261, 303), (387,232, 380), (397, 414, 211),(29, 368, 61), (150, 118, 369), (205, 65, 489), (116,350, 124), (223, 359, 458),(458, 201, 137), (13, 98, 242), (338, 186, 18), (58, 445, 314), (428, 152, 166),(212, 156, 281), (208, 330, 93),(100, 333, 420) }

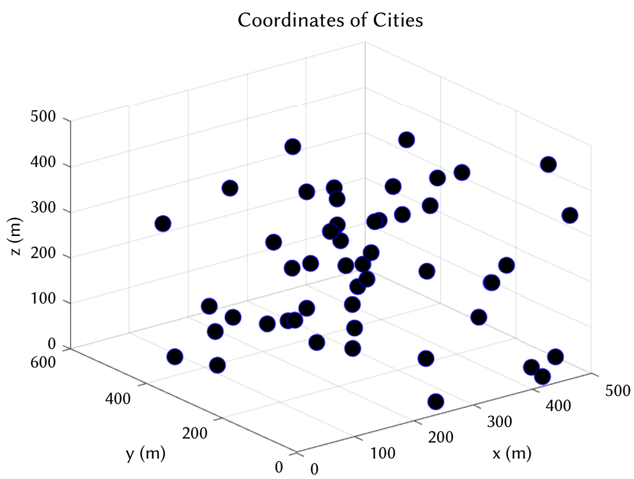


Fig. 3. The coordinates of the sensor nodes to be visited by the AUV.

By considering different subsets of these parameters, we evaluate the performance of the following algorithms by tackling the EAPP problem as a TSP problem: Ant Colony Optimization (ACO)-based Approach, Particle Swarm Optimization (PSO)-based Approach, Teaching Learning Based Optimization (TLBO)-based Algorithm, Genetic Algorithm (GA)-based Approach, Grey Wolf Optimizer (GWO)-based Approach.

In the following subsubsections, we present the solutions achieved by ACO-based Approach, PSO-based Approach, TLBO-based Algorithm, GA-based Approach, and GWO-based Approach as a result of 1000 iterations.

1. ACO-Based Approach

In this subsubsection, we present an ACO-based solution for the 3D TSP problem.

Fig. 4 exhibits ACO's achieved path planning solution in 1000 iterations for visiting the 50 nodes in Fig. 3.

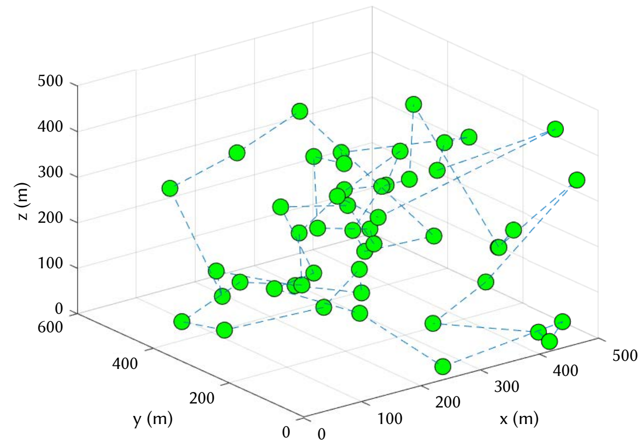


Fig. 4. The achieved path planning solution for visiting the 100 nodes by AUV with ACO in 1000 iterations.

2. PSO-Based Approach

In this subsubsection, we present a PSO-based solution for the 3D TSP problem.

Fig. 5 exhibits the PSO's achieved path planning solution in 1000 iterations for visiting the 50 nodes in Fig. 3.

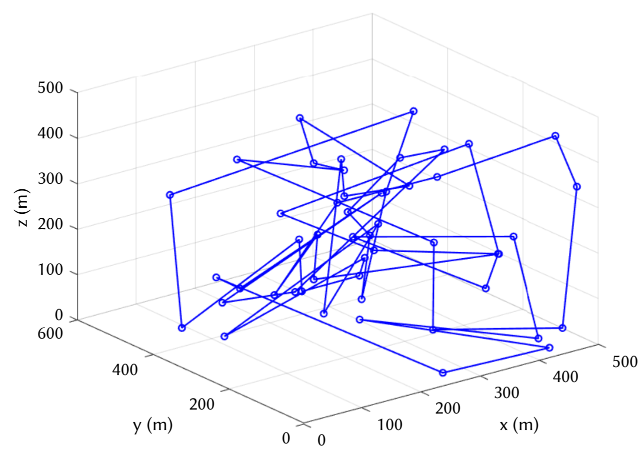


Fig. 5. The achieved path planning solution for visiting the 100 nodes by AUV with PSO in 1000 iterations.

3. TLBO-Based Algorithm

In this subsubsection, we present a TLBO-based solution for the 3D TSP problem.

Fig. 6 exhibits the TLBO's achieved path planning solution in 1000 iterations for visiting the 50 nodes in Fig. 3.

4. GWO-Based Approach

In this subsubsection, we present a GWO-based solution for the 3D TSP problem.

Fig. 7 exhibits the GWO's achieved path planning solution in 1000 iterations for visiting the 50 nodes in Fig. 3.

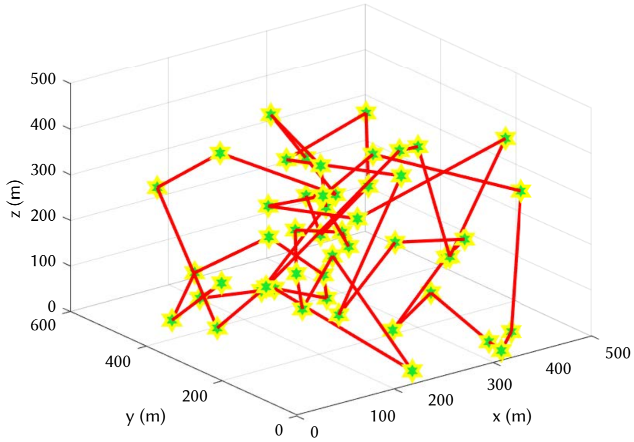


Fig. 6. The achieved path planning solution for visiting the 100 nodes by AUV with TLBO in 1000 iterations.

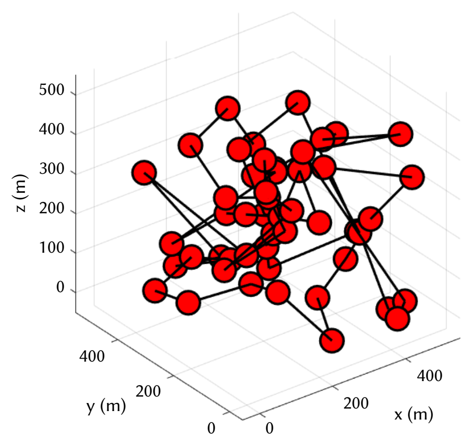


Fig. 7. The achieved path planning solution for visiting the 100 nodes by AUV with GWO in 1000 iterations.

5. GA-Based Approach

In this subsubsection, we present a GA-based solution for the 3D TSP problem.

Fig. 8 exhibits the GA's achieved path planning solution in 1000 iterations for visiting the 50 nodes in Fig. 3.

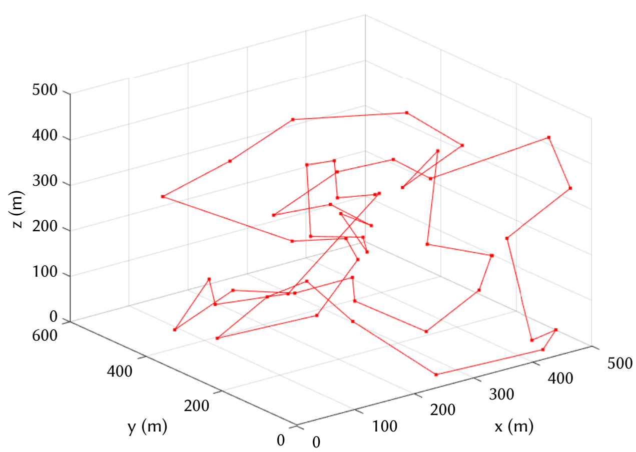


Fig. 8. The achieved path planning solution for visiting the 100 nodes by AUV with GA in 1000 iterations.

6. Comparison and Discussion

Considering the general trend, the GA-based Approach shows better performance than the ACO-based Approach, PSO-based Approach, TLBO-based Approach, and GWO-based Approach.

Fig. 9 shows the total traveled distance by AUV with different algorithms (the ACO-based Approach, PSO-based Approach, TLBO-based Approach, GWO-based Approach, and GA-based Approach) for visiting the 50 nodes, which are located initially as given in Fig. 3.

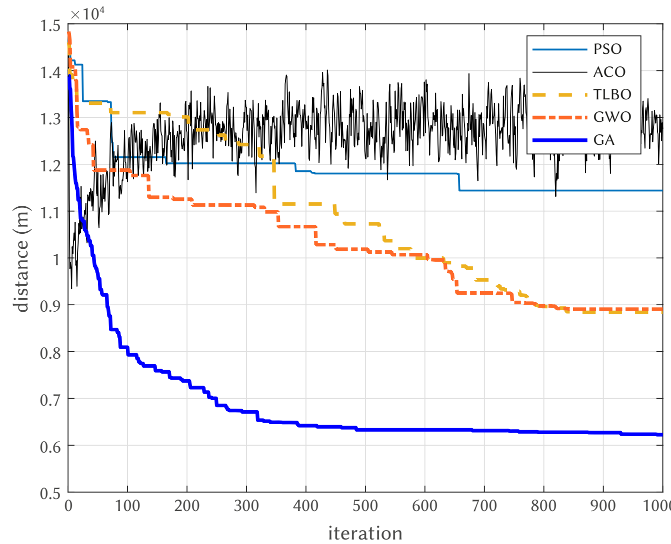


Fig. 9. The achieved path lengths for visiting the 100 nodes by all the algorithms (the ACO-based Approach, PSO-based Approach, TLBO-based Approach, GWO-based Approach, and GA-based Approach) in 1000 iterations.

From Fig. 9, we can make the following observations on the performance of the algorithms for the 50-node scenario. The ACO-based approach achieves better in less number of iterations such that its performance becomes worse with the increasing number of iterations after achieving its minimum. Although the PSO-based approach achieves better with an increasing number of iterations, it achieves better than just the ACO-based approach for 1000 iterations. Although the GWO-based approach converges faster than TLBO-based approach, both achieve almost the same performance for 1000 iterations, which is considerably better than the ACO-based approach and PSO-based approach. The GA-based approach not only achieves much better than all of ACO-based Approach, PSO-based Approach, TLBO-based Approach, GWO-based Approach but also converges faster than PSO-based Approach, TLBO-based Approach, GWO-based Approach (Only ACO-based approach converges to minimum very fast.)

From Table II, we can make the following observations. At the beginning (in the first iteration), all the algorithms except the GA-based approach have similar performance with at most 2.5% difference (366 m difference between ACO-based approach and GWO-based approach) while the GA-based approach achieves 4.0% difference better than ACO-based approach. In iteration 100, the TLBO-based approach achieves the worst performance while the GA-based approach achieves considerably better than the other approaches (3534 m, namely, 30.4% less than the ACO-based approach which is the second best). In iteration 300, the ACO-based approach achieves the worst performance while the GA-based approach achieves considerably better than the other approaches (4418 m, namely, 39.7% less than the GWO-based approach which is the second best). In addition, PSO and TLBO-based approaches achieve closely to each other in iteration 300. In iteration 600, the ACO-based approach achieves the worst performance while the GA-based approach achieves considerably

TABLE II. Total Distance for Visiting the 50 Nodes By the Algorithms (the ACO-Based Approach, PSO-Based Approach, TLBO-Based Approach, GWO-Based Approach, and GA-Based Approach) With Respect to Iteration Number (Note That Iteration 1 Is Considered as the Beginning Instead of Iteration 0)

Iteration	1	100	200	300	400	500	600	700	800	900	1000
ACO	14468	11628	12683	13214	12748	12130	13177	12652	12936	12981	12502
PSO	14787	12150	12019	12019	11851	11802	11802	11439	11439	11439	11439
TLBO	14603	13103	13012	12421	11153	10729	10069	9533	8963	8906	8906
GWO	14834	11874	11254	11130	10669	10186	9996	9251	8977	8837	8837
GA	13889	8094	7374	6712	6420	6330	6330	6314	6279	6271	6227

better than the other approaches (3666 m, namely, 36.7% less than the GWO-based approach which is the second best). In addition, GWO and TLBO-based approaches achieve very closely to each other in iteration 600 (just 0.73% difference). In iteration 1000, the ACO-based approach achieves the worst performance while the GA-based approach achieves considerably better than the other approaches (2610 m, namely, 29.5% less than the GWO-based approach which is the second best). In addition, GWO and TLBO-based approaches achieve very closely to each other in iteration 1000 (just 0.77% difference), which is much better than the ACO-based approach and PSO-based approach.

B. 100-Node Scenario

In this subsection, we will consider a scenario with 100 nodes and a single AUV. Fig. 10 illustrates the locations of the 100 nodes in the 500 m × 500 m × 500 m space.

Locations of the 100 nodes are given as { (408, 82, 323), (453, 398, 190), (64, 156, 406), (457, 265, 267), (317, 83, 176), (49, 301, 470), (140, 132, 438), (274, 328, 276), (479, 345, 312), (483, 375, 294), (79, 226, 104), (486, 42, 151), (479, 115, 236), (243, 457, 116), (401, 77, 423), (71, 413, 98), (211, 270, 113), (458, 499, 86), (397, 40, 114), (480, 222, 218), (328, 54, 156), (18, 481, 462), (425, 3, 216), (467, 388, 93), (340, 409, 453), (379, 435, 490), (372, 43, 220), (197, 200, 56), (328, 130, 130), (86, 401, 205), (354, 216, 298), (16, 456, 132), (139, 91, 302), (24, 132, 356), (49, 73, 111), (412, 69, 59), (348, 435, 149), (159, 290, 160), (476, 275, 213), (18, 73, 254), (220, 427, 43), (191, 312, 132), (383, 176, 401), (398, 257, 15), (94, 201, 465), (245, 38, 366), (223, 120, 245), (324, 62, 290), (355, 92, 119), (378, 120, 230), (139, 209, 482), (340, 25, 274), (328, 452, 261), (82, 473, 116), (60, 246, 245), (250, 245, 313), (480 169, 340), (171, 451, 198), (293, 185, 184), (112, 56, 494), (376, 391, 19), (128, 195, 443), (253, 121, 457), (350, 202, 399), (446, 49, 50), (480, 66, 131), (274, 472, 168), (70, 479, 340), (75, 288, 69), (129, 30, 361), (421, 118, 54), (128, 177, 327), (408, 411, 248), (122, 8, 390), (465, 22, 358), (175, 85, 452), (99, 325, 446), (126, 366, 168), (309, 324, 350), (237, 226, 99), (176, 274, 16), (416, 149, 373), (293, 373, 251), (275, 95, 240), (459, 344, 453), (143, 92, 305), (379, 185, 309), (377, 313, 430), (191, 391, 403), (284, 41, 289), (38, 465, 92), (27, 388, 120), (266, 244, 444), (390, 218, 15), (468, 224, 245), (65, 154, 84), (285, 255, 490), (235, 256, 357), (6, 409, 251), (169, 398, 236)}.

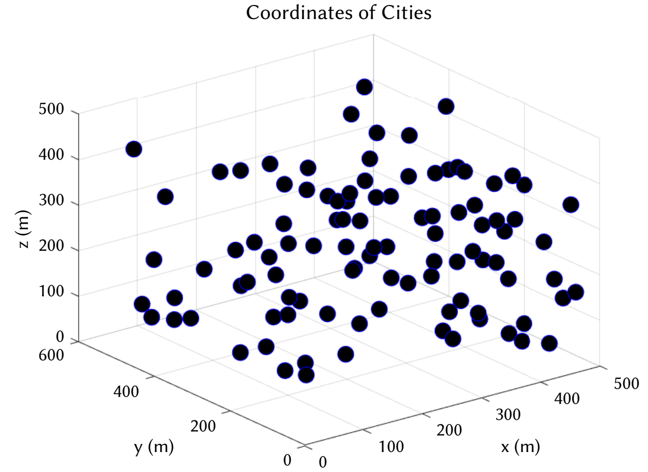


Fig. 10. The coordinates of the 100 sensor nodes to be visited by the AUV.

By considering different subsets of these parameters, we evaluate the performance of the following algorithms by tackling the EAPP problem as a TSP problem: Ant Colony Optimization (ACO)-based Approach, Particle Swarm Optimization (PSO)-based Approach, Teaching Learning Based Optimization (TLBO)-based Algorithm, Genetic Algorithm (GA)-based Approach, Grey Wolf Optimizer (GWO)-based Approach.

In the following subsubsections, we present the solutions achieved by ACO-based Approach, PSO-based Approach, TLBO-based Algorithm, GA-based Approach, and GWO-based Approach as a result of 1000 iterations.

1. ACO-Based Approach

In this subsubsection, we present an ACO-based solution for the 3D TSP problem.

Fig. 11 exhibits ACO's achieved path planning solution in 1000 iterations for visiting the 100 nodes in Fig. 10.

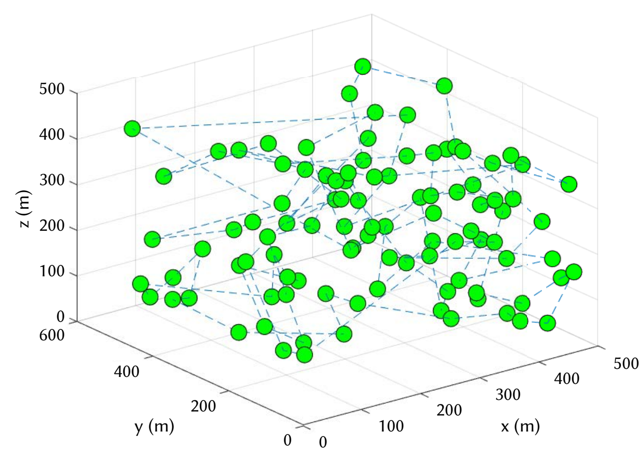


Fig. 11. The achieved path planning solution for visiting the 100 nodes by AUV with ACO in 1000 iterations.

2. PSO-Based Approach

In this subsubsection, we present a PSO-based solution for the 3D TSP problem.

Fig.12 exhibits the PSO's achieved path planning solution in 1000 iterations for visiting the 100 nodes in Fig. 10.

3. TLBO-Based Algorithm

In this subsubsection, we present a TLBO-based solution for the 3D TSP problem.

Fig. 13 exhibits the TLBO's achieved path planning solution in 1000 iterations for visiting the 100 nodes in Fig. 10.

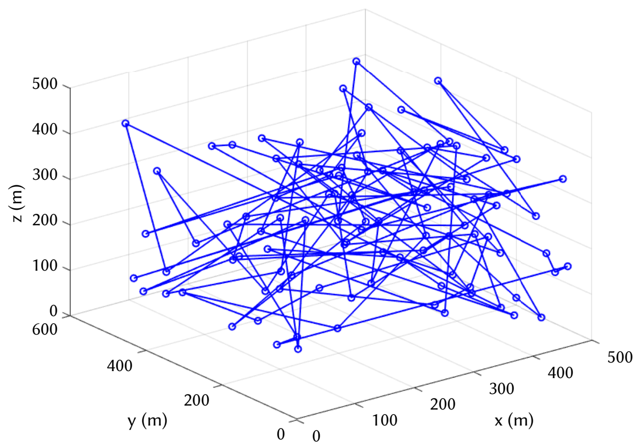


Fig. 12. The achieved path planning solution for visiting the 100 nodes by AUV with PSO in 1000 iterations.

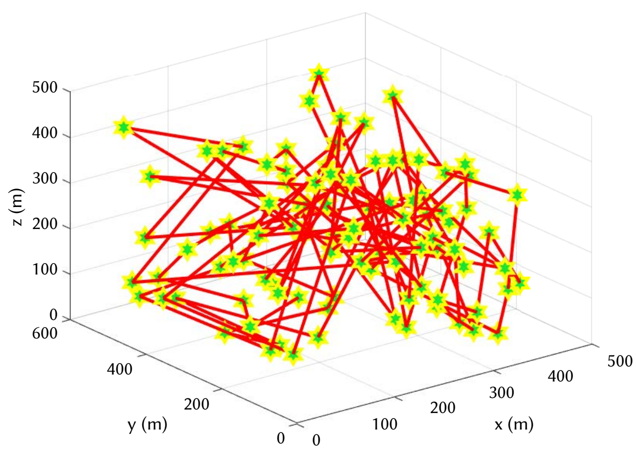


Fig. 13. The achieved path planning solution for visiting the 100 nodes by AUV with TLBO in 1000 iterations.

4. GWO-Based Approach

In this subsubsection, we present a GWO-based solution for the 3D TSP problem.

Fig. 14 exhibits the GWO's achieved path planning solution in 1000 iterations for visiting the 100 nodes in Fig. 10.

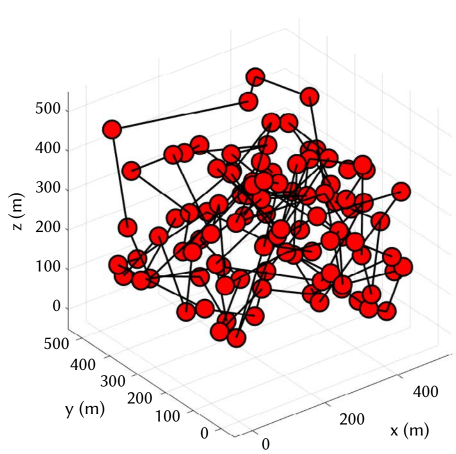


Fig. 14. The achieved path planning solution for visiting the 100 nodes by AUV with GWO in 1000 iterations.

5. GA-Based Approach

In this subsubsection, we present a GA-based solution for the 3D TSP problem. Fig. 15 exhibits the GA's achieved path planning solution in 1000 iterations for visiting the 100 nodes in Fig. 10.

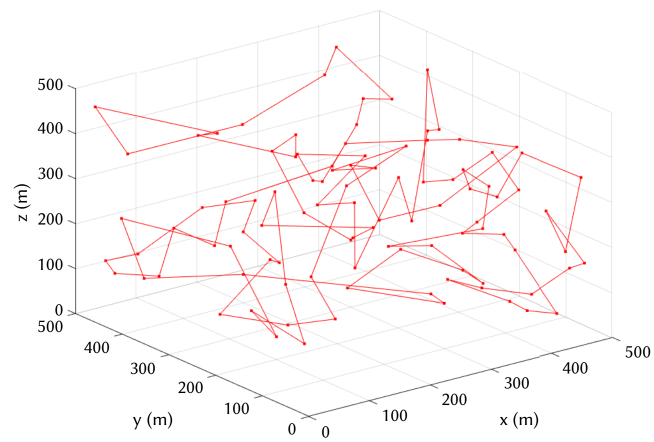


Fig. 15. The achieved path planning solution for visiting the 100 nodes by AUV with GA in 1000 iterations.

6. Comparison and Discussion

Considering the general trend, the GA-based Approach shows better performance than the ACO-based Approach, PSO-based Approach, TLBO-based Approach, and GWO-based Approach.

Fig. 16 shows total traveled distance by AUV with different algorithms (ACO-based Approach, PSO-based Approach, TLBO-based Approach, GWO-based Approach, and GA-based Approach) for visiting the 100 nodes, which are located as given in Fig. 10.

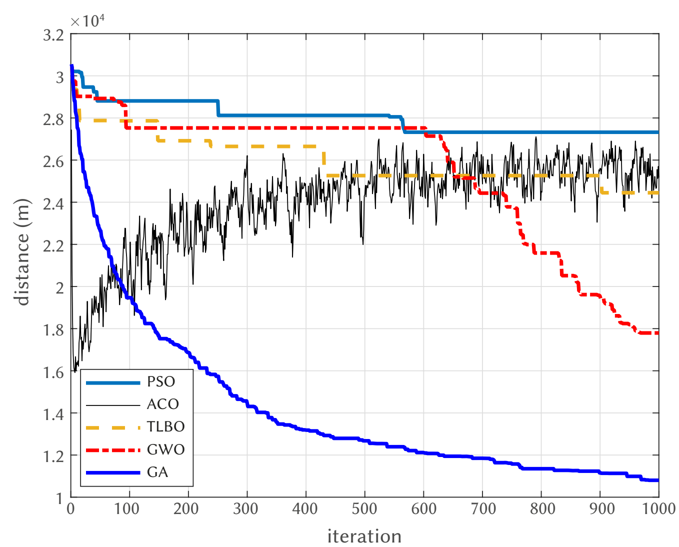


Fig. 16. The achieved path lengths for visiting the 100 nodes by all the algorithms (the ACO-based Approach, PSO-based Approach, TLBO-based Approach, GWO-based Approach, and GA-based Approach) in 1000 iterations.

From Fig. 16, we can make the following observations on the performance of the algorithms for the 100-node scenario. Although the PSO-based approach achieves better with an increasing number of iterations, it achieves worse than all the other approaches for 1000 iterations. The ACO-based approach achieves better in less number of iterations such that its performance becomes worse with the increasing number of iterations after achieving its minimum. The GWO-based

TABLE III. Total Distance for Visiting the 100 Nodes By the Algorithms (ACO-Based Approach, PSO-Based Approach, TLBO-Based Approach, GWO-Based Approach, and GA-Based Approach) With Respect to Iteration Number (Note That Iteration 1 Is Considered as the Beginning Instead of Iteration 0)

Iteration	1	100	200	300	400	500	600	700	800	900	1000
ACO	27446	20564	21526	26213	24262	24379	13177	12652	12936	12981	12502
PSO	30230	28808	28808	28120	28120	28120	27323	27323	27323	27323	27323
TLBO	29366	27864	26917	26649	26649	25264	25264	25264	25264	25264	24448
GWO	30418	27525	27525	27525	27525	27525	27525	24432	21589	19529	17794
GA	30549	19471	16861	14455	13182	12676	12120	11848	11351	11159	10803

approach converges almost as fast as the TLBO-based approach in the first 600 iterations while it converges to its minimum faster than the TLBO-based approach. The TLBO-based approach achieves almost the same as the ACO-based approach whereas the GWO-based approach achieves considerably better than all of the TLBO-based approach, the ACO-based approach and PSO-based approach. The GA-based approach not only achieves much better than all of the ACO-based Approach, PSO-based Approach, TLBO-based Approach, GWO-based Approach but also converges faster than PSO-based Approach, TLBO-based Approach, GWO-based Approach (Only ACO-based approach converges to its minimum very fast.) Considering all these, increasing the number of nodes from 50 to 100 nodes makes a considerable difference in some of these algorithms, especially the PSO-based approach and TLBO-approach.

From Table III, we can make the following observations. At the beginning (in the first iteration), all of the algorithms except the GA-based approach have similar performance with at most 2.5% difference (366 m difference between ACO-based approach and GWO-based approach) while the GA-based approach achieves 4.0% difference better than ACO-based approach. In iteration 100, the TLBO-based approach achieves the worst performance while the GA-based approach achieves considerably better than the other approaches (3534 m, namely, 30.4% less than the ACO-based approach which is the second best). In iteration 300, the ACO-based approach achieves the worst performance while the GA-based approach achieves considerably better than the other approaches (4418 m, namely, 39.7% less than the GWO-based approach which is the second best). In addition, PSO and TLBO-based approaches achieve closely to each other in iteration 300. In iteration 600, the ACO-based approach achieves the worst performance while the GA-based approach achieves considerably better than the other approaches (3666 m, namely, 36.7% less than the GWO-based approach which is the second best). In addition, GWO and TLBO-based approaches achieve very closely to each other in iteration 600 (just 0.73% difference). In iteration 1000, the ACO-based approach achieves the worst performance while the GA-based approach achieves considerably better than the other approaches (2610 m, namely, 29.5% less than the GWO-based approach which is the second best). In addition, GWO and TLBO-based approaches achieve very closely to each other in iteration 1000 (just 0.77%difference), which is much better than the ACO-based approach and PSO-based approach.

C. Safety-Awareness in 50-Node Scenario and 100-Node Scenario

In this subsection, we evaluate the performance of the Safety-Aware Genetic Algorithm (SAGA)-based solution for the 3D TSP problem under the limitation where visiting node i just after node $i-1$ has an extreme distance cost so impossible to visit.

1. Safety-Aware Genetic Algorithm (SAGA)-Based Approach

In this subsubsection, we observe the Safety-Aware Genetic Algorithm (SAGA)-based solution for the 3D TSP problem. Fig. 17 exhibits the SAGA's achieved path planning solution in 1000 iterations for visiting the 50 nodes in Fig. 3.

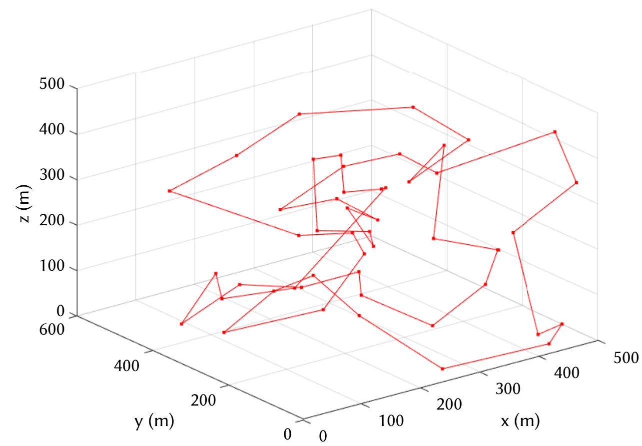


Fig. 17. The achieved path planning solution for visiting the 50 nodes by AUV with SAGA in 1000 iterations under limitations.

Fig. 18 exhibits the SAGA's achieved path planning solution in 1000 iterations for visiting the 100 nodes in Fig. 10.

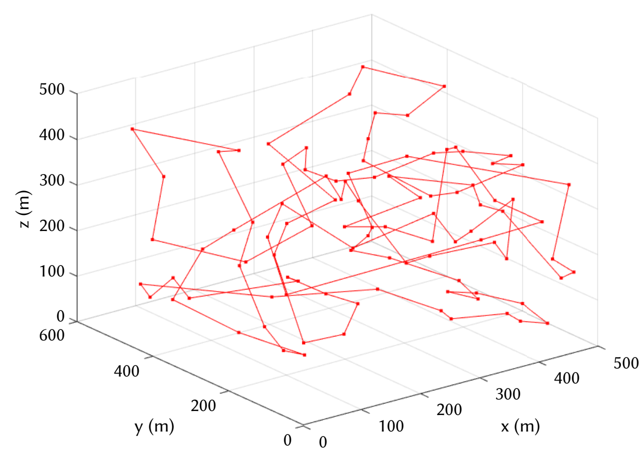


Fig. 18. The achieved path planning solution for visiting the 100 nodes by AUV with SAGA in 1000 iterations under limitations.

2. Comparison and Discussion

Considering the general trend, the GA-based Approach shows better performance than SAGA-based Approach. Fig. 19 shows the total traveled distance by AUV with GA-based Approach and SAGA-based Approach under the 50-node scenario in Fig. 3 and the 100-node scenario in Fig. 10.

From Fig. 19, we can make the following observations on the performance of the GA-based Approach and SAGA-based Approach under the 50-node scenario and the 100-node scenario. Under the 50-node scenario, the GA-based Approach and SAGA-based Approach achieve almost the same performance with just a slight difference.

TABLE IV. Total Distance for Visiting the Nodes By the GA-Based Approach and the SAGA-Based Approach Under 50-Node Scenario and 100-Node Scenario With Respect to Iteration Number (Note That Iteration 1 Is Considered as the Beginning Instead of Iteration 0)

Iteration	1	100	200	300	400	500	600	700	800	900	1000
GA with 50 node	13889	8094	7374	6712	6420	6330	6330	6314	6279	6271	6227
SAGA with 50 node	14999	8604	7392	7197	6964	6738	6558	6524	6491	6491	6404
GA with 100 node	30549	19471	16861	14455	13182	12676	12120	11848	11351	11159	10803
SAGA with 100 node	31356	19108	15030	13402	12713	12278	11907	11419	11321	11123	10996

Similarly, under the 100-node scenario, the GA-based Approach and SAGA-based Approach achieve almost the same performance with just a slight difference.

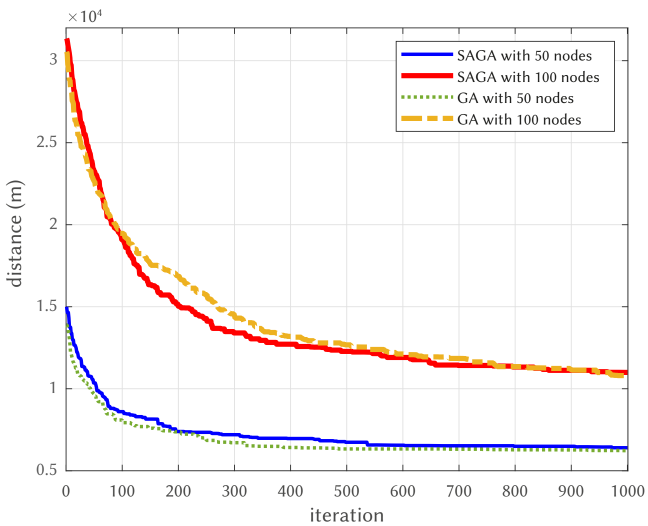


Fig. 19. Total distance for visiting the nodes by the GA-based Approach and the SAGA-based Approach under 50-node scenario and 100-node scenario in 1000 iterations.

From Table IV, we can make the following observations. Under the 50-node scenario, the GA-based Approach and SAGA-based Approach achieve almost the same performance with just a slight difference. In iteration 1000, the difference between the GA-based approach and the SAGA-based approach becomes very subtle (177 m, namely, 2.76% less than the SAGA-based approach). Similarly, under the 100-node scenario, the GA-based Approach and SAGA-based Approach achieve almost the same performance with just a slight difference. Just between iterations 100 and 400, the difference increases. However, after iteration 500, the difference decreases again. In iteration 1000, the difference between the GA-based approach and the SAGA-based approach becomes very subtle (193 m, namely, 1.76% less than the SAGA-based approach).

In iteration 1000, the difference between the GA-based approach and the SAGA-based approach becomes very subtle (193 m, namely, 1.76% less than the SAGA-based approach).

VII. Conclusions

Because there is a growing need for ocean exploration these days, research is focusing on longer range and greater exploration ranges. In this research, we present an efficient path-planning approach using an autonomous underwater vehicle with limited battery power for charging the underwater wireless sensor network (UWSN) and theoretically analyze its total energy usage. Due to the limited energy supply of the UWSN, we tackle the problem from the charging perspective. Several AUVs are a good approach to charge the UWSN to extend the exploration network. Furthermore, the charging efficiency and the range of exploration can be significantly increased by selecting suitable dive sites and designing a path that considers the node's location and data flow.

Data collection problems with autonomous underwater vehicles (AUV) can be handled by the following AI-based algorithms; Ant Colony Optimization-based Approach, Particle Swarm Optimization-based Approach, Teaching Learning-based Optimization-based Algorithm, Genetic Algorithm-based Approach, Grey Wolf Optimizer-based Approach. Simulations demonstrate that the AUV route planning system finds a better solution and converges more quickly than previous algorithms by using a genetic algorithm-based approach.

Different from the related literature, this work also considers the scenario where it is better not to use direct travel paths between specific pairs of nodes because of several reasons, such as flows below the surface, places too small for UAV movement, and extra hazards like electromagnetic waves. We propose a modified genetic algorithm-based approach, the Safety-Aware Genetic Algorithm (SAGA)-based Approach that introduces a very high cost for using the direct paths linking those nodes to tackle this more difficult scenario; thus, these direct paths will not be preferred during the path planning. In this scenario, the SAGA-based approach provides just a 3%longer path than the path provided by the GA-based approach. This shows that the SAGA-based approach performs very robustly for scenarios where it is better not to use direct travel paths between specific pairs of nodes for several reasons.

In the future, we can consider more complicated scenarios where the distance cost matrix can be defined instead of considering direct blockage in the links between the nodes through which direct traveling is very challenging because of the several reasons.

References

[1] T. Acarer, "The Turkish Model For Improving Imo Survey Results And Reducing Ship Accidents," *Dokuz Eylül University Maritime Faculty Journal*, vol. 11, no. 1, 2019, pp. 117-135.

[2] C. Aygün, *Türkiye ile Avrupa Birliği'nde Uygulanan Deniz Ulaştırma Politikaları ve Ekonomiye Etkileri*, T.C: İstanbul Üniversitesi Deniz Bilimleri ve İşletmeciliği Enstitüsü, Yüksek Lisans Tezi, İstanbul, 2012.

[3] International Chamber of Shipping, (2020). Description of the Subject International Chamber of Shipping (ICS), accessible at http://en.reingex.com/Chamber-Shipping.shtml#:~:text=The%20purpose%20of%20the%20International,transporting%20all%20types%20of%20cargo

[4] UNCTAD/ RMT/ 2018: REVIEW OF MARITIME TRANSPORT 2018, United Nations publication issued by the United Nations Conference on Trade and Development, UNCTAD: Geneva, Switzerland, 2018, accessible at https://unctad.org/en/PublicationsLibrary/rmt2018_en.pdf.

[5] N. Butt, D. Johnson, K. Pike, N. Pryce-Roberts, N. Vigar, "15 Years of Shipping Accidents: A review for WWF," Southampton Solent University, 2012.

[6] G. Kodak, T. Acarer, "Evaluation of the effect of maritime traffic regulations on the accident rate in the strait of Istanbul," *Aquatic Research*, vol. 4, no. 2, pp. 181-207, 2021, doi: 10.3153/AR21015.

[7] BBC News, "Egypt's Suez Canal blocked by huge container ship," 24 March 2021, available at https://www.bbc.com/news/world-middle-east-56505413

[8] J. M. Lee, E. Y. Wong, "Suez Canal blockage: an analysis of legal impact,

risks and liabilities to the global supply chain," *MATEC Web Conf.*, 339 01019, 2021, doi: 10.1051/matecconf/202133901019.

[9] O. Ozkanlisoy, E. Akkartal, "The Effect of Suez Canal Blockage on Supply Chains," *Dokuz Eylül University Maritime Faculty Journal*, vol. 14, no. 1, pp. 51-79, 2022, doi:10.18613/deudfd.933816.

[10] T. Gao & J. Lu, "The impacts of strait and canal blockages on the transportation costs of the Chinese fleet in the shipping network," *Maritime Policy & Management*, vol. 46, no. 6, pp. 669-686, 2019, doi: 10.1080/03088839.2019.1594423.

[11] S. Fan, Z. Yang, J. Wang, J. Marsland, "Shipping accident analysis in restricted waters: Lesson from the Suez Canal blockage in 2021," *Ocean Engineering*, vol. 266, no. 5, 113119, 2022, doi: 10.1016/j.oceaneng.2022.113119.

[12] T. Acarer, "New Possibilities Provided by Industrial Developments in the Management of Ships Belonging to Maritime Businesses and Unmanned Ships," *Mersin University Journal of Maritime and Logistics Research*, vol. 5, no. 2, pp. 122-153, 2023.

[13] T. Acarer, "Possibilities to be Provided by Maritime Businesses for Ship Management by Transforming VHF Short Distance Maritime Communication to Data Communication," *Maritime Research Journal: Amfora*, vol. 2, no. 3, pp. 15-31, 2023.

[14] E. Felemban, F. K. Shaikh, "Underwater sensor network applications: A comprehensive survey," *International Journal of Distributed Sensor Networks*, vol. 11, pp. 896832–896845, 2015, doi: 10.1155/2015/89.

[15] T. Qiu, Z. Zhao, "Underwater Internet of Things in Smart Ocean: System Architecture and Open Issues," *IEEE Transactions on Industrial Informatics*, vol. 16, pp. 4297–4307, 2020.

[16] I. F. Akyildiz, D. Pompili, "Underwater acoustic sensor networks: Research challenges," *Ad Hoc Networks*, vol. 3, no. 3, pp. 257–279, 2005.

[17] Q. Li, X. Du, "Energy-efficient data compression for underwater wireless sensor networks," *IEEE Access*, vol. 8, pp. 73395–73406, 2020.

[18] D.R. Pendergast, E.P. DeMauro, "A rechargeable lithium-ion battery module for underwater use," *Journal of Power Sources*, vol. 196, no. 2, pp. 793–800, 2011.

[19] D. R. Blidberg, "The development of autonomous underwater vehicles (AUV); a brief summary," *in Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2001)*, Seoul, Republic of Korea, vol. 4, pp. 122–129, 2001.

[20] H. Ghafoor, Y. Noh, "An overview of next-generation underwater target detection and tracking: An integrated underwater architecture," *IEEE Access*, vol. 7, pp. 98841–98853, 2019.

[21] L. Xie, Y. Shi, Y. Thomas Hou, W. Lou, H. D. Sherali, H. Zhou, "Rechargeable sensor networks with magnetic resonant coupling," Rechargeable Sensor Networks: Technology, Theory, and Application, 2014, pp. 31–68, doi: 10.1142/9789814525466_0002.

[22] J. Lee, N. Yun, "A focus on comparative analysis: Key findings of MAC protocols for underwater acoustic communication according to network topology," *in Proceedings of the Multimedia, Computer Graphics and Broadcasting: International Conference*, Jeju Island, Korea, 8–10 December 2011.

[23] N. Z. Zenia, M. Aseeri, "Energy-efficiency and reliability in MAC and routing protocols for underwater wireless sensor network: A survey," *Journal of Network and Computer Applications*, vol. 71, pp. 72–85, 2016, doi: 10.1016/j.jnca.2016.06.005.

[24] M.T.R. Khan, S.H. Ahmed, "An energy-efficient data collection protocol with AUV path planning in the internet of underwater things," *Journal of Network and Computer Applications*, vol. 135, pp. 20–31, 2019, doi: 10.1016/j.jnca.2019.02.025.

[25] Y. Su, Y. Xu, "HCAR: A Hybrid-Coding-Aware Routing Protocol for Underwater Acoustic Sensor Networks," *IEEE Internet Things Journal*, vol. 10, no. 12, pp. 10790–10801, 2023, doi: 10.1109/JIOT.2023.3240827.

[26] V. Kumar, D. Sandeep, "Multi-hop communication based optimal clustering in hexagon and voronoi cell structured WSNs," *AEU - International Journal of Electronics and Communications*, vol. 93, pp. 305–316, 2018, doi: 10.1016/j.aeue.2018.05.025.

[27] R. Xie, X. Jia, "Transmission-efficient clustering method for wireless sensor networks using compressive sensing," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 3, pp. 806–815, 2013, doi: 10.1109/TPDS.2013.90.

[28] S. Yadav, V. Kumar, "Hybrid compressive sensing enabled energy efficient transmission of multi-hop clustered UWSNs," *AEU - International Journal of Electronics and Communications*, vol. 110, pp. 152836–152851, 2019, doi: 10.1016/j.aeue.2019.152836.

[29] Y. Sun, M. Zheng, X. Han, S. Li, J. Yin, "Adaptive clustering routing protocol for underwater sensor networks," *Ad Hoc Networks*, vol. 136, pp. 102953–102965, 2022, doi: 10.1016/j.adhoc.2022.102953.

[30] R. Fan, Z. Jin, "A time-varying acoustic channel-aware topology control mechanism for cooperative underwater sonar detection network," *Ad Hoc Networks*, vol. 149, 103228, 2023, doi: 10.1016/j.adhoc.2023.103228.

[31] C. F. Liu, Z. Zhao, "A distributed node deployment algorithm for underwater wireless sensor networks based on virtual forces," *Journal of Systems Architecture*, vol. 97, pp. 9–19, 2019 doi: 10.1016/j.sysarc.2019.01.010.

[32] L. Wei, J. Han, "Topology Control Algorithm of Underwater Sensor Network Based on Potential-Game and Optimal Rigid Sub-Graph," *IEEE Access*, vol. 8, pp. 177481–177494, 2020, doi: 10.1109/ACCESS.2020.3024742.

[33] R. Zhu, A. Boukerche, "A trust management-based secure routing protocol with AUV-aided path repairing for Underwater Acoustic Sensor Networks," *Ad Hoc Networks*, vol. 149, pp. 103212–103225, 2023, doi: 10.1016/j.adhoc.2023.103212.

[34] Z. Yan, Y. Li, "Data collection optimization of ocean observation network based on AUV path planning and communication," *Ocean Engineering*, vol. 282, pp. 114912–114927, 2023, doi: 10.1016/j.oceaneng.2023.114912.

[35] G. Shen, X. Zhu, "Research on phase combination and signal timing based on improved K-medoids algorithm for intersection signal control," *Wireless Communications and Mobile Computing*, vol. 2020, 3240675, pp. 1-11, 2020, doi: 10.1155/2020/3240675.

[36] T. Kan, R. Mai, "Design and analysis of a Three-Phase wireless charging system for lightweight autonomous underwater vehicles," *IEEE Transactions on Power Electronics*, vol. 33, no. 8, pp. 6622–6632, 2018, doi: 10.1109/TPEL.2017.2757015.

[37] A.G. Ramos, V.J. García-Garrido, "Lagrangian coherent structure assisted path planning for transoceanic autonomous underwater vehicle missions," *Scientific Reports*, vol. 8, 4575, 2018, doi: 10.1038/s41598-018-23028-8.

[38] C. Cheng, Q. Sha, "Path planning and obstacle avoidance for AUV: A review," *Ocean Engineering*, vol. 235, pp. 109355–109368, 2021, doi: 10.1016/j.oceaneng.2021.109355.

[39] S.V. Kumar, R. Jayaparvathy, "Efficient path planning of AUVs for container ship oil spill detection in coastal areas," *Ocean Engineering*, vol. 217, pp. 107932–107945, 2020, doi: 10.1016/j.oceaneng.2020.107932.

[40] E. Golen, F. Mishra, "An underwater sensor allocation scheme for a range dependent environment," *Computer Networks*, vol. 54, pp. 404–415, 2010, doi: 10.1016/j.comnet.2009.09.016.

[41] Y. Yi, G. S. Yang, "Energy balancing and path plan strategy for rechargeable underwater sensor network," *in Proceedings of the 2022-4th International Conference on Advances in Computer Technology*, Suzhou, China, 22–24 April 2022.

[42] Y. Cui, P. Zhu, G. Lei, P. Chen, G. Yang, "Energy-Efficient Multiple Autonomous Underwater Vehicle Path Planning Scheme in Underwater Sensor Networks," *Electronics*, vol. 12, no. 15, 3321, 2023, doi: 10.3390/electronics12153321.

[43] Q. Ding, X. Xu, "Improved GWO Algorithm for UAV Path Planning on Crop Pest Monitoring," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 7, no. 5, pp. 30-39, 2022, doi:10.9781/ijimai.2022.07.002.

[44] Z. Zeng, K. Sammut, L. Lian, F. He, A. Lammas, Y. Tang, "A comparison of optimization techniques for AUV path planning in environments with ocean currents," *Robotics and Autonomous Systems*, vol. 82, pp. 61-72, 2016, doi: 10.1016/j.robot.2016.03.011.

[45] K. P. Carroll, S.R. McClaran, E.L. Nelson, D.M. Barnett, D.K. Friesen, G. William, "AUV path planning: an A* approach to path planning with consideration of variable vehicle speeds and multiple, overlapping, time-dependent exclusion zones," *in Proceedings of the 1992 Symposium on Autonomous Underwater Vehicle Technology (AUV'92)*, 1992, pp. 79–84.

[46] S. M. Lavalle, "Rapidly-Exploring Random Trees: A New Tool for Path Planning," Technical Report, Computer Science Dept., I.S.U. (TR 98–11), 1998.

[47] K. I. Tsianos, I.A. Sucan, and L.E. Kavraki, "Sampling-Based Robot Motion Planning: Towards Realistic Applications," *Computer Science Review*, vol.

1, no. 1, pp. 2-11, 2007. doi:10.1016/j.cosrev.2007.08.002.

[48] L. Kenye, R. Kala, "Optimistic Motion Planning Using Recursive Sub-Sampling: A New Approach to Sampling-Based Motion Planning," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 7, no. 4, pp. 87-99, 2022, doi:10.9781/ijimai.2022.04.001

[49] L. Kavraki, J. Latombe, "Randomized Preprocessing of Configuration for Fast Path Planning," *in Proceedings of the 1994 IEEE International Conference on Robotics and Automation*, vol. 2133, 1994, pp. 2138-2145, doi:10.1109/ROBOT.1994.350966.

[50] J. Witt, M. Dunbabin, "Go with the flow: Optimal AUV path planning in coastal environments," *in Australian conference on robotics and automation*, vol. 2008, no. 2, 2008.

[51] O. M. Gul, T. Acarer, "Robust and Energy-Aware Path Planning by Autonomous Underwater Vehicle in Underwater Wireless Sensor Networks for Monitoring Maritime Transportation," *EMO Scientific Journal*, vol. 14, no. 2, pp. 71–85, 2024.

[52] P. C. Pop, O. Cosma, C. Sabo, C. P. Sitar, "A comprehensive survey on the generalized traveling salesman problem," *European Journal of Operational Research*, vol. 314, no. 3, 2024, pp. 819-835, 2024, doi: 10.1016/j.ejor.2023.07.022.

[53] D. Davendra, *Travelling Salesman Problem, Theory and Applications*, InTech, 2010.

[54] D.S. Johnson, L.A. McGeoch, *The Traveling Salesman Problem: A Case Study, Local Search in Combinatorial Optimization*, pp. 215–310, John Wiley & Sons, 1997.

[55] K.L. Hoffman, M. Padberg, "Traveling salesman problem," *In Encyclopedia of Operations Research and Management Science*, eds: S.I. Gass, C.M. Harris, Springer, New York, NY, 2001.

[56] G. Gutin, A. Punnen, *The Traveling Salesman Problem and Its Variations. Combinatorial Optimization*, vol. 12. Kluwer, Dordrecht, 2002.

[57] M. Dorigo, G. DiCaro, *The ant colony optimization meta-heuristic*, in New Ideas in Optimization, McGraw-Hill, 1999.

[58] M. Dorigo, M. Birattari and T. Stutzle, "Ant colony optimization," *in IEEE Computational Intelligence Magazine*, vol. 1, no. 4, pp. 28-39, 2006, doi: 10.1109/MCI.2006.329691.

[59] J. Yu, X. You, S. Liu, "A heterogeneous guided ant colony algorithm based on space explosion and long–short memory," *Applied Soft Computing*, vol. 113, Part B, 2021, 107991, doi: 10.1016/j.asoc.2021.107991.

[60] J. Kennedy and R. Eberhart, "Particle swarm optimization," *Proceedings of ICNN'95 - International Conference on Neural Networks*, Perth, WA, Australia, 1995, pp. 1942-1948 vol.4, doi: 10.1109/ICNN.1995.488968.

[61] M. R. Bonyadi, Z. Michalewicz, "Particle swarm optimization for single objective continuous space problems: a review," *Evolutionary Computation*, vol. 25, no. 1, pp. 1–54, 2017, doi:10.1162/EVCO_r_00180.

[62] R. Thangaraj, M. Pant, A. Abraham, V. Snasel, "Modified particle swarm optimization with time varying velocity vector," *International Journal of Innovative Computing, Information and Control*, vol. 8, no. 1, pp. 201–218, 2012.

[63] A. G. Gad, "Particle Swarm Optimization Algorithm and Its Applications: A Systematic Review," *Archives of Computational Methods in Engineering*, vol. 29, pp. 2531–2561, 2022, doi: 10.1007/s11831-021-09694-4

[64] R. V. Rao, V.J. Savsani, D.P. Vakharia, "Teaching–learning-based optimization: A novel method for constrained mechanical design optimization problems," *Computer-Aided Design*, vol. 43, no. 3, pp. 303-315, 2011, doi: 10.1016/j.cad.2010.12.015.

[65] Y. Xu, Y. Peng, X. Su, Z. Yang, C. Ding, X. Yang, "Improving teaching–learning-based-optimization algorithm by a distance-fitness learning strategy," *Knowledge-Based Systems*, vol. 257, 2022, 108271, doi: 10.1016/j.knosys.2022.108271.

[66] S. Mirjalili, S.M. Mirjalili, A. Lewis, "Grey wolf optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014, doi: 10.1016/j.advengsoft.2013.12.007.

[67] D.E.Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley Longman Publishing Co., Inc.,1989.

[68] E. Bonabeau, M. Dorigo, G. Theraulaz, *Swarm intelligence: from natural to artificial systems*, Oxford university press, 1999.

Tayfun Acarer

He received a B.S. degree in electronics engineering from Istanbul Technical University (ITU), Istanbul, Türkiye, in 1980, the M.S. and Ph.D. degrees from Istanbul University, Istanbul, in 1992 and 1995, respectively. He has worked at several companies in the ICT Sector for more than two decades, he worked as Chair of the ICT Regulatory Body (ICTA) from 2005 to 2015. He started to work as an Assistant Professor at Bilgi University, Istanbul, in 2016. Since 2018, he has been with the National Metrology Institute (UME), Gebze, Turkey, where he is currently a Board Member. He has given lectures on Information Technologies and Marine Communication and Electronics Navigation Systems at several universities in Turkey and Technical University Berlin (TU Berlin) in Germany for nearly two decades. He is currently working as a faculty at Maritime Transportation and Management Vocational School of Higher Education, Piri Reis University in Tuzla, Istanbul, Türkiye.

Evaluating the Impact of Pumping on Groundwater Level Prediction in the Chuoshui River Alluvial Fan Using Artificial Intelligence Techniques

Yu-Sheng Su [1]*, Yu-Cheng Hu[2], Yun-Chin Wu[2], Ching-Teng Lo[2]

[1] Department of Computer Science and Information Engineering, National Chung Cheng University, Chiayi county (Taiwan)
[2] Department of Computer Science and Engineering, National Taiwan Ocean University, Keelung city (Taiwan)

* Corresponding author: ccucssu@gmail.com

Received 15 June 2023 | Accepted 21 December 2023 | Early Access 8 April 2024

Abstract

Over the past decade, excessive groundwater extraction has been the leading cause of land subsidence in Taiwan's Chuoshui River Alluvial Fan (CRAF) area. To effectively manage and monitor groundwater resources, assessing the effects of varying seasonal groundwater extraction on groundwater levels is necessary. This study focuses on the CRAF in Taiwan. We applied three artificial intelligence techniques for three predictive models: multiple linear regression (MLR), support vector regression (SVR), and Long Short-Term Memory Networks (LSTM). Each prediction model evaluated the extraction rate, considering temporal and spatial correlations. The study aimed to predict groundwater level variations by comparing the results of different models. This study used groundwater level and extraction data from the CRAF area in Taiwan. The dataset we constructed was the input variable for predicting groundwater level variations. The experimental results show that the LSTM method is the most suitable and stable deep learning model for predicting groundwater level variations in the CRAF, Taiwan, followed by the SVR method and finally the MLR method. Additionally, when considering different distances and depths of pumping data at groundwater level monitoring stations, it was found that the Guosheng and Hexing groundwater level monitoring stations are best predicted using pumping data within a distance of 20 kilometers and a depth of 20 meters.

Keywords

Artificial Intelligence, Chuoshui River Alluvial Fan, Groundwater Level Prediction, Water Pumping.

DOI: 10.9781/ijimai.2024.04.002

I. Introduction

Groundwater is the most important freshwater resource, and is widely used for industrial, commercial and agricultural purposes. The Chuoshui River Alluvial Fan (CRAF) is an important agricultural and industrial area that uses a significant amount of groundwater resources. However, excessive groundwater use can lead to serious environmental problems, including changes in river flow, land subsidence, and seawater intrusion [1]-[2]. In order to manage groundwater resources sustainably, and to detect the changes in groundwater levels, the interaction between pumping rate and groundwater level is studied at different temporal and spatial scales [2]. The pumping rate of the wells can be calculated using a numerical model and the relationship between electricity consumption and pumping rate. A time-series analysis was used to establish a time-dependent groundwater level processing model, and artificial intelligence was applied to predict groundwater level variations and analyze how much groundwater extraction would lead to irreversible land subsidence [3]. In this study, the calculated pumping rate ws used in conjunction with the time-dependent groundwater level processing model to establish a groundwater extraction prediction mechanism to achieve the goal of predicting and warning about groundwater level changes.

The establishment of groundwater resource management requires a better understanding and monitoring of the relationship between groundwater level fluctuations and the spatial distribution of land subsidence [4]-[5]. During the data processing, analysis, and modeling, groundwater sensors were used to collect real-time groundwater level data over a long period of time. However, abnormal conditions of groundwater sensors, measurement failures, network connectivity problems, human errors, and other factors can lead to abnormal data at monitoring stations, resulting in data errors or loss [6]. Previous studies have utilized time series techniques to analyze the spatio-temporal distribution of groundwater level data and systematically clean and impute missing values. Time series techniques can quantify the

Please cite this article as: Y. S. Su, Y. C. Hu, Y. C. Wu, C. T. Lo. Evaluating the Impact of Pumping on Groundwater Level Prediction in the Chuoshui River Alluvial Fan Using Artificial Intelligence Techniques, International Journal of Interactive Multimedia and Artificial Intelligence, vol. 8, no. 7, pp. 28-37, 2024, http://dx.doi.org/10.9781/ijimai.2024.04.002

autoregressive characteristics of the data as well as the corresponding variabilities associated with short-term fluctuations [7]. In the field of time series techniques, this study utilized the recursive seasonal-trend decomposition method to analyze the trend variations of groundwater levels in the Chuoshui River area. In addition, the Bayesian maximum entropy method was applied to impute the uncertain groundwater level data to ensure the integrity of the groundwater level dataset. Artificial intelligence techniques can efficiently compute non-linear data and use optimal algorithmic functions to model and analyze the groundwater environment. Prediction of groundwater level fluctuations is a critical aspect of ensuring effective groundwater management and availability in the future [8]-[12].

The dataset used in this study consists of groundwater level and extraction data from the CRAF area in Taiwan. The aim was to assess the occurrence of groundwater level decline in the CRAF area, and to analyze the relationship between groundwater extraction and the level of decline to determine groundwater use. We then used regression techniques to predict the trend changes in groundwater levels in the absence of extraction [11], [13]. This approach involves establishing a spatio-temporal data-driven framework that analyzes the temporal nature of the groundwater level dataset to obtain spatiotemporal characteristics of groundwater use in the CRAF area. The extracted spatio-temporal characteristics were then incorporated into the groundwater level prediction model together with the extraction data [12]-[13]. We compared three artificial intelligence techniques in our prediction methods: multiple linear regression (MLR), support vector regression (SVR), and Long Short-Term Memory Networks (LSTM) [7] [14], [15]. Each predictive model can be used to validate the effectiveness of these techniques and compare the accuracy of groundwater level predictions. The research question was to explore suitable predictive models to forecast the trend of groundwater level variations from 2020 to 2021. From the results of the predictive assessment, the strengths and weaknesses of each predictive model were identified and a suitable pumping dataset was found for predicting groundwater level variations. The generation of an appropriate pumping dataset was based on the location and depth of the groundwater monitoring station as a central zone, collecting data from nearby pumping wells.

The remainder of the paper is organized as follows: Section II explains the related works. Section III explains the research methodology. Section IV gives an overview of the experimental results, and Section V presents the conclusions.

II. Related Works

A. Linear Regression Applied to Groundwater Level Prediction

The study of groundwater levels begins with data collection and observation through sensors installed at groundwater monitoring stations [16]. Under the influence of specific ecological and climatic processes, the normal fluctuations in groundwater levels show stability, resulting in regular patterns of rise and fall. This results in time series data with a continuous distribution [17]-[18]. Groundwater level data exhibit linear correlations, and previous studies have utilized linear models to achieve the best possible fit and use of such data. The principle behind this is to use the method of least squares to model the relationship between one or more independent variables and a dependent variable in regression analysis [19]-[20]. Due to their suitability for handling regression problems with continuous data, linear regression models have been widely used in previous research for statistical analysis and prediction of groundwater data [21]-[23]. In the study by Yan et al., a linear regression model was developed to predict groundwater levels in the coastal plains of eastern China using data such as precipitation, evaporation, river water levels, and tides. By analyzing the trend of groundwater level variations and performing linear regression analysis, satisfactory prediction results can be obtained with effective data and computational models [23].

We are investigated how pumping behavior affects groundwater level variations. However, the pumping data have a non-normal distribution and are difficult to fit into a multiple linear regression model. Considering this, the multiple linear regression model was not the optimal choice, and so we opted for machine learning models based on classifiers for data analysis and modeling [14], [24], [25]. In this study, we incorporated support vector regression and Long Short-Term Memory Networks (LSTM), a type of recurrent neural network, to handle both linear and non-linear relationships in the data, thus achieving better analytical results.

B. Support Vector Regression Applied to Groundwater Level Prediction

Excessive groundwater pumping may affect groundwater level fluctuations and cause land subsidence in the CRAF area in Taiwan. Previous studies have focused on the problem of irreversible land subsidence caused by excessive groundwater extraction [24], [26]. In the field of artificial intelligence, the support vector regression (SVR) model is utilized to analyze the trend changes between linear groundwater levels and non-linear extraction data. As a machine learning model based on classifier design, SVR minimizes structural risk and exhibits strong adaptability, global optimization and excellent generalization ability with respect to the data [27]. Previous studies have demonstrated that the SVR model has better predictive performance when analyzing non-linear data [11], [28] - [30]. El Bilali [30] used four machine learning models, including adaptive boosting, random forest, artificial intelligence and support vector regression, to assess and predict the water quality of the Berrechid aquifer in northwestern Morocco. The results of the research showed that Support Vector Regression had less sensitivity to input variables and better generalization capabilities compared to Adaptive Boosting and Random Forest. Therefore, it was more suitable for evaluating and predicting different types of data [30]. Mirarabi et al. [31] conducted a performance comparison between support vector regression and artificial intelligence models using groundwater data from the Hashtgerd Plain in Alborz Province, Iran. They found that the accuracy of both models declined over time.

Based on the above, we have found that Support Vector Regression (SVR) can be effectively utilized for both linear and non-linear regression analysis. This model can improve its predictive performance by optimizing the model parameters and performing data pre-processing [32]-[35]. Therefore, optimizing the parameters of the SVR model in research has a crucial positive impact on improving model performance.

C. Recurrent Neural Network Applied to Groundwater Level Prediction

In recent years, the extensive utilization of sensor techniques in groundwater monitoring has resulted in a greater number of influencing factors that need to be analyzed and investigated [36]. In the past, predictive models were often limited to capturing shallow correlations in the data, and were unable to uncover deeper relationships, resulting in inaccurate predictions [37]-[38]. Hinton and colleagues proposed the use of unsupervised learning with Deep Belief Networks (DBNs), which have the advantage of using hierarchical feature representations to model deep and complex nonlinear relationships [39]. The exceptional performance of such networks has made deep learning a trendsetter. Among them, Recurrent Neural Networks (RNNs) are widely regarded as effective methods for capturing the temporal dependencies in sequential data [40]-[42]. However, according to relevant studies, traditional RNNs are only suitable for processing short sequences. When applied to long sequential data, they can suffer from problems such as vanishing gradients, where the network struggles to remember long-term information, or exploding gradients [43]-[44]. To address this problem, Hochreiter and Schmidhuber [45] proposed Long Short-Term Memory Networks (LSTMs), a model that introduced memory cells capable of storing information for longer periods of time. This breakthrough allowed for more effective modeling and prediction of long sequential data. Gers et al. [46] improved the LSTM by introducing the forget gate mechanism, which allows the model to selectively retain or discard information from the previous cell state. This enhancement significantly improved the predictive performance of the model.

The LSTM method has gained considerable prominence in groundwater research. Zhang et al. [47] used it to predict the depth of groundwater levels in agricultural regions of China. The results indicated that LSTM excels in capturing the intricate relationships between linear and non-linear dynamics present in long-term sequential data, making it a key factor in improving the efficiency of agricultural irrigation and groundwater management. Vu et al. [48] utilized the LSTM method to reconstruct missing groundwater level data in the Normandy region of France. The result demonstrated the effectiveness of this approach in successfully reconstructing the missing groundwater level data, thereby improving the accuracy and reliability of hydrological forecasting and management.

Based on the current literature review, this study used the MLR, SVR, and LSTM methods to analyze the data from groundwater monitoring stations and the electricity consumption data from pumping wells in the Chuoshui River alluvial fan area of Taiwan. These models are utilized for deep learning modeling and analysis. In addition, this study incorporated optimization algorithms to enhance the predictive performance of the models, thereby improving the accuracy of groundwater level prediction.

III. Methods

A. Study Area

The CRAF is the largest alluvial fan plain in Taiwan [26]. It stretches from the Wu River in the north to the Beigang River in the south, from the Taiwan Strait in the west to the Bagua Mountain Plateau and the Douliu Hills in the east. The area is about 2,100 square kilometers. The main river is the Chuoshui River. Flowing from east to west, the Chuoshui River crosses the alluvial fan in the central mountain range before emptying into the Taiwan Strait [49]-[50]. Fig. 1 shows the geographical extent of the entire CRAF area, represented by solid lines.

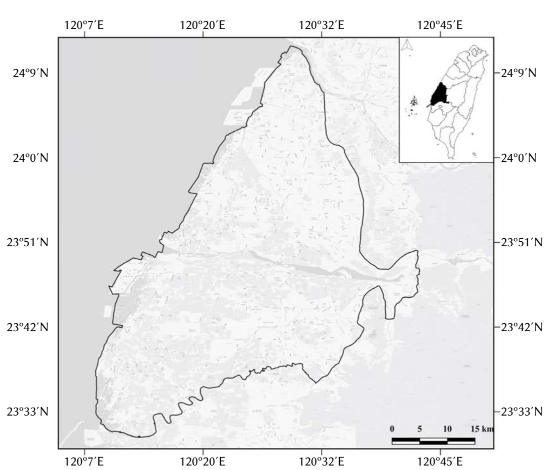


Fig. 1. The geographical location of the Chuoshui River Alluvial Fan.

The CRAF area consists of four underground aquifer layers and three aquitard layers. To access the groundwater resources, most of the pumping wells extract water from aquifer 1 and aquifer 2 [26]. In addition, the groundwater monitoring stations have a numerical code (1, 2, 3, 4) appended to their names, indicating the specific aquifer layer in which the groundwater monitoring station is located.

B. Groundwater Level Data Acquisition and Processing

This study used groundwater level data collected from 2007 to 2021. The data were obtained from the Integrated Service System for Ground Subsidence Monitoring of the Water Resources Agency and the Hydrological Information Network of the Ministry of Economic Affairs, Water Resources Agency. The data collection period for the Integrated Service System for Ground Subsidence Monitoring extended until December 31, 2022, with groundwater level data provided at 24-hour intervals. The data collection period for the Hydrological Information Network of the Ministry of Economic Affairs started from January 1, 2020 to the present, with groundwater level data provided at 10-minute intervals.

The Python programming (version 3.8.10) can be used to pre-process and store groundwater level data. Python libraries such as Pandas, NumPy, Requests, and InfluxDB_Client can be used to write a web crawler program that downloads the raw groundwater level data and stores it in an InfluxDB (version 2.0.9) database [51]. Although the collected groundwater level data come from different sources, there are common elements between the two datasets. The dataset from the Hydrological Information Network of the Ministry of Economic Affairs can be matched with the code data from the Integrated Service System for Ground Subsidence Monitoring of the Water Resources Agency. If the time series and water level values of the groundwater data are the same, this indicates that the information has been collected from the same groundwater level monitoring station. The processed groundwater level information can be organized as shown in Table I.

TABLE I. Data on Groundwater Level

Name	Descriptions	Examples
ST_NO	Groundwater level monitoring station code	07010111
NAME_C	Chinese name of the groundwater level monitoring station, with a number indicating the aquifer level.	Guosheng (1)
Time	Timestamp of observation	2022-01-01T00:00:00Z
Water_Level(m)	Groundwater level height, with data units in metres.	16.943

From the Integrated Service System for Ground Subsidence Monitoring of the Water Resources Bureau and the Hydrological Information Network of the Ministry of Economic Affairs, a total of 260 groundwater monitoring stations within the CRAF area were considered for organizing the time-series groundwater level datasets. However, not every monitoring station has complete groundwater level data available. Therefore, a comparison was made between the data collected from the groundwater monitoring stations and the map of abandoned groundwater monitoring wells [52]. It was found that there were anomalies and abandoned statuses in the data collection for 95 monitoring stations.

In the end, we had 165 monitoring stations that were currently active and had complete data. These 165 groundwater monitoring stations are located at different depths. Fig. 2 shows the distribution and number of groundwater monitoring stations in the different aquifers. (A) The first aquifer of the CRAF contains 62 groundwater monitoring stations, most of which are located at depths between 25 and 100 meters. (B) The second aquifer of the CRAF contains 55

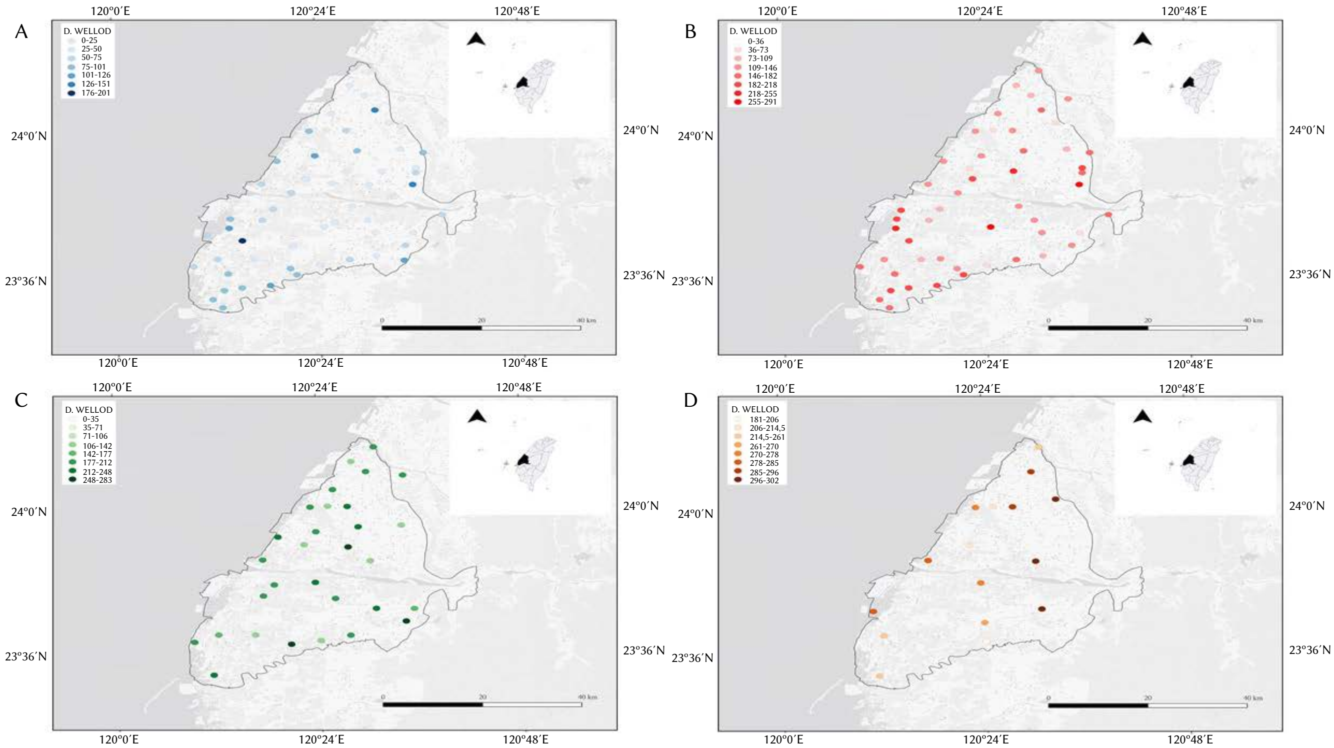


Fig. 2. Map showing the distribution of groundwater level monitoring stations in the Chuoshui River Alluvial Fan.

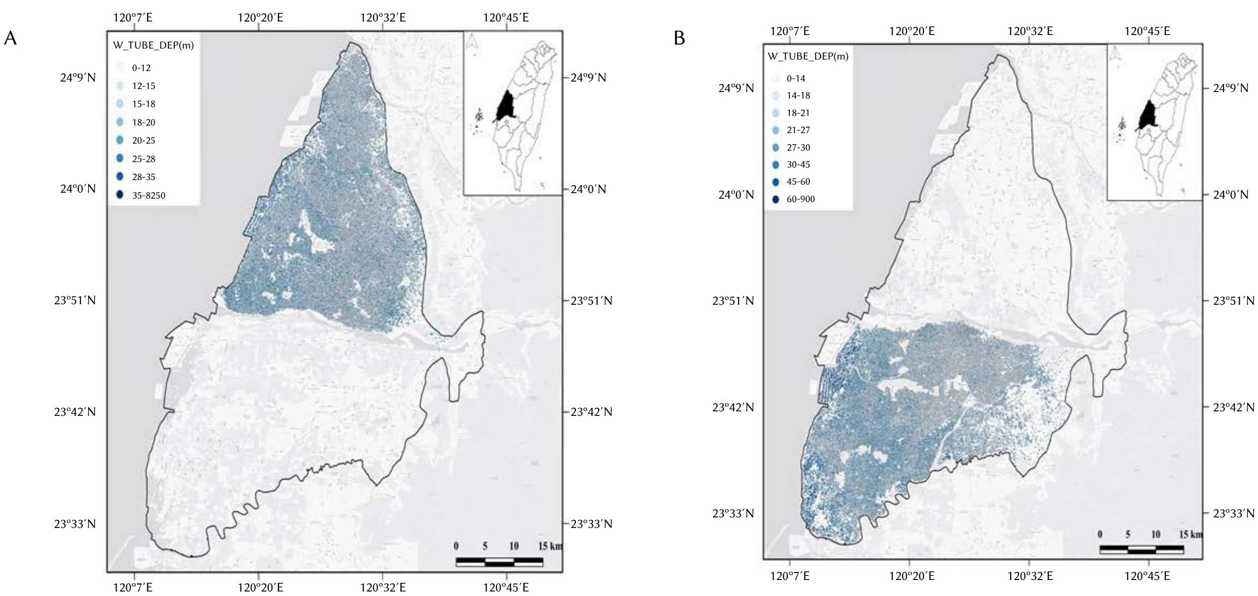


Fig. 3. Map of pumping well station distribution in the Chuoshui River Alluvial Fan.

groundwater monitoring stations, most of which are located at depths between 73 and 182 meters. (C) The third aquifer of the CRAF contains 32 groundwater monitoring stations, with most stations located at depths ranging from 142 to 248 meters. (D) The fourth aquifer of the CRAF contains 18 groundwater monitoring stations, most of which are located at depths between 218 and 302 meters. In Fig. 2, it is evident that each groundwater level monitoring station is situated at a different aquifer depth. Therefore, this study aimed to compare the accuracy of different deep learning models in predicting groundwater level variations within the context of the same aquifer depth. This allowed us to construct groundwater level prediction models for similar groundwater environments.

C. Water Pumping Data Acquisition and Processing

The study collected water extraction data from the CRAF area in Taiwan between 2007 and 2021. This included electricity consumption data from pumping wells located in Changhua County and Yunlin County. The electricity consumption of the pumping wells was collected using electricity meters, with data provided at monthly intervals. The electricity consumption data were used to simulate the groundwater extraction volume for each month. The water extraction data record the coding of the pumping wells, their latitude and longitude coordinates, installation depths, and electricity consumption, as shown in Table II.

There are a total of 242,586 pumping wells in the CRAF area, including 125,905 wells in Changhua County and 116,681 wells in Yunlin County. Fig. 3 shows the locations of the pumping wells installed in Changhua County and Yunlin County, indicating the depth ranges of the pumping well stations.

TABLE II. Data on Pumping Wells

Name	Descriptions	Examples
WELL_NO	Pumping well code	10237500000004
TIME	Observation timestamp	2007-01-01T00:00:00
Pumping_well_power	Electricity consumption data of pumping wells	0.0
Lon	Longitude (WGS84)	120.506997
Lat	Latitude (WGS84)	24.075657
W_TUBE_DEP	Depth at which the station is located underground in metres	16.0

This study used the method of determining flow rate based on the relationship between electricity consumption and pumping volume to calculate the pumping volume of water wells in the Yunlin and Changhua regions. The hybrid pumping equipment includes an electronic water meter, an electronic electricity meter, and a pumping motor on/off time recorder. The electronic electricity meter records the monthly electricity consumption. By utilizing the electricity consumption data and the water/electricity ratio specific to each local pumping well, the pumping flow rate can be calculated to determine the pumping volume for each month. The water/electricity ratio is expressed in cubic meters per kilowatt-hour and represents the pumping volume of water per unit of electricity consumed. It can be obtained by dividing the pumping flow rate by the power consumption [53]. Fig. 4 shows the time series chart of the pumping volume for pumping station 10237500000004. The water/electricity ratio parameter for this pumping station is 12.9 cubic meters per kilowatt-hour. The pumping period is from 2007 to July 2021, and the unit of water volume is cubic meters.

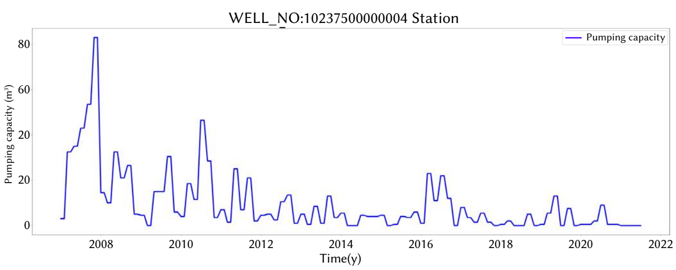


Fig. 4. The time series chart of the pumping volume for pumping station with ID 10237500000004.

D. Groundwater Level Prediction Techniques

To analyze the impact of pumping rates on groundwater level changes, this study proposes a groundwater level prediction model to determine how much the pumping wells within a certain range and depth will influence the accuracy of predicting groundwater levels. This method provides an actionable approach to manage and monitor groundwater levels, allowing an understanding of how the groundwater level changes at a single groundwater observation station are influenced by pumping wells within specific ranges and depths. In this study, we employed the method of determining flow rate based on the relationship between electricity consumption and pumping volume to accurately calculate the pumping data. Three models were used to predict the variations in groundwater levels: multiple linear regression, SVR, and LSTM. By analyzing the accuracy of predicting groundwater levels, we were able to evaluate the strengths and weaknesses of these models. The workflow of prediction modeling and validation settings is illustrated in Fig. 5.

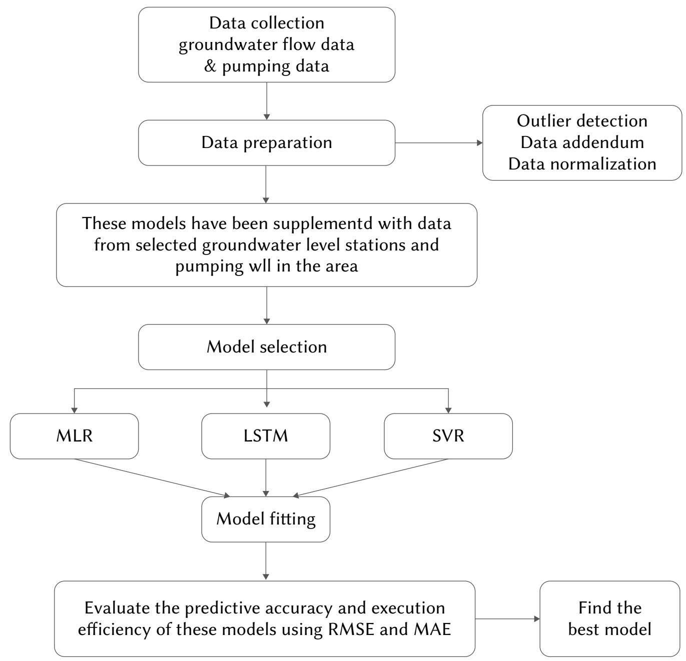


Fig. 5. The flowchart of the groundwater level prediction model.

In Fig. 5, this study begins with the selection of potential features from the collected data. Both pumping volume and groundwater level data are subjected to outlier detection, imputation, and normalization processes to obtain the appropriate format [1], [54], [55]. However, the relative position and depth between groundwater monitoring stations and pumping wells can affect the model construction and the accuracy of groundwater level prediction. During the model execution, the information from groundwater observation stations is selected as the target for model construction. Additionally, data from pumping wells located at different distances and depths were collected based on the position and depth information of the groundwater observation stations [56]. The collected pumping volume dataset served as the input dataset, and the groundwater level data served as the output dataset. Both datasets are used together to train, validate, and test the candidate deep learning models, including multiple linear regression, support vector regression, and LSTM models. These models were used to estimate the groundwater levels at the groundwater monitoring stations. The aim was to understand how pumping wells within a certain radius distance influence the prediction accuracy of groundwater levels at the observation stations.

In the final step, the estimated groundwater data and the observed groundwater data were used to test the models. The prediction accuracy of the models was compared using metrics such as root mean square error and mean absolute error. This analysis helps to identify which ranges and depths of pumping volume datasets are suitable for the groundwater monitoring stations.

E. Evaluation Metrics

Accurately assessing the performance and accuracy of deep learning models is crucial. We seek to evaluate the predictive performance of multiple linear regression, support vector regression, and LSTM models in predicting groundwater levels [1]. The accuracy of the aforementioned deep learning models is assessed using root mean square error (RMSE) and mean absolute error (MAE) for evaluation purposes. These two metrics are commonly used to measure the accuracy of predicting groundwater levels. The RMSE is more sensitive to outliers and is suitable for errors that follow a Gaussian distribution. On the other hand, MAE calculates the average weighted error for all errors [11], [14]. The evaluation of the RMSE and MAE in this study represents the accuracy of the deep learning model in predicting the groundwater level sequence from 2020 to August 2021. By comparing the predicted groundwater levels obtained from the deep learning model with the original groundwater levels, RMSE and MAE results were generated. When MAE = 0, RMSE = 0, or if they approach 0, it indicates the highest consistency between the predicted values and the observed values, demonstrating a better performance of the model in predicting groundwater levels.

IV. Results

This study focused on the alluvial fan area of the Chuoshui River to investigate the relationship between groundwater level variations and pumping behavior. In this study, groundwater level monitoring stations in Changhua County, including Guosheng, Tianwei and Hexing, and in Yunlin County, including Dongguang, Wencuo and Fengrong, were selected as research objects. These stations were selected because they showed more than 98% of normal values after analysis using the recursive seasonal trend decomposition method, which requires minimal adjustment for abnormal values [57], [58]. By using the installation locations and depths of each groundwater observation station, we created different datasets of pumping rates under various geographic conditions. The generation of pumping rate datasets primarily focused on pumping well stations located within a radius of 10 and 20 kilometers and at depths of ±10, 15, and 20 meters relative to each groundwater observation station. Table III shows the number of pumping well stations requiring data collection and organization under five distance and depth conditions. The collected data were aggregated as input variables to predict the groundwater level variations at the monitoring stations.

TABLE III. Number of Well Used to Collect Pumping Station Data

Sampling range	**Groundwater monitoring station**					
	Guo sheng	Tian wei	Hexing	Dong guang	Wen cuo	Feng rong
10km/10m	968	4275	7494	5322	1195	2219
10km/15m	1565	6629	11790	8002	1862	3399
10km/20m	2475	9235	17402	10716	2510	4676
20km/15m	7680	12038	20598	16450	11000	10583
20km/20m	12186	16699	30856	22227	15035	14431

Using data from pumping well stations as input variables, we constructed MLR, SVR and LSTM models to predict groundwater level variations from January 2020 to August 2021. Despite their different implementation methods, MLR, SVR, and LSTM models all belong to the field of machine learning and deep learning techniques. In this study, the pumping well station data were pre-processed and normalized to reduce errors. 89% of the pumping data were used for model building to analyze the groundwater level variations between 2007 and 2020 and to determine the initial prediction values. Training performance was evaluated by selecting the lowest RMSE and correlation coefficient. In addition, 11% of the pumping data were used for testing and predicting subsequent groundwater level variations [59]. In this study, the MLR model used the linear regression algorithm. The SVR model was configured with parameters such as a linear kernel, an epsilon value of 0.01, gamma set to auto mode, and a soft margin (C) value of 1. The LSTM model was constructed with 3 input layers, 2 dropout layers with a dropout rate of 0.2, and 1 output layer. The optimizer used for the LSTM was the Adam optimizer with a learning rate of 0.0001 [60]. All of these models were used to predict groundwater level fluctuations, and Tables IV to IX show the tested groundwater level monitoring stations paired with different samples of the pumping environment. By using the MLR, SVR and LSTM models to predict groundwater level variations for each groundwater level monitoring station from January 1, 2020 to August 1, 2021, the RMSE and MAE values were obtained to evaluate the prediction accuracy of the models [61].

The datasets from each groundwater observation station and pumping station are input into the artificial intelligence prediction process for predicting groundwater levels. Tables IV to IX display the predictive accuracy for each groundwater observation station. By identifying the lowest RMSE (Root Mean Square Error) and MAE (Mean Absolute Error) values, it is possible to determine the suitable pumping station datasets within a certain range of distance and depth that match the location and observation depth of each groundwater observation station.

TABLE IV. Performance Evaluation of Using Three Models to Predict Groundwater Level Variations at Guosheng (1) Station

	MLR		**SVR**		**LSTM**	
Sampling range	RMSE	MAE	RMSE	MAE	RMSE	MAE
10KM/10M	0.536	0.388	0.346	0.293	0.308	0.256
10KM/15M	0.512	0.367	0.510	0.398	0.291	0.242
10KM/20M	0.450	0.366	0.424	0.311	0.344	0.268
20KM/15M	0.451	0.327	0.373	0.309	0.341	0.255
20KM/20M	0.227	0.181	0.425	0.328	0.287	0.236

The installation location of the Guosheng groundwater monitoring station is located at longitude 120° 56' 91" and latitude 24° 09' 26", at a depth of 24 meters. Table IV shows the prediction accuracy for the Guosheng groundwater monitoring station. It was found that the most suitable pumping volume dataset for predicting groundwater level variations at the Guosheng groundwater monitoring station is within a distance difference of 20 kilometers and a depth difference of 20 meters. Furthermore, based on the RMSE and MAE values, it was determined that the Multiple Linear Regression (MLR) model is the most suitable for predicting groundwater level variations at the Guosheng groundwater monitoring station. Therefore, Fig. 6 shows the groundwater level time series plot for the Guosheng groundwater monitoring station. The red line represents the actual groundwater level time series, while the blue line represents the groundwater level time series predicted by the MLR model.

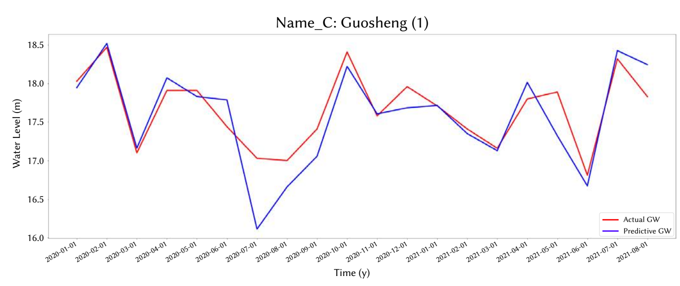


Fig. 6. The actual groundwater level and predicted groundwater level chart for the Guosheng groundwater monitoring station.

TABLE V. Performance Evaluation of Using Three Models to Predict Groundwater Level Variations at Tianwei (1) Station

	MLR		SVR		LSTM	
Sampling range	RMSE	MAE	RMSE	MAE	RMSE	MAE
10KM/10M	0.408	0.354	0.484	0.368	0.350	0.261
10KM/15M	0.390	0.324	0.306	0.242	0.285	0.228
10KM/20M	0.506	0.405	0.484	0.367	0.388	0.252
20KM/15M	0.372	0.299	0.361	0.283	0.547	0.378
20KM/20M	0.388	0.310	0.390	0.297	0.225	0.185

The installation location of the Tianwei groundwater monitoring station is at longitude 120° 52’ 73” and latitude 23° 89’ 13”, at a depth of 36 meters. Table V presents the prediction accuracy for the Tianwei groundwater monitoring station. It was found that the most suitable pumping volume dataset for predicting groundwater level variations at the Tianwei groundwater monitoring station is within a distance difference of 10 kilometers and a depth difference of 15 meters. Furthermore, based on the RMSE and MAE values within a distance of 10 kilometers and a depth of 15 meters, it was determined that the Long Short-Term Memory (LSTM) model is the most suitable for predicting groundwater level variations at the Tianwei groundwater monitoring station. Therefore, Fig. 7 shows the groundwater level time series chart for the Tianwei groundwater monitoring station. The red line represents the actual groundwater level time series, while the blue line represents the groundwater level time series predicted by the LSTM model.

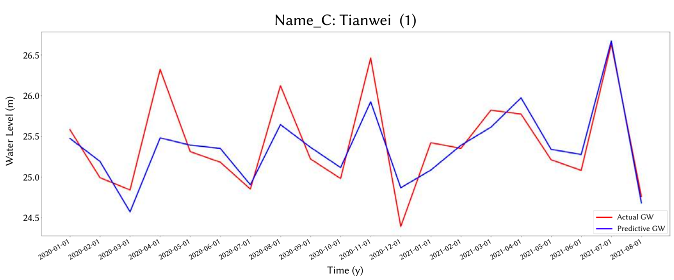


Fig. 7. The actual groundwater level and predicted groundwater level chart for the Tianwei groundwater monitoring station.

TABLE VI. Performance Evaluation of Using Three Models to Predict Groundwater Level Variations at Hexing (1) Station

	MLR		SVR		LSTM	
Sampling range	RMSE	MAE	RMSE	MAE	RMSE	MAE
10KM/10M	0.721	0.418	0.692	0.445	0.521	0.365
10KM/15M	0.596	0.390	0.551	0.360	0.604	0.415
10KM/20M	0.701	0.446	0.687	0.463	0.447	0.341
20KM/15M	0.555	0.365	0.885	0.544	0.513	0.406
20KM/20M	0.727	0.437	0.229	0.186	0.319	0.250

The installation location of the Hexing groundwater monitoring station is at longitude 120° 45’ 81” and latitude 23° 89’ 40”, at a depth of 23 meters. Table VI represents the prediction accuracy for the Hexing groundwater monitoring station. It was found that the most suitable pumping volume dataset for predicting groundwater level variations at the Hexing groundwater monitoring station is within a distance difference of 20 kilometers and a depth difference of 20 meters. Furthermore, based on the RMSE and MAE values within a distance of 20 kilometers and a depth of 20 meters, it was determined that the Support Vector Regression (SVR) model is the most suitable for predicting groundwater level variations at the Hexing groundwater observation station. Therefore, Fig. 8 shows the time series of groundwater levels for the Hexing groundwater monitoring station. The red line represents the actual groundwater level time series, while the blue line represents the groundwater level time series predicted by the SVR model.

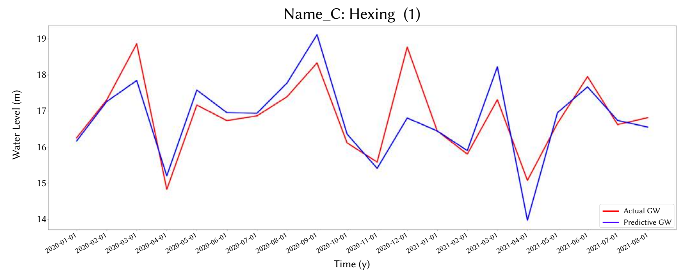


Fig. 8. The actual groundwater level and predicted groundwater level chart for the Hexing groundwater monitoring station.

TABLE VII. Performance Evaluation of Using Three Models to Predict Groundwater Level Variations at Dongguang (1) Station

	MLR		SVR		LSTM	
Sampling range	RMSE	MAE	RMSE	MAE	RMSE	MAE
10KM/10M	1.074	0.852	1.515	1.073	0.738	0.574
10KM/15M	1.585	1.297	1.245	1.021	0.633	0.500
10KM/20M	1.493	1.102	1.281	1.071	0.904	0.686
20KM/15M	1.185	0.887	1.537	1.103	0.875	0.652
20KM/20M	1.483	1.060	1.562	1.272	0.811	0.653

The installation location of the Dongguang groundwater monitoring station is at longitude 120° 27’ 20” and latitude 23° 65’ 19”, at a depth of 33 meters. Table VII represents the prediction accuracy for the Dongguang groundwater observation station. It was found that the most suitable pumping volume dataset for predicting groundwater level variations at the Dongguang groundwater observation station is within a distance difference of 10 kilometers and a depth difference of 15 meters. Furthermore, based on the RMSE and MAE values within a distance of 10 kilometers and a depth of 15 meters, it was determined that the Long Short-Term Memory (LSTM) model is the most suitable for predicting groundwater level variations at the Dongguang groundwater monitoring station. Therefore, Fig. 9 shows the groundwater level time series plot for the Dongguang groundwater monitoring station. The red line represents the actual groundwater level time series, while the blue line represents the groundwater level time series predicted by the LSTM model.

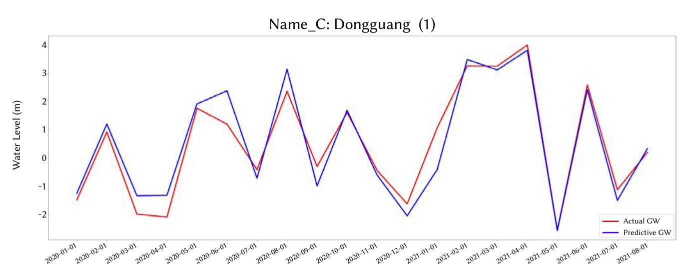


Fig. 9. The actual groundwater level and predicted groundwater level chart for the Dongguang groundwater monitoring station.

TABLE VIII. Performance Evaluation of Using Three Models to Predict Groundwater Level Variations at Wencuo (1) Station

	MLR		SVR		LSTM	
Sampling range	RMSE	MAE	RMSE	MAE	RMSE	MAE
10KM/10M	0.483	0.408	0.466	0.398	0.460	0.351
10KM/15M	0.397	0.331	0.441	0.341	0.441	0.341
10KM/20M	0.493	0.348	0.449	0.373	0.339	0.246
20KM/15M	0.387	0.299	0.395	0.282	0.446	0.322
20KM/20M	0.468	0.364	0.347	0.303	0.390	0.296

The installation location of the Wencuo groundwater monitoring station is at longitude 120° 51' 20" and latitude 23° 65' 77", at a depth of 35.67 meters. Table VIII represents the prediction accuracy for the Wencuo groundwater observation station. It was found that the most suitable pumping volume dataset for predicting groundwater level variations at the Wencuo groundwater monitoring station is within a distance difference of 10 kilometers and a depth difference of 20 meters. Furthermore, based on the RMSE and MAE values within a distance of 10 kilometers and a depth of 20 meters, it was determined that the Long Short-Term Memory (LSTM) model is the most suitable for predicting groundwater level variations at the Wencuo groundwater monitoring station. Therefore, Fig. 10 shows the groundwater level time series plot for the Wencuo groundwater monitoring station. The red line represents the actual groundwater level time series, while the blue line represents the groundwater level time series predicted by the LSTM model.

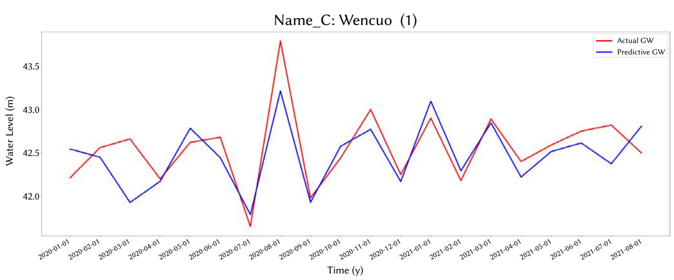


Fig. 10. The actual groundwater level and predicted groundwater level chart for the Wencuo groundwater monitoring station.

TABLE IX. Performance Evaluation of Using Three Models to Predict Groundwater Level Variations at Fengrong (1) Station

	MLR		SVR		LSTM	
Sampling range	RMSE	MAE	RMSE	MAE	RMSE	MAE
10KM/10M	0.693	0.593	0.872	0.704	0.961	0.670
10KM/15M	0.947	0.745	0.770	0.608	0.812	0.558
10KM/20M	0.700	0.510	0.797	0.661	0.672	0.495
20KM/15M	0.922	0.652	0.642	0.492	0.813	0.530
20KM/20M	0.646	0.512	0.700	0.524	0.753	0.577

The installation location of the Fengrong groundwater monitoring station is at longitude 120° 31' 09" and latitude 23° 79' 07", at a depth of 51.82 meters. Table IX represents the prediction accuracy for the Fengrong groundwater monitoring station. It was found that the most suitable pumping volume dataset for predicting groundwater level variations at the Fengrong groundwater observation station is within a distance difference of 10 kilometers and a depth difference of 20 meters. Furthermore, based on the RMSE and MAE values within a distance of 10 kilometers and a depth of 20 meters, it was determined that the Long Short-Term Memory (LSTM) model is the most suitable for predicting groundwater level variations at the Fengrong groundwater observation station. Therefore, Fig. 11 shows the groundwater level time series chart for the Fengrong groundwater observation station. The red line represents the actual groundwater level time series, while the blue line represents the groundwater level time series predicted by the LSTM model.

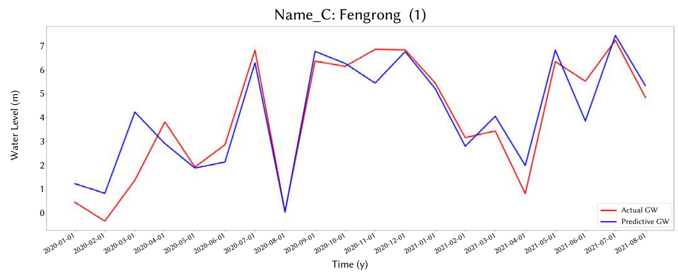


Fig. 11. The actual groundwater level and predicted groundwater level chart for the Fengrong groundwater monitoring station.

Based on the results from Tables 4 to 9, we obtained the most suitable applied pumping dataset for predicting groundwater level variations for each groundwater monitoring station. Consequently, from the optimal execution environment, the RMSE and MAE values from the optimal execution environment were used to identify the best and most stable deep learning model. The RMSE values for the MLR method ranged from 0.2 to 1.6, and the MAE values ranged from 0.18 to 1.3. For the SVR method, the RMSE values ranged from 0.2 to 1.3, and the MAE values ranged from 0.18 to 1.1. For the LSTM method, the RMSE values ranged from 0.28 to 0.7, and the MAE values ranged from 0.2 to 0.5. The findings of this research indicate that the best model for predicting groundwater level variations is the LSTM method, followed by the SVR method, and finally the MLR method.

V. Conclusion

This study applied artificial intelligence techniques to predict groundwater level variations in the CRAF area of Taiwan from 2020 to August 2021. We investigate the performances of the MLR, SVR, and LSTM methods in predicting groundwater levels with limited data. The dataset includes groundwater level and pumping data collected from the CRAF area. The pumping dataset was constructed by extracting pumping well data from the positions and depths around groundwater level monitoring stations as input variables, while the groundwater level data obtained from groundwater level measurement stations serve as output variables. The positions and depths of each groundwater level observation point acted as reference points, and the collection of pumping data at different distances and depths affected the accuracy of predicting groundwater level variations using the MLR, SVR, and LSTM methods. From the results, it was observed that the Guosheng and Hexing groundwater level measurement stations are suitable for executing the groundwater level prediction procedure using pumping data within a radius of 20 kilometers and a depth of 20 meters. In the experimental results, we found that the LSTM model shows stability, strong generalization capabilities, and high prediction accuracy in groundwater level prediction. By comparing the results of applying the best pumping conditions at six groundwater level monitoring stations, it is evident that the MAE and RMSE values of the LSTM method tend to be smaller than those produced by the MLR and SVR methods. Additionally, the LSTM method provides the best predictive models for groundwater level at four groundwater le vel monitoring stations. Therefore, the results of this study will contribute to the planning and management of groundwater resources.

Acknowledgment

This research was supported by the Taiwan Comprehensive University System (TCUS) and the National Science and Technology Council, the Republic of China, under grants NSTC 111-MOEA-M-008-001, NSTC 112-MOEA-M-008-001, and NSTC 111-2410-H-019-006-MY3. The authors thank the Water Resource Agency (WRA) and the Central Geological Survey (CGS) of Taiwan for providing the hydrological and geological data.

References

[1] O. H. Kombo, S. Kumaran, Y. H. Sheikh, A. Bovim, and K. Jayavel, "Long-term groundwater level prediction model based on hybrid KNN-RF technique," *Hydrology*, vol. 7, no. 3, pp. 59, 2020.

[2] D. S. Shih, C. J. Chen, M. H. Li, C. S. Jang, C. M. Chang, and Y. Y. Liao, "Statistical and numerical assessments of groundwater resource subject to excessive pumping: Case study in Southwest Taiwan," *Water*, vol. 11, no. 2, pp. 360, 2019.

[3] M. Z. Ali, H. J. Chu, and T. J. Burbey, "Mapping and predicting subsidence from spatio-temporal regression models of groundwater-drawdown and subsidence observations," *Hydrogeology Journal*, vol. 28, no. 8, pp. 2865-2876, 2020.

[4] H. J. Chu, M. Z. Ali, and T. J. Burbey, "Development of spatially varying groundwater-drawdown functions for land subsidence estimation," *Journal of Hydrology: Regional Studies*, vol. 35, pp. 100808, 2021.

[5] A. Sharafati, S. B. H. S. Asadollah, and A. Neshat, "A new artificial intelligence strategy for predicting the groundwater level over the Rafsanjan aquifer in Iran," *Journal of Hydrology*, vol. 591, pp. 125468, 2020.

[6] C. Lehr, and G. Lischeid, "Efficient screening of groundwater head monitoring data for anthropogenic effects and measurement errors," *Hydrology and Earth System Sciences*, vol. 24, no. 2, pp. 501-513, 2020.

[7] M. Sakizadeh, M. Mohamed, and H. Klammler, "Trend analysis and spatial prediction of groundwater levels using time series forecasting and a novel spatio-temporal method," *Water Resources Management*, vol. 33, pp. 1425-1437, 2019.

[8] S. Sahoo, T. A. Russo, J. Elliott, and I. Foster, "Machine learning algorithms for modeling groundwater level changes in agricultural regions of the US," *Water Resources Research*, vol. 53, no. 5, pp. 3878-3895, 2017.

[9] J. Yin, J. Medellín-Azuara, A. Escriva-Bou, and Z. Liu, "Bayesian machine learning ensemble approach to quantify model uncertainty in predicting groundwater storage change," *Science of The Total Environment*, vol. 769, pp. 144715, 2021.

[10] A. I. A. Osman, A. N. Ahmed, Y. F. Huang, P. Kumar, A. H. Birima, M. Sherif, and A. El-Shafie, "Past, present and perspective methodology for groundwater modeling-based machine learning approaches," *Archives of Computational Methods in Engineering*, vol. 29, no. 6, pp. 3843-3859, 2022.

[11] E. A. Hussein, C. Thron, M. Ghaziasgar, A. Bagula, and M. Vaccari, "Groundwater prediction using machine-learning tools," *Algorithms*, vol. 13, no. 11, pp. 300, 2020.

[12] A. Ahmadi, M. Olyaei, Z. Heydari, M. Emami, A. Zeynolabedin, A. Ghomlaghi, and M. Sadegh, "Groundwater level modeling with machine learning: a systematic review and meta-analysis," *Water*, vol. 14, no. 6, pp. 949, 2022.

[13] Y. Tang, C. Zang, Y. Wei, and M. Jiang, "Data-driven modeling of groundwater level with least-square support vector machine and spatial–temporal analysis," *Geotechnical and Geological Engineering*, vol. 37, no. 3, pp. 1661-1670, 2019.

[14] X. Huang, L. Gao, R. S. Crosbie, N. Zhang, G. Fu, and R. Doble, "Groundwater recharge prediction using linear regression, multi-layer perception network, and deep learning," *Water*, vol. 11, no. 9, pp. 1879, 2019.

[15] M. Panahi, N. Sadhasivam, H. R. Pourghasemi, F. Rezaie, and S. Lee, "Spatial prediction of groundwater potential mapping based on convolutional neural network (CNN) and support vector regression (SVR)," *Journal of Hydrology*, vol. 588, pp. 125033, 2020.

[16] Y. S. Su, and Y. C. Hu, "Applying Cloud Computing and Internet of Things Technologies to Develop a Hydrological and Subsidence Monitoring Platform," *Sens. Mater*, vol. 34, pp. 1313-1321, 2022.

[17] P. Esling, and C. Agon, "Time-series data mining," *ACM Computing Surveys (CSUR)*, vol. 45, no. 1, pp. 1-34, 2012.

[18] F. Huang, Y. Zhang, D. Zhang, and X. Chen, "Environmental groundwater depth for groundwater-dependent terrestrial ecosystems in arid/semiarid regions: A review," *International journal of environmental research and public health*", vol. 16, no. 5, pp. 763, 2019.

[19] D. C. Montgomery, E. A. Peck, and G. G. Vining, Introduction to linear regression analysis: John Wiley & Sons, 2021.

[20] M. Haenlein, and A. M. Kaplan, "A beginner's guide to partial least squares analysis," *Understanding statistics*, vol. 3, no. 4, pp. 283-297, 2004.

[21] K. A. Mogaji, H. S. Lim, and K. Abdullah, "Modeling of groundwater recharge using a multiple linear regression (MLR) recharge model developed from geophysical parameters: a case of groundwater resources management," *Environmental Earth Sciences*, vol. 73, pp. 1217-1230, 2015.

[22] M. Mokarram, "Modeling of multiple regression and multiple linear regressions for prediction of groundwater quality (case study: north of Shiraz)," *Modeling earth systems and environment*, vol. 2, pp. 1-7, 2016.

[23] S. F. Yan, S. E. Yu, Y. B. Wu, D. F. Pan, and J. G. Dong, "Understanding groundwater table using a statistical model," *Water Science and Engineering*, vol. 11, no. 1, pp. 1-7, 2018.

[24] S. Sahoo, and M. K. Jha, "Groundwater-level prediction using multiple linear regression and artificial neural network techniques: a comparative assessment," *Hydrogeology Journal*, vol. 21, no. 8, pp. 1865, 2013.

[25] S. Kouadri, C. B. Pande, B. Panneerselvam, K. N. Moharir, and A. Elbeltagi, "Prediction of irrigation groundwater quality parameters using ANN, LSTM, and MLR models," *Environmental Science and Pollution Research*, pp. 1-25, 2021.

[26] T. Chang, K. Wang, S. Wang, C. Hsu and C. Hsu, "Evaluation of the Groundwater and Irrigation Quality in the Zhuoshui River Alluvial Fan between Wet and Dry Seasons," *Water*, vol. 14, no. 9, pp. 1494, 2022.

[27] P. C. Deka, "Support vector machine applications in the field of hydrology: a review," *Applied soft computing*, vol. 19, pp. 372-386, 2014.

[28] C. A. Ramezani and K. M. Zounemat, "Evaluation of the efficiency of support vector regression, multi-layer perceptron neural network and multivariate linear regression on groundwater level prediction (Case study: Shahrekord Plain)," *Journal of Watershed Management Research*, vol. 8, no. 15, pp. 1–12, 2017.

[29] S. Afrifa, T. Zhang, P. Appiahene and V. Varadarajan, "Mathematical and Machine Learning Models for Groundwater Level Changes: A Systematic Review and Bibliographic Analysis," *Internet*, vol. 14, no. 9, pp. 259, 2022.

[30] A. El Bilali, A. Taleb and Y. Brouziyne, "Groundwater quality forecasting using machine learning algorithms for irrigation purposes," *Agricultural Water Management*, vol. 245, pp. 106625, 2021.

[31] A. Mirarabi, H. R. Nassery, M. Nakhaei, J. Adamowski, A. H. Akbarzadeh and F. Alijani, "Evaluation of data-driven models (SVR and ANN) for groundwater-level prediction in confined and unconfined systems," *Environmental Earth Sciences*, vol. 78, pp. 1-15, 2019.

[32] M. Shafaei and O. Kisi, "Lake level forecasting using wavelet-SVR, wavelet-ANFIS and wavelet-ARMA conjunction models," *Water Resources Management*, vol. 30, pp. 79-97, 2016.

[33] P. Aghelpour, B. Mohammadi and S. M. Biazar, "Long-term monthly average temperature forecasting in some climate types of Iran, using the models SARIMA, SVR, and SVR-FA," *Theoretical and Applied Climatology*, vol. 138, no. 3-4, pp. 1471-1480, 2019.

[34] S. Mozaffari, S. Javadi, H. K. Moghaddam and T. O. Randhir, "Forecasting groundwater levels using a hybrid of support vector regression and particle swarm optimization," *Water Resources Management*, vol. 36, no. 6, pp. 1955-1972, 2022.

[35] H. J. Chu and T. J. Burbey, "Estimating future (next-month's) spatial groundwater response from current regional pumping and precipitation rates," *Journal of Hydrology*, vol. 604, pp. 127160, 2022.

[36] O. H. Kombo, S. Kumaran and A. Bovim, "Design and Application of a Low-Cost, Low- Power, LoRa-GSM, IoT Enabled System for Monitoring of Groundwater Resources With Energy Harvesting Integration," *IEEE Access*, vol. 9, pp. 128417-128433, 2021.

[37] O. I. Abiodun et al., "Comprehensive Review of Artificial Neural Network Applications to Pattern Recognition," *IEEE Access*, vol. 7, pp. 158820-158846, 2019.

[38] W. Huang, G. Song, H. Hong and K. Xie, "Deep Architecture for Traffic Flow Prediction: Deep Belief Networks With Multitask Learning," *IEEE*

Transactions on Intelligent Transportation Systems, vol. 15, no. 5, pp. 2191-2201, Oct 2014.

[39] G. E. Hinton, S. Osindero and Y. W. Teh, "A fast learning algorithm for deep belief nets," *Neural computation*, vol. 18, no. 7, pp. 1527-1554, 2006.

[40] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu and Y. Zhang, "Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 841-851, Jan 2019.

[41] Z. Zhao, W. Chen, X. Wu, P. C. Chen and J. Liu, "LSTM network: a deep learning approach for short-term traffic forecast," *IET Intelligent Transport Systems*, vol. 11, no. 2, pp. 68-75, 2017.

[42] X. H. Le, H. V. Ho, G. Lee and S. Jung, "Application of long short-term memory (LSTM) neural network for flood forecasting," *Water*, vol. 11, no. 7, pp. 1387, 2019.

[43] Y. Bengio, P. Simard and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Transactions on Neural Networks*, vol. 5, no. 2, pp. 157-166, March 1994.

[44] S. Hochreiter, Y. Bengio, P. Frasconi and J. Schmidhuber, "Gradient flow in recurrent nets: the difficulty of learning long-term dependencies," *IEEE press*, 2001.

[45] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735-1780, 1997.

[46] F. A. Gers, J. Schmidhuber and F. Cummins, "Learning to forget: Continual prediction with LSTM," *Neural computation*, vol. 12, no. 10, pp. 2451-2471, 2000.

[47] J. Zhang, Y. Zhu, X. Zhang, M. Ye and J. Yang, "Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas," *Journal of hydrology*, vol. 561, pp. 918-929, 2018.

[48] M. T. Vu, A. Jardani, N. Massei and M. Fournier, "Reconstruction of missing groundwater level data by using Long Short-Term Memory (LSTM) deep neural network," *Journal of Hydrology*, vol. 597, pp. 125776, 2021.

[49] W. C. Hsu, H. C. Chang, K. T. Chang, E. K. Lin, J. K. Liu and Y. A. Liou, "Observing land subsidence and revealing the factors that influence it using a multi-sensor approach in Yunlin County, Taiwan," *Remote sensing*, vol. 7, no. 6, pp. 8202-8223, 2015.

[50] Central Geological Survey, "Groundwater Recharge Geologically Sensitive area Delineation Plan—G0001 Zhuoshui River Alluvial Fan," MOEA., Taipei, Taiwan, 2014.

[51] H. Yoon, Y. Hyun, K. Ha, K. K. Lee and G. B. Kim, "A method to improve the stability and accuracy of ANN-and SVM-based time series models for long-term groundwater level predictions," *Computers & geosciences*, vol. 90, pp. 144-155, 2016.

[52] Republic of China. Water Resources Agency., AbandonedGroundwaterWellLocation, Taipei, 2020. [Online]. Available: https://data.gov.tw/dataset/129316. [Accessed: Aug 24, 2020].

[53] H. J. Chu, C. W. Lin, T. J. Burbey, and M. Z. Ali, "Spatiotemporal analysis of extracted groundwater volumes estimated from electricity consumption," *Groundwater*, vol. 58, no. 6, pp. 962-972, 2020.

[54] P. Sihag, S. Mohsenzadeh Karimi and A. Angelaki, "Random forest, M5P and regression analysis to estimate the field unsaturated hydraulic conductivity," *Applied Water Science*, vol. 9, pp. 1-9, 2019.

[55] M. J. Shin, S. H. Moon, K. G. Kang, D. C. Moon and H. J. Koh, "Analysis of groundwater level variations caused by the changes in groundwater withdrawals using long short-term memory network," *Hydrology*, vol. 7, no. 3, pp. 64, 2020.

[56] H. J. Chu, M. Z. Ali and T. J. Burbey, "Development of spatially varying groundwater-drawdown functions for land subsidence estimation," *Journal of Hydrology: Regional Studies*, vol. 35, pp. 100808, 2021.

[57] N. Le Duy, T. V. K. Nguyen, D. V. Nguyen, A. T. Tran, H. T. Nguyen, I. Heidbüchel, ... and H. Apel, "Groundwater dynamics in the Vietnamese Mekong Delta: Trends, memory effects, and response times," *Journal of Hydrology: Regional Studies*, vol. 33, pp. 100746, 2021.

[58] B. Ghazi, E. Jeihouni and Z. Kalantari, "Predicting groundwater level fluctuations under climate change scenarios for Tasuj plain, Iran," *Arabian Journal of Geosciences*, vol. 14, pp. 1-12, 2021.

[59] S. M. Guzman, J. O. Paz, M. L. M. Tagert and A. E. Mercer, "Evaluation of seasonally classified inputs for the prediction of daily groundwater levels: NARX networks vs support vector machines," *Environmental Modeling & Assessment*, vol. 24, pp. 223-234, 2019.

[60] F. Kratzert, D. Klotz, G. Shalev, G. Klambauer, S. Hochreiter and G. Nearing, "Benchmarking a catchment-aware long short-term memory network (LSTM) for large-scale hydrological modeling," *Hydrology and Earth System Sciences Discussions*, vol. 2019, pp. 1-32, 2019.

[61] D. Kumar, T. Roshni, A. Singh, M. K. Jha and P. Samui, "Predicting groundwater depth fluctuations using deep learning, extreme learning machine and Gaussian process: a comparative study," *Earth Science Informatics*, vol. 13, pp. 1237-1250, 2020.

Yu-Sheng Su

Yu-Sheng Su is currently an associate professor of department of computer science and information engineering at National Chung Cheng University, Taiwan. His research interests include Cloud Computing, Big Data Analytics, Intelligent System, and Metaverse.

Yu-Cheng Hu

Yu-Cheng Hu received M.S. degree in department of computer science and engineering at National Taiwan Ocean University, Keelung, Taiwan. His research interests include Time Series Analysis, Cloud Computing, Big Data Analysis, and Machine Learning/Deep Learning.

Yun-Chin Wu

Yun-Chin Wu received M.S. degree in department of computer science and engineering at National Taiwan Ocean University, Taiwan. His research interests include AIoT, Cloud Computing, Big Data Analysis, and Machine Learning/Deep Learning.

Ching-Teng Lo

Ching-Teng Lo is a master's degree student in department of computer science and engineering at National Taiwan Ocean University, Taiwan. His research interests include Edge Computing, AIoT, and Generative Artificial Intelligence.

Semi-Supervised Machine Learning Approaches for Thyroid Disease Prediction and its Integration With the Internet of Everything

Melih Agraz

Division of Applied Math, Brown University (USA)

Department of Statistics, Giresun University (Turkey)

* Corresponding author: melih_agraz@brown.edu; melih.agraz@giresun.edu.tr

Received 29 June 2023 | Accepted 7 June 2024 | Early Access 26 July 2024

LA UNIVERSIDAD EN INTERNET

Abstract

Thyroid disorders are critical conditions that considerably affect a person's general health, and may lead to additional health complications. Notably, these conditions often remain undetected in individuals who show "normal" results on traditional thyroid function tests. To enhance the diagnostic accuracy for thyroid disorders, such as hypothyroidism and hyperthyroidism, this study leveraged digital health records and explored semi-supervised learning methods. We intentionally removed the labels from subjects initially categorized as "normal," incorporating them into our dataset as unlabeled data. The goal was to overcome the limitations of conventional diagnostic techniques, which may fail to detect subtle imbalances in thyroid hormones. In pursuit of this objective, we employed a combination of semi-supervised learning methods, namely FixMatch, Co-training, and self-training, in conjunction with supervised learning algorithms, specifically Naive Bayes and logistic regression. Our findings indicate that the FixMatch algorithm surpassed traditional supervised learning methods in various metrics, including accuracy (0.9054), sensitivity (0.9494), negative predictive value (0.9365), and F1 score (0.9146). Additionally, we propose a framework for integrating these diagnostic tools into the Internet of Everything (IoE) to promote early detection and facilitate improved healthcare outcomes. This research highlights the potential of semi-supervised learning techniques in the diagnosis of thyroid disorders and offers a roadmap for harnessing the IoE in healthcare advancement.

Keywords

FixMatch, IoE Medical Systems, Machine Learning, Thyroid Disease.

DOI: 10.9781/ijimai.2024.07.006

I. Introduction

A. Research Background and Motivations

The Internet of Everything (IoE) is increasingly recognized as the future of the internet. It enables intelligent interconnections among diverse elements, including things (devices), processes, people, and data, as illustrated in Fig. 1. This concept harnesses advanced technologies like 5G and artificial intelligence (AI) to enhance internet connectivity, making it faster, smarter, and more tailored to individual needs [1], [2]. Furthermore, the advancement in smart devices has given rise to a new dimension of the IoE, emphasizing the connection of all devices to the internet [3].

The Internet of Everything (IoE) holds the potential to revolutionize numerous industries and various aspects of daily life. For instance, it can be utilized to enhance transportation efficiency by integrating vehicles, traffic signals, and road sensors. Similarly, IoE can aid in improving energy efficiency through the interconnectedness of buildings, appliances, and power grids. Moreover, it has the capability to transform healthcare by establishing connections among patients, doctors, and medical devices. This study discusses the development of machine learning methods, specifically semi-supervised learning (SSL) and supervised learning (SL) algorithms, for monitoring and detecting thyroid diseases. Additionally, it delves into the integration of these methods with the IoE, highlighting their potential synergistic benefits.

B. Research Objectives

In this study, we aimed to investigate several critical questions: Can we predict subtypes of thyoid diseases? Can individuals with normal thyroid levels still be suffering from thyroid disorders? Is it possible to use pseudo-labels to develop robust machine learning models for early detection? Beyond addressing these medical queries, we are also proposing IoE models for researchers in the field of thyroid diseases. Specifically, we explore how pseudo-labels can enhance

Please cite this article as:
M. Agraz. Semi-Supervised Machine Learning Approaches for Thyroid Disease Prediction and its Integration With the Internet of Everything, International Journal of Interactive Multimedia and Artificial Intelligence, vol. 8, no. 7, pp. 38-47, 2024, http://dx.doi.org/10.9781/ijimai.2024.07.006

the effectiveness of machine learning models. As a result, this study focuses on predicting hypothyroid and hyperthyroid diseases in thyroid patients. We utilized the UC Irvine (UCI) thyroid dataset [4], comprising three classes (hypo/hyper/normal), for this prediction. To consider the possibility of individuals with normal thyroid function actually being euthyroid, we treated them as unlabeled data in our analysis i.e., we removed 'normal' labels, we introduced pseudo-labels during the semi-supervised learning process. This approach allowed us to validate the model's performance by comparing predicted labels with the true class distributions identified through subsequent validation steps. Our comprehensive study employs both SL techniques, including Naive Bayes and Logistic Regression, and SSL approaches, such as Fixmatch, Co-training, and self-training. We conducted experiments using both labeled (hypo/hyper) and unlabeled data to evaluate the effectiveness of our proposed SSL methods in comparison with traditional SL methods. Additionally, we encourage researchers to utilize our findings in their IoE research studies, as the insights gained could be pivotal in developing early detection tools for thyroid disorders.

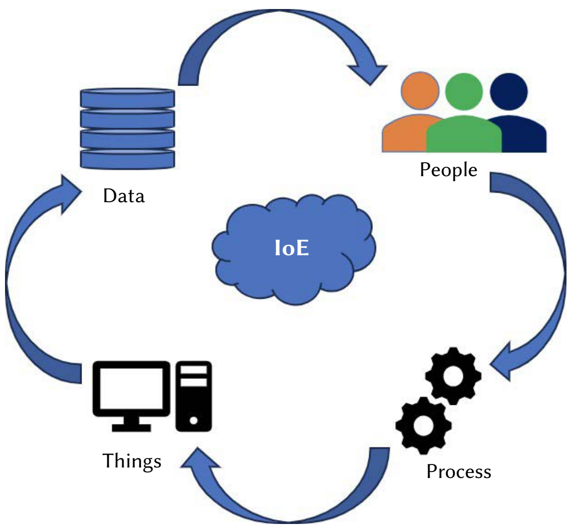


Fig. 1. Integral Elements of the Internet of Everything (IoE). This illustration depicts the interconnected nature of IoE, highlighting the continuous interaction among four critical components: Data, People, Processes, and Things. It emphasizes the dynamic flow of information and the synergy that fuels the functionality of IoE ecosystems.

The structure of the study is as follows: Section II presents a literature review. Section III explains the workflow of thyroid prediction and the proposed IoE application for thyroid disorders. Section IV details the data, features, machine learning methods used in the study, and performance measures, including a subsection for Discussion where the results are analyzed. Finally, Section V discusses the conclusions and future directions of this research.

II. Literature Review

Numerous studies have combined the Internet of Everything (IoE) with medicine [5]–[10]. For instance, Roy and Singh [9] explored the use of the Internet of Medical Things (IoMT) for remote patient monitoring during the COVID-19 pandemic. Their machine learning model achieved a remarkable F1-score of 0.997 in predicting infections. Ahamed et al. [10] developed a Cardiovascular Disease Prediction System using IoT and machine learning for early diagnosis of heart diseases, achieving notable effectiveness with a hyperparameter-tuned Random Forest model. Sivaparthipan et al. [8] proposed a robotic system employing machine learning for treating Parkinson's disease, focusing on processing large datasets to predict patient mobility patterns. In their study on predictive analytics, Rghioui et al. [5] introduced a unique system for monitoring diabetes patients. They evaluated four different machine learning algorithms (Naive Bayes, Random Forest, OneR, and J48), ultimately selecting Random Forest as the most effective prediction algorithm. Maghawry et al. [6]'s study allowed for the development of advisory systems that combine biosensor data with historical medical and social network data to provide accurate alerts and recommendations for various diseases. Additionally, Raja and Chakraborty [7] enhanced medical access in remote areas by using wearable sensors to collect health metrics, stored in cloud storage, and monitored by doctors for effective treatment.

Thyroid gland disorders are significant and frequently encountered conditions that are often overlooked in clinical diagnoses. In areas where goiters are common, about 15 to 30% of adults suffer from this issue [11], [12]. Based on hormonal levels, individuals with thyroid dysfunction are categorized into hypothyroidism, euthyroidism, and hyperthyroidism [13]. These disorders are typically associated with either excessive (hyper) or insufficient (hypo) secretion of thyroid hormones. Factors such as previous thyroid surgery, ionizing radiation exposure, chronic thyroid inflammation, iodine deficiency, enzyme deficiencies, and certain medications can lead to hypothyroidism [14]. Graves' disease is a common cause of hypothyroidism, characterized by the body's production of proteins that stimulate excessive thyroid hormone production [14], [15]. Euthyroidism refers to a state of normal thyroid hormone production and serum levels [16], [17]. In this context, understanding thyroid gland functional data is crucial for accurate diagnosis. Ozyılmaz and Yıldırım [14] highlighted the importance of interpreting this data in diagnosing thyroid diseases, demonstrating the effectiveness of feedforward neural network structures. There have been numerous studies in machine learning to explore both SL [10], [14], [18]–[21] and SSL [22], [22]–[26] approaches. Keramidas et al. [18] proposed a k-nearest neighborhood (k-NN) algorithm to detect thyroid nodules in ultrasound images. Razia [19] introduced a model using both unsupervised and SL methods for thyroid disease diagnosis. Zhang et al. [24] developed a semi-supervised graph convolutional deep learning model, Semi-GCNs-DA, for cross-device adaptation in identifying thyroid nodules. Turk et al. [23] suggested semi-supervised methods for detecting thyroid nodules in ultrasound data, proposing an encoder-based neural network model with high recall and sensitivity. Yang et al. [25] developed a dual-path semi-supervised conditional generative adversarial network (DScGAN) model for thyroid nodule detection, demonstrating its effectiveness in SSL with limited medical datasets and insufficient labels. Requena et al. [26] introduced an innovative SSL approach using Encoder-Decoder Convolutional Neural Networks for Human Activity Recognition (HAR) in healthcare. This method combines public labeled and extensive private unlabeled raw sensor data, enhancing the model's generalization to real-world scenarios. They demonstrated its effectiveness in a case study involving overweight patients, accurately classifying movement patterns from large volumes of accelerometer data. Martin et al. [22] addressed the challenge of selecting an appropriate distance metric for clustering algorithms like k-means. They introduced a semi-supervised clustering algorithm that learns a linear combination of multiple dissimilarities, utilizing incomplete knowledge through pairwise constraints. Enhanced with a regularization term to prevent overfitting, this method showed superior performance in identifying tumor samples using gene expression profiles, outperforming standard semi-supervised techniques and those relying on a single dissimilarity, particularly in noisy environments.

In this study, we achieved the highest accuracy of 90.54% using the FixMatch method, which incorporates several improvements from previous studies. For instance, Razia et al. [27]achieved an accuracy of 73.29% using a combination of SVM, Multiple Linear Regression, Naïve Bayes, and Decision Trees with 22 attributes. Chaubey et al. [28] reached 88.54% using Decision Trees, KNN, and Logistic Regression with a reduced attribute set of 5. Most recently, Peya et al. [29] reported an impressive accuracy of 95.85% using KNN, Naïve Bayes, and Decision Trees with 22 attributes. Our use of the semi-supervised learning approach, particularly FixMatch, allows for significant improvements in scenarios where labeled data is limited, thereby providing a robust solution in resource-constrained settings.

Despite significant advancements in combining IoE with medical applications, current studies predominantly rely on fully supervised learning methods, which require extensive labeled data. This limitation restricts their applicability in scenarios with limited labeled data. Moreover, few studies have explored SSL techniques, which can leverage unlabeled data to improve diagnostic accuracy. Additionally, the integration of SSL methods with IoE frameworks remains underexplored, particularly in the context of real-time monitoring and early detection of thyroid diseases. These gaps highlight the need for research that combines SSL methods with IoE to enhance diagnostic accuracy and enable continuous health monitoring.

III. Research Methodology

This study established a comprehensive research model aimed at addressing the challenges associated with thyroid disease prediction. Our methodology encompasses experimental design, performance evaluation, and the implementation of various machine learning techniques.

Workflow of the Thyroid Prediction Study. The study's workflow involved a comparative analysis of SL and SSL models, as depicted in Fig. 2. In this process, we discarded the normal labels from the thyroid dataset, which comprises three labels (hypo/hyper/normal), resulting in the creation of unlabeled data. Consequently, this left us with labeled data (hypo/hyper) and unlabeled data. The SL approach was exclusively applied to the labeled data, while the SSL approach was utilized for both labeled and unlabeled data. We compared the performance of these two methods to assess their effectiveness. To address the issue of significant imbalance within the model, random undersampling was conducted during the training phase to balance the training data. The machine learning methods employed in the SL approach included Logistic Regression and Naive Bayes. In contrast, the SSL approach incorporated FixMatch, Co-training, and self-training techniques.

Workflow of the Proposed Internet of Everything Thyroid Study. While our current machine learning model operates with laboratory data (electronic health records in the Fig. 3), we propose a future-oriented model that integrates data from wearable devices, as depicted in Fig. 3.

1. **Continuous Monitoring of Thyroid-Related Physiological Parameters**: Our envisioned device is designed for the ongoing monitoring of various physiological indicators critical to thyroid function. These include heart rate, body temperature, and physical activity levels. This continuous data stream provides a solid foundation for in-depth analysis, making it an effective tool for real-time health tracking.
2. **Data-Driven Analysis through Machine Learning Algorithms**: The device goes beyond simple data gathering. It incorporates sophisticated machine learning algorithms to analyze and interpret the collected data. These algorithms are specifically tailored to evaluate the thyroid's functional status, offering an automated yet highly accurate analysis.
3. **Immediate Alerting Mechanisms for Anomalies**: In case of detection of abnormalities indicative of thyroid dysfunction, the device promptly issues alerts to the user. This rapid response facilitates early lifestyle adjustments or urgent medical consultations, helping to prevent potential health complications.
4. **Secure Data Management via Cloud Infrastructure**: Recognizing the paramount importance of data security in healthcare, our system employs robust cloud storage for data management. This approach guarantees the safety and privacy of user data and supports long-term health monitoring and analysis.
5. **User Interface Designed for Accessibility and Customization**: To enhance user interaction, the system features a user-friendly mobile app and web dashboard. These interfaces not only provide access to real-time and historical data but also enable users to add additional information, like medication and symptoms. This extra layer of data enriches the precision and dependability of the predictive models used by the device.

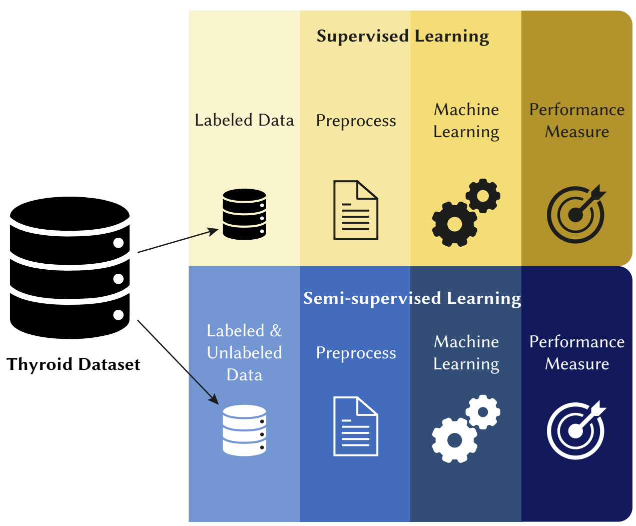


Fig. 2. Comparative Main Workflow of Supervised vs. Semi-Supervised Learning Using the Thyroid Dataset. This flowchart illustrates the study's methodology, highlighting the differences in data preprocessing, machine learning application, and performance measurement stages between the supervised and semi-supervised learning paradigms when applied to thyroid health data.

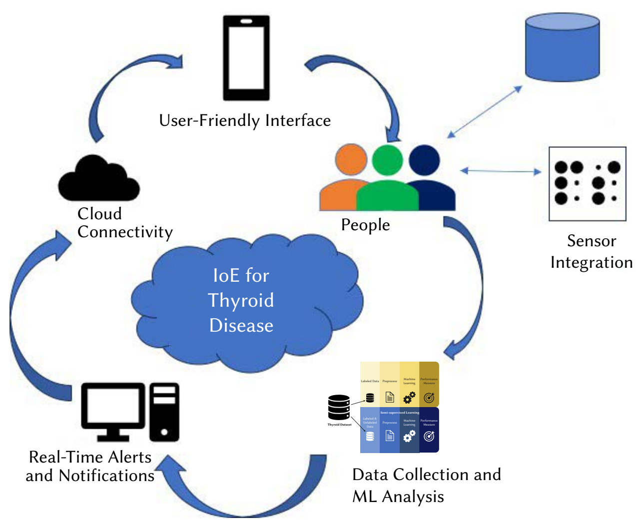


Fig. 3. Workflow of the IoE Framework for Thyroid Disease Management. This diagram presents a comprehensive IoE strategy, highlighting the essential steps from sensor- based data acquisition, through machine learning-powered analysis, to timely alerts and notifications, ultimately leading to a user-centric interface enhanced by cloud technology.

IV. Experimental Design and Performance Evaluation

A. Dataset Collection

In this study, we leveraged the thyroid disease dataset, freely accessible to the public, from the UCI repository [4]. This dataset comprises 10 distinct databases, all sourced from the Garvan Institute in Sydney, Australia. Our specific focus was on the thyroid-related data within this collection of datasets available on the UCI repository [4]. Additionally, users can access the data at Github repository at https://github.com/melihagraz/SSL_AND_SL/tree/main/data.

The dataset offers comprehensive information on 7,200 patients and encompasses 21 features, as outlined in Table I. Patients are categorized into three groups: normal (1), hyperthyroidism (2), and hypothyroidism (3). The distribution of these categories is as follows: 166 patients are classified as normal, 368 as having hyperthyroidism, and 6666 as having hypothyroidism, as illustrated in Fig. 4. Regarding data preprocessing, the dataset was found to be complete, requiring no imputation for missing values. Additionally, we retained all observations, and we worked with the normalized data.

1. Data Security and Ethics

While the IoE offers unprecedented opportunities for real-time data collection and analysis in thyroid disease research, it also poses challenges in terms of data compatibility and scalability. Existing medical databases may require substantial modifications to integrate with IoE devices. Additionally, the sheer volume of data generated could potentially overwhelm current data storage solutions.

The growing digitalization of healthcare raises a variety of ethical opportunities and challenges [30]. As noted by Jacquemard et al. [31], ethical values should inform all stages of the electronic health records lifecycle, from design and development to implementation and practical application. These ethical considerations become particularly important when discussing the integration of the IoE with medical data. Obtaining informed consent for data collection becomes increasingly complex, especially when dealing with sensitive health information. Moreover, there is a risk that the benefits of this advanced technology may not be equally accessible across all populations, potentially exacerbating existing healthcare inequalities.

Data security is a critical concern when integrating IoE technologies into medical research. This integration involves the transmission and storage of sensitive personal health information, which may be vulnerable to unauthorized access or misuse [32]. The expanded network of interconnected devices amplifies the potential points of failure or unauthorized access, making the protection of sensitive health data increasingly challenging. Therefore, rigorous encryption methods and robust cybersecurity protocols are essential to mitigate these risks.

2. Potential Challenges

There could be potential challenges of IoE with thyroid disease research.

1. Technical Challenges: Integrating IoE in thyroid disease research involves handling various types of devices and data sources. Ensuring compatibility, interoperability, and continuous data flow between these devices and systems can be complex. There may also be issues related to data accuracy, reliability, and real-time processing.
2. Data Management and Analysis: The sheer volume of data generated by IoE devices poses a significant challenge. Efficiently storing, managing, and analyzing this data to extract meaningful insights for thyroid disease research requires advanced data analytics tools and expertise in big data management.

TABLE I. Dataset Variables Overview. This Table Lists the Features of the Thyroid UCI Dataset, Including Variable Names, Their Descriptions, and Classification as Continuous or Discrete Data Types. The Dataset Encompasses Demographic Details, Treatment History, and Clinical Test Results, Which Are Pivotal for Thyroid Disease Analysis [4]

Variable	Description	Data types
Age	Age of the patient	Continuous
Sex	Sex of the patient	Discrete
On_thyroxine	On thyroxine	Discrete
Query_on_thyroxine	Query on thyroxine	Discrete
On_antithyroid _medication	On antithyroid medication	Discrete
Sick	Sick	Discrete
Pregnant	Pregnant	Discrete
Thyroid_surgery	Thyroid surgery	Discrete
I131_treatment	I131 treatment	Discrete
Query_hypothyroid	Query hypothyroid	Discrete
Query_hyperthyroid	Query hyperthyroid	Discrete
Lithium	Lithium	Discrete
Goitre	Goitre	Discrete
Tumor	Tumor	Discrete
Hypopituitary	Hypopituitary	Discrete
Psych	Psych	Discrete
TSH	Amount of TSH	Continuous
T3	Amount of T3	Continuous
TT4	Amount of TT4	Continuous
T4U	Amount of T4U	Continuous
FTI	Amount of FTI	Continuous

3. Infrastructure and Cost: Setting up the infrastructure for IoE integration in thyroid disease research can be costly. This includes the cost of devices, data storage solutions, analytics software, and securing the network. Additionally, there is a need for continuous maintenance and upgrades.
4. User Acceptance and Training: For IoE to be effective in research, it's essential that all collaborators (researchers, clinicians, patients) understand and accept the technology. This might require extensive training and education to ensure proper usage and interpretation of IoE-generated data.

B. Outcome and Features

1. Outcome

In this study, the output of the raw data has three labels, due to the possibility of hypo/hyperthyroidism in normal individuals, the labels of individuals with a normal condition were removed. Therefore, only labeled (hypo/hyper) for SL analysis and labeled and unlabeled data for SSL analysis were used.

2. Features

The analyzes performed in this study took advantage of the comprehensive set of features listed in Table I, allowing for a more comprehensive and in-depth exploration of the research objectives. In addition, considering that the dataset is sufficient, feature selection or feature reduction methods were not applied.

We can explain the variables and classify them as the following format.

- **Patient Details**
  - **Age:** Age of the patient. This is a continuous variable, meaning it can take any numerical value within a given range.
  - **Sex:** Sex of the patient (usually Male or Female). This is a discrete variable, meaning it can only take certain predefined values.
- **Treatment and Medication**
  - **On_thyroxine:** Indicates whether the patient is on thyroxine medication. Discrete variable (usually Yes or No).
  - **On_antithyroid_medication:** Specifies if the patient is taking medication for thyroid issues. Discrete variable (usually Yes or No).
- **Health Conditions**
  - **Sick:** Indicates whether the patient is currently sick. Discrete variable (usually Yes or No).
  - **Pregnant:** Indicates whether the patient is pregnant. Discrete variable (usually Yes or No).
  - **Thyroid_surgery:** Indicates whether the patient has had thyroid surgery. Discrete variable (usually Yes or No).
  - **I131_treatment:** Indicates whether the patient has undergone Iodine-131 treatment, usually for thyroid cancer or hyperthyroidism. Discrete variable (usually Yes or No).
- **Queries and Concerns**
  - **Query_hypothyroid:** Indicates whether there is a question or query about the patient having hypothyroidism. Discrete variable (usually Yes or No).
  - **Query_hyperthyroid:** Indicates whether there is a question or query about the patient having hyperthyroidism. Discrete variable (usually Yes or No).
  - **Query_on_thyroxine:** Indicates whether there is a question or query about the patient being on thyroxine. Discrete variable (usually Yes or No).
- **Other Medications and Conditions**
  - **Lithium:** Indicates whether the patient is on lithium, which can affect the thyroid. Discrete variable (usually Yes or No).
  - **Goitre:** Indicates whether the patient has a goitre (enlarged thyroid). Discrete variable (usually Yes or No).
  - **Tumor:** Indicates whether the patient has a tumor. Discrete variable (usually Yes or No).
  - **Hypopituitary:** Indicates whether the patient has hypopituitarism (underactive pituitary gland). Discrete variable (usually Yes or No).
  - **Psych:** Likely indicates whether the patient has a psychiatric condition. Discrete variable (usually Yes or No).
- **Hormone Levels**
  - **TSH (Thyroid-Stimulating Hormone):** Amount of TSH in the blood. Continuous variable.
  - **T3 (Triiodothyronine):** Amount of T3 in the blood. Continuous variable.
  - **TT4 (Total Thyroxine):** Amount of TT4 in the blood. Continuous variable.
  - **T4U (Thyroxine Uptake):** Amount of T4U in the blood. Continuous variable.
  - **FTI (Free Thyroxine Index):** Amount of FTI in the blood. Continuous variable.

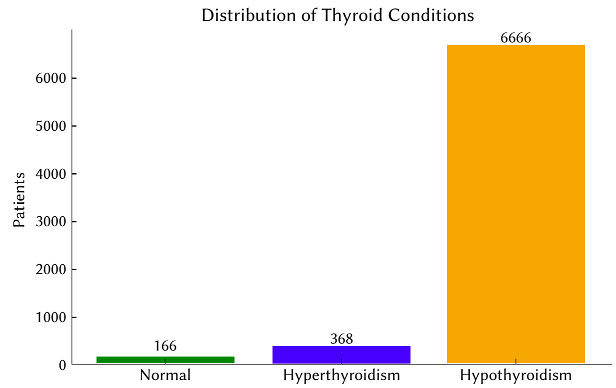


Fig. 4. Distribution of Patient Diagnoses in the Thyroid Dataset. The bar chart presents the number of patients categorized into three groups: those diagnosed with hyperthyroidism, hypothyroidism, and normal (unlabeled) labels group.

C. Data Imbalance

We unlabeled normal patients from the thyroid data, and we observed an imbalance in the ratio of Hypo/Hyperthyroid patients, with a ratio of 1:18. To address this imbalance, random undersampling was performed during the training phase by randomly selecting samples from the majority class, resulting in a balanced ratio of 1:1 in the training of the machine learning algorithm.

D. Machine Learning Methods

1. Supervised Learning

Supervised learning (SL) is a widely used learning method in the field of data science. In this method, a machine learning model learns to predict the associated output value using a given input data. In this study, the following two SL methods were preferred.

1. Logistic regression.
2. Naive Bayes.

Consensus methods for choosing the type of statistical model have not been established [33]. As a result, we selected logistic regression and Naive Bayes for their respective advantages. Logistic regression is highly valued in medical research for its interpretability, suitability for binary outcomes, ability to adjust for confounders, flexibility, and the appropriateness of predicted probabilities. It is also recognized for its robustness in assessing model fit and accuracy. These attributes make logistic regression an invaluable tool for handling the complexities of medical data effectively [34]. On the other hand, the Naive Bayes algorithm is acclaimed for requiring minimal training data, straightforward computation, ease of implementation, time efficiency, and its ability to handle large datasets and incomplete data. It is also resilient to irrelevant features and data noise [35]–[38]. For these reasons, we chose logistic regression and Naive Bayes. Given the limited variety of models in semi-supervised methods compared to supervised methods, we tried to select effective methods such as FixMatch and self-training among the limited options available.

2. Semi-Supervised Learning

Semi-supervised learning (SSL) is a learning method used in the field of machine learning. This method works on a dataset that includes both labeled (with correct output values) and unlabeled (without correct output values) data samples. In this system, unlabeled data is used either to improve the model's performance or to augment labeled data.

Self-training method. This method serves as the foundational approach for SSL. As illustrated in Fig. 5, the model is initially trained using labeled data, same to the procedure in SL. Subsequently, this trained model is utilized to predict the labels for the unlabeled data.

Observations with a prediction confidence exceeding a predetermined threshold are identified and assigned pseudo-labels. These pseudo-labels are incorporated into the existing pool of labeled data, and the model undergoes retraining. The iterative process persists until a stopping criterion is met, when no more pseudo-labels can be predicted in this study. The final trained model is then used to assess its performance on test data.

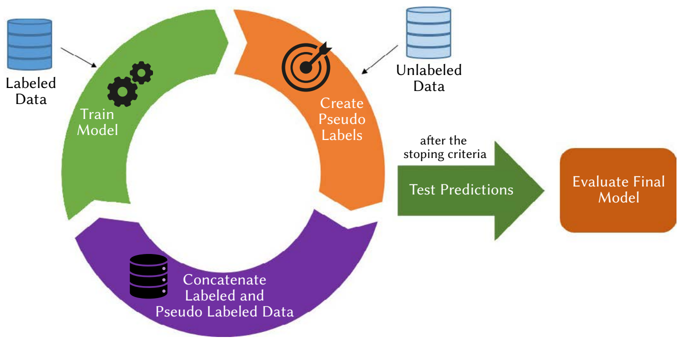


Fig. 5. Overview of the Self-training method workflow. This diagram illustrates the cyclical process of training a model with labeled data, generating pseudo-labels for unlabeled data, testing predictions after meeting the stopping criteria, and evaluating the final model performance, followed by concatenation of labeled and pseudo-labeled data for further iterations.

FixMatch method. FixMatch is one of the SSL methods that focuses on effectively utilizing a small amount of labeled data along with a larger amount of unlabeled data. It trains the model using both weak and strong augmentations, enabling the model to make more general and robust predictions. In FixMatch, “augmentation” refers to data augmentation, which is typically used to increase data diversity. The concept of “weak” and “strong” augmentations is employed in FixMatch. Weak augmentation makes slight modifications to the data, such as rotating or slightly zooming an image, while preserving its basic characteristics. Strong augmentation, on the other hand, introduces more radical changes like cropping, flipping, or significant color adjustments, exposing the model to a greater variety of data. In our approach, we apply weak augmentation to add a low level of Gaussian noise to the data, and strong augmentation to introduce a higher level of Gaussian noise. Conventional SSL methods use both labeled and unlabeled data, while this method encourages consistency between predictions obtained by applying different data augmentations on the same data points. This method, developed by Sohn et al. [39], is a highly effective SSL method that enables the learning of the system even with a single labeled data point. It has been proven to be more effective than other SSL methods, and it is used to build high-accuracy models with a limited number of labeled data.

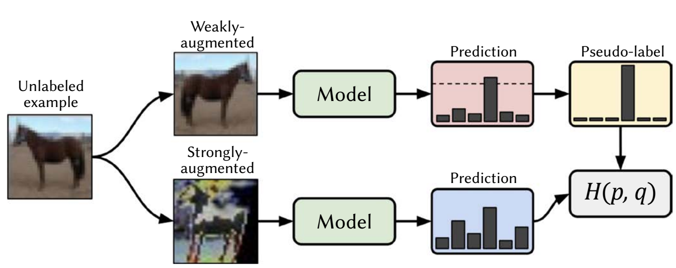


Fig. 6. Schematic representation of the FixMatch algorithm applied in semi-supervised learning [39].

The FixMatch method, as applied in our study of thyroid tabulated data [39], consists of the following steps, as seen in the general form of FixMatch in Fig. 6 and the adaptation of the FixMatch method for thyroid disease prediction in Fig. 7:

1. **Data Augmentation Function:** The most critical aspect of FixMatch is the data augmentation phase. Unlabeled data are augmented using two distinct strategies: strong and weak augmentation. Strong augmentation aims to make the data more challenging to learn, while weak augmentation aims to simplify the learning process.
2. **Training on Labeled Data:** The model is initially trained on labeled data, and the loss, denoted as L_s, is calculated.
3. **Pseudo-labeling:** The $H(p, q)$ (cross entropy loss) component shown in Fig. 6 is generated during this step and the subsequent one. The classification model processes the weakly augmented data, and values exceeding a predetermined threshold are assigned pseudo-labels.
4. **Loss Calculation:** Data that were strongly augmented are processed through the classification model, and the loss, L_{us}, is calculated. This loss is computed using data identified in the previous step as pseudo-labels. The total loss, L, is computed as the sum of L_s and L_{us} as seen in Fig. 7.
5. **Iteration:** Steps 2-4 are iteratively repeated until the model converges.

Additionally, we updated the working schema as represented in Fig. 7. Originally, the FixMatch method was developed for image data. However, in our study, we adapted it for tabulated data by employing data augmentation techniques based on a Gaussian distribution, which could be another novelty of the study.

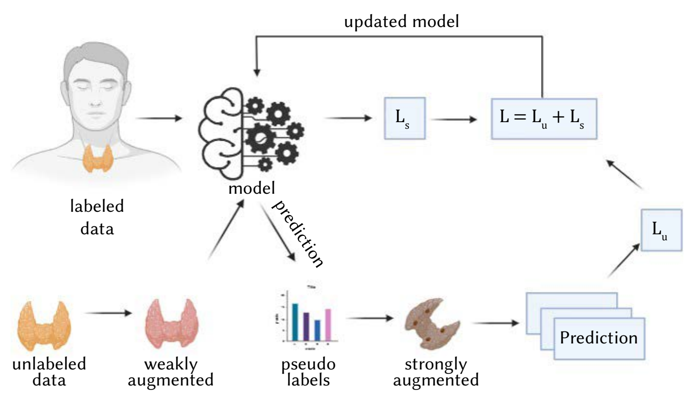


Fig. 7. Adaptation of the FixMatch method for thyroid disease prediction, depicting the process from labeled data through model refinement to loss calculation, highlighting the semi-supervised learning cycle.

Co-Traininig method. The Co-Training method, first proposed by Blum [40], is a semi-supervised machine learning technique that involves splitting labeled data and generating the most effective pseudo-labels from unlabeled data through two different models as represented in Fig. 8. According to the Fig. 8, the data is first split into two distinct views, X1 and X2, each having the same outputs. Each set of view is trained using a separate classifier (Classifier 1 (M1) and Classifier 2 (M2)). After the training, these classifiers are applied to unlabeled data and pseudo-labels are predicted. The common pseudo-labels selected from M1 and M2 are identified and added to the labeled views X1 and X2. The model repeats this process either until a predetermined number of iterations is reached or until there is no more unlabeled data. In this study, we have set the number of iterations to 30, following Blum’s approach [40]. In some cases, SSL is used to augment unlabeled data, while in others, it enhances the performance of labeled data. In this study, it is preferred to increase the performance of the algorithms. Performance measures of the final iterative results are listed. This study explores the following structure.

TABLE II. Performance Metrics of Machine Learning Algorithms. Co-Training M1: Co-Training Classifier1, Co-Training M2: Co-Training Classifier2, NB: Naive Bayes, LR: Logistic Regression; Sens: Sensitivity; Spec: Specificity; PPV: Positive Predictive Value; NPV: Negative Predictive Value; F1: F1-measure; Acc: Accuracy. Best Performance Measures Are Showed in Bold

Metric	FixMatch	Self-training	NB	LR	Co-Training M1	Co-Training M2
Sens	**0.9494**	0.2024	0.0642	0.6794	0.1961	0.2487
Spec	0.8551	0.9543	**0.9865**	0.7978	0.9605	0.9605
PPV	0.8824	0.9883	0.9891	0.7706	0.9886	**0.9910**
NPV	**0.9365**	0.0619	0.0549	0.7133	0.0639	0.0680
F1	**0.9146**	0.2420	0.1201	0.7222	0.3273	0.3976
Acc	**0.9054**	0.3350	0.1126	0.7386	0.2374	0.2871

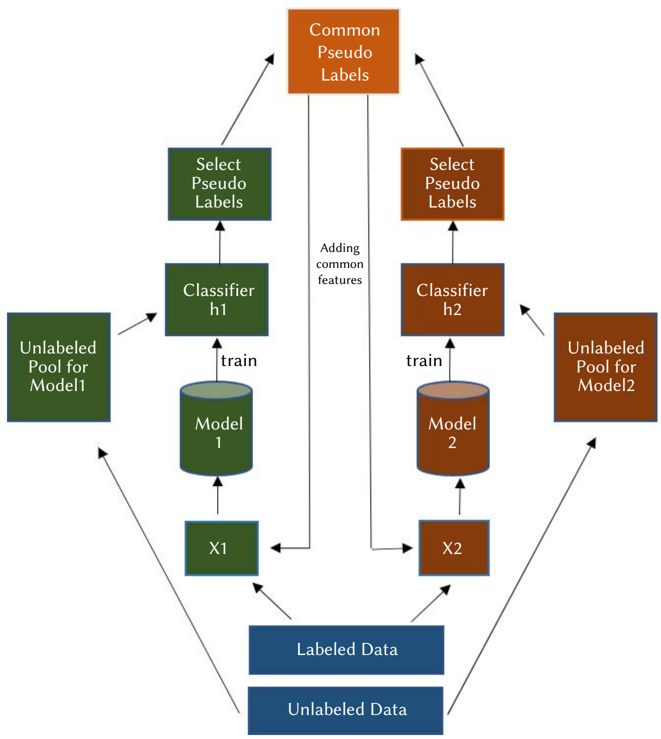


Fig. 8. Workflow of the Co-training method, illustrating the synergistic training of two models on labeled and unlabeled data with a focus on the iterative selection and refinement of pseudo-labels to improve classification performance.

1. **Split Labeled Data**

- Split the labeled data into two parts: a training set and a test set. Use the training set for training the model, while the test set is utilized for assessing its performance.

2. **Prepare Data for Co-training**

- Divide the features of the training set into two subsets views (X1 and X2). Each subset will be used to train a separate model.

3. **Co-training Process**

- Initial Training: Train two separate models, each on a different view of features from the labeled training data.
- Iteration Loop: Repeat the following steps for a set number of iterations (e.g., 30 times):
  - Predict on Unlabeled Data: Use both models to make predictions on the unlabeled data, utilizing their respective feature subsets.
  - Identify Agreements: Find instances where both models agree on the prediction. These are treated as pseudo-labeled data.
  - Update Training Sets: Add the pseudo-labeled instances to the respective training sets of each model, effectively increasing the labeled dataset.
  - Retrain Models: Re-train both models on their updated training sets, which now include the newly pseudo-labeled data.
  - Evaluate Models: At each iteration, evaluate the performance of both models on the separate test set and track the accuracy.

4. **Model Evaluation**

- After the final iteration, assess the performance of the models using the test set. Calculate sensitivity, specificity, positive predictive value, negative predictive value, F1 and accuracy.

E. Performance Evaluation

The core of our study involved comparing the performance of the SL and SSL models, as depicted in Fig. 2. The SL method was applied to labeled data, while the SSL method incorporated both labeled and unlabeled data. Performance metrics were meticulously computed and are outlined in Table II, these include Sens: Sensitivity; Spec: Specificity; PPV: Positive Predictive Value; NPV: Negative Predictive Value; F1: F1-measure; Acc: Accuracy.

F. Discussion

Thyroid disease is a significant condition that can lead to important consequences if left untreated. The thyroid gland, an essential part of the body's endocrine system, is in charge of producing the hormones that regulate metabolism. When thyroid disorders occur, they disrupt the normal functioning of the thyroid gland, resulting in imbalances in hormone production. These imbalances can have a profound impact on individuals' overall health and quality of life. Common symptoms include fatigue, weight changes, mood swings, and cognitive issues [41]. It is crucial to promptly address thyroid disorders as untreated cases can lead to serious complications. Cardiovascular problems, fertility issues, and mental health concerns are among the potential outcomes of untreated or poorly managed thyroid disease. Thyroid disorders can be diagnosed and understood through a combination of laboratory results and imaging techniques. The term "euthyroid" refers to the enlargement of the thyroid gland despite the values obtained from thyroid function tests being within normal limits. It highlights the importance of considering additional factors beyond laboratory results when evaluating thyroid health. Thyroid disorders can also manifest in individuals who have normal results from laboratory tests, emphasizing the need to take into account factors beyond laboratory results when assessing thyroid health.

In this study, we propose the use of SSL methods in conjunction with IoE for the diagnosis of thyroid diseases. Initially, labels of individuals classified as normal were discarded. These individuals were then included in the analysis as unlabeled data in the SSL methods. In the self-training process, pseudo-labels were assigned to this unlabeled data, enabling the identification of individuals with hypo- or hyperthyroidism within the group initially labeled as normal. In the Co-training method, the dataset was split into two different views, and each dataset was trained separately. Common pseudo-labels were created, obtained from each view, and the model's performance was increased. In the FixMatch method, unlabeled data was utilized to enhance the model's performance through data augmentation, specifically by applying Gaussian noise in this study. This approach was aimed at improving the prediction accuracy for hypo/hyperthyroidism using the SSL method.

As shown in Table II and Fig. 9, the SSL method, particularly the FixMatch approach, yielded the highest values for Accuracy, Sensitivity, NPV, and F1 score. Additionally, Naive Bayes provided the highest values for Specificity, and the highest PPV result was obtained from the Co-training method. Accordingly, the best overall results were achieved using the SSL method, especially the FixMatch method.

V. Conclusion

A. Research Contribution

In this study, thyroid disease prediction was performed using SL and SSL methods based on data generated from individuals who do not have thyroid disease. The model performances were compared with the outputs of the SL method. According to the results obtained from SSL, especially in the FixMatch method, high values were achieved for accuracy, sensitivity, NPV and F1 score. However, in the SL methods, Naive Bayes yielded high values for Specificity and the highest PPV was ontained from the Co-training algorithms. In summary, it can be said that the FixMatch method demonstrated better performance in hypo/hyperthyroid prediction by utilizing unlabeled data from normal individuals. For this reason, based on the machine learning analysis the device could generate real-time alerts and notifications to the user, indicating any abnormalities or potential thyroid-related problems and those with euthyroid can be detected, so that the patients labeled as normal can be double checked to see if they have thyroid. In addition, hypo/hyper thyroid diseases of the people can be predicted in the same way, and people could be warned and take precautions beforehand.

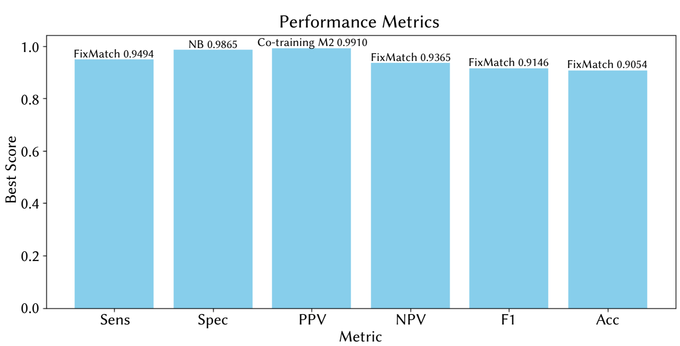


Fig. 9. Best performance measures obtained from the thyroid dataset, displaying the highest scores for each metric: Sensitivity (Sens), Specificity (Spec), Positive Predictive Value (PPV), Negative Predictive Value (NPV), F1 Score (F1), and Accuracy (Acc). The Naive Bayes (NB) algorithm achieved a Spec score of 0.9865, while the Co-Training M2 algorithm showed superior performance in PPV with a score of 0.9910. Fix Match scores were notable for F1 (0.9146), and NPV (0.9365), with an overall Accuracy (Acc) of 0.9054.

B. Future Works and Research Limitations

Although numerous research efforts have investigated the use of ultrasound imaging to detect thyroid nodules, the datasets utilized in these studies are predominantly inaccessible to the public. Accumulating a substantial dataset is hindered by several factors, including time constraints, the specialized nature of medical practices, patient involvement, and the costs associated with ultrasound equipment acquisition [42]. This study aims not only to develop a model that can differentiate between thyroid subtypes but also to leverage labels of cases that are marked as normal yet may have a high likelihood of thyroid disease to enhance model performance. Additionally, we considered integrating this model with an IoT system, initially using publicly available datasets from the UCI repository. Our future goal is to refine our model using sensor data suitable for IoT applications.

Additionally, in our forthcoming research endeavors, our primary emphasis will be on the development of the proposed IoE methodology for thyroid monitoring using wearable devices. Our aim here will be to monitor and detect various thyroid-related diseases, including thyroid cancer. As an essential initial step, we will collect extensive thyroid-related data through these wearable devices, including image datasets. Subsequently, our objective is to delve into the realm of advanced deep learning algorithms.

The choice to explore deep learning algorithms is driven by the complex and multifaceted nature of thyroid data. Deep learning is highly effective at identifying complex relationships and drawing meaningful conclusions from extensive and complicated data sets. With thyroid health being influenced by numerous factors and characterized by subtle variations, deep learning techniques, such as convolutional neural networks (CNNs) [43], [44], long-short term memory (LSTM) [45] and recurrent neural networks (RNNs) [46], hold the potential to uncover hidden relationships within the data.

Moreover, we aim to undertake a comparative analysis between SSL methods, which we have already employed, and deep learning algorithms. This comparative study will allow us to evaluate the efficacy of deep learning in extracting valuable information from our thyroid dataset. By doing so, we hope to gain a comprehensive understanding of which approaches are most suitable for thyroid monitoring, ultimately advancing the field of thyroid healthcare.

Appendix

The SSL and SL code used in this study are available at https://github.com/melihagraz/SSL_AND_SL. Additionally, dataset is reachable at https://github.com/melihagraz/SSL_AND_SL/tree/2af0fecb9adcdd53000a3fac8ef97cc59471b242/data.

References

[1] Q. Yu, J. Ren, Y. Fu, Y. Li, W. Zhang, "Cybertwin: An origin of next generation network architecture," *IEEE Wireless Communications*, vol. 26, no. 6, pp. 111–117, 2019.

[2] S. Abdelwahab, B. Hamdaoui, M. Guizani, A. Rayes, "Enabling smart cloud services through remote sensing: An internet of everything enabler," *IEEE Internet of Things Journal*, vol. 1, no. 3, pp. 276–288, 2014.

[3] M. Deore, U. Kulkarni, "MDFRCNN: Malware detection using faster region proposals convolution neural network," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 7, no. 4, pp. 146–162, 2022.

[4] R. Quinlan, "Thyroid Disease." UCI Machine Learning Repository, 1987. DOI: https://doi.org/10.24432/C5D010.

[5] A. Rghioui, A. Naja, J. L. Mauri, A. Oumnad, "An iot based diabetic patient monitoring system using machine learning and node mcu," in *Journal of Physics: Conference Series*, vol. 1743, 2021, p. 012035, IOP Publishing.

[6] N. E. Maghawry, S. Ghoniemy, "A proposed internet of everything framework for disease prediction," *International Journal of Online & Biomedical Engineering*, vol. 15, no. 4, 2019.

[7] G. B. Raja, C. Chakraborty, "Internet of things based effective wearable healthcare monitoring system for remote areas," in *Implementation of Smart Healthcare Systems using AI, IoT, and Blockchain*, Elsevier, 2023, pp. 193–218.

[8] C. Sivaparthipan, B. A. Muthu, G. Manogaran, B. Maram, R. Sundarasekar, S. Krishnamoorthy, C.-H. Hsu, K. Chandran, "Innovative and efficient method of robotics for helping the parkinson's disease patient using iot in big data analytics," *Transactions on Emerging Telecommunications Technologies*, vol. 31, no. 12, p. e3838, 2020.

[9] A. Singh, A. Kumar, S. Namasudra, "Dnacds: Cloud ioe big data security and accessing scheme based on dna cryptography," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 18, no. 1, pp. 181–801, 2023.

[10] J. Ahamed, A. Manan Koli, K. Ahmad, A. Jamal, B. Gupta, *et al.*, "CDPS-IoT: cardiovascular disease prediction system based on iot using machine learning," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 7, no. 4, pp. 78–86, 2022.

[11] A. Mostbeck, G. Galvan, P. Bauer, O. Eber, K. Atefie, K. Dam, H. Feichtinger, H. Fritzsche, H. Haydl, H. Köhn, *et al.*, "The incidence of hyperthyroidism in austria from 1987 to 1995 before and after an increase in salt iodization in 1990," *European Journal of Nuclear Medicine*, vol. 25, pp. 367–374, 1998.

[12] P. A. Farling, "Thyroid disease," *British Journal of Anaesthesia*, vol. 85, no. 1, pp. 15–28, 2000.

[13] J. Muñoz-Ortiz, M. C. Sierra-Cote, E. Zapata-Bravo, L. Valenzuela-Vallejo, M. A. Marin-Noriega, P. Uribe- Reina, J. P. Terreros-Dorado, M. Gómez-Suarez, K. Arteaga-Rivera, A. De-La-Torre, "Prevalence of hyperthyroidism, hypothyroidism, and euthyroidism in thyroid eye disease: a systematic review of the literature," *Systematic Reviews*, vol. 9, no. 1, pp. 1–12, 2020.

[14] L. Ozyilmaz, T. Yildirim, "Diagnosis of thyroid disease using artificial neural network methods," in *Proceedings of the 9th International Conference on Neural Information Processing, 2002, ICONIP'02*, vol. 4, 2002, pp. 2033–2036, IEEE.

[15] I. Kravets, "Hyperthyroidism: diagnosis and treatment," *American family physician*, vol. 93, no. 5, pp. 363–370, 2016.

[16] J. Kvetny, "The significance of clinical euthyroidism on reference range for thyroid hormones," *European Journal of Internal Medicine*, vol. 14, no. 5, pp. 315–320, 2003.

[17] K. Hedayat, J. Lapraz, "A new approach to biological modeling: Introduction to the biology of functions," *Theory Endobiogeny*, vol. 15, pp. 215–254, 2019.

[18] E. G. Keramidas, D. K. Iakovidis, D. Maroulis, S. Karkanis, "Efficient and effective ultrasound image analysis scheme for thyroid nodule detection," in *Image Analysis and Recognition: 4th International Conference, ICIAR 2007, Montreal, Canada, August 22-24, 2007. Proceedings 4*, 2007, pp. 1052–1060, Springer.

[19] S. Razia, M. Rama Narasingarao, "A neuro computing frame work for thyroid disease diagnosis using machine learning techniques," *Journal of Theoretical and Applied Information Technology*, vol. 95, pp. 1996–2005, 05 2017.

[20] A. Andueza, M. Arco-Osuna, B. Fornés, R. Gonzalez Crespo, J. Martín Álvarez, "Using the statistical machine learning models arima and sarima to measure the impact of covid-19 on official provincial sales of cigarettes in spain," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 8, no. 1, pp. 73–87, 2023.

[21] M. Qasim Gandapur, E. Verdú, "ConvGRU-CNN: Spatiotemporal deep learning for real-world anomaly detection in video surveillance system," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 8, no. 4, pp. 88–95, 2023.

[22] M. Martín Merino, A. J. López Rivero, V. Alonso, M. Vallejo, A. Ferreras, "A clustering algorithm based on an ensemble of dissimilarities: An application in the bioinformatics domain," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 7, no. 6, pp. 6–13, 2022.

[23] G. Turk, M. Ozdemir, R. Zeydan, Y. Turk, Z. Bilgin, E. Zeydan, "On the identification of thyroid nodules using semi-supervised deep learning," *International Journal for Numerical Methods in Biomedical Engineering*, vol. 37, no. 3, p. e3433, 2021.

[24] K. Zhang, Z. Li, C. Cai, J. Liu, D. Xu, C. Fang, P. Huang, Y. Wang, M. Yang, S. Chang, "Semi- supervised graph convolutional networks for the domain adaptive recognition of thyroid nodules in cross-device ultrasound images," *Medical Physics*, vol. 50, pp. 7806–7821, 2023.

[25] W. Yang, J. Zhao, Y. Qiang, X. Yang, Y. Dong, Q. Du, G. Shi, M. B. Zia, "Dscgans: Integrate domain knowledge in training dual-path semi-supervised conditional generative adversarial networks and s3vm for ultrasonography thyroid nodules classification," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part IV 22*, 2019, pp. 558–566, Springer.

[26] S. H. Requena, J. M. G. Nieto, A. Popov, I. N. Delgado, "Human activity recognition from sensorised patient's data in healthcare: A streaming deep learning-based approach," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 8, no. 1, pp. 23–37, 2023.

[27] S. Razia, M. N. Rao, "Machine learning techniques for thyroid disease diagnosis-a review," *Indian Journal of Science and Technology*, vol. 9, no. 28, pp. 1–9, 2016.

[28] G. Chaubey, D. Bisen, S. Arjaria, V. Yadav, "Thyroid disease prediction using machine learning approaches," *National Academy Science Letters*, vol. 44, no. 3, pp. 233–238, 2021.

[29] Z. J. Peya, M. K. N. Chumki, K. M. Zaman, "Predictive analysis for thyroid diseases diagnosis using machine learning," in *2021 International conference on science & contemporary technologies (ICSCT)*, 2021, pp. 1–6, IEEE.

[30] L. Royakkers, J. Timmer, L. Kool, R. Van Est, "Societal and ethical issues of digitization," *Ethics and Information Technology*, vol. 20, pp. 127–142, 2018.

[31] T. Jacquemard, C. P. Doherty, M. B. Fitzsimons, "The anatomy of electronic patient record ethics: a framework to guide design, development, implementation, and use," *BMC Medical Ethics*, vol. 22, no. 1, pp. 1–14, 2021.

[32] K. M. Besher, Z. Subah, M. Z. Ali, "Iot sensor initiated healthcare data security," *IEEE Sensors Journal*, vol. 21, no. 10, pp. 11977–11982, 2020.

[33] M. E. Shipe, S. A. Deppen, F. Farjah, E. L. Grogan, "Developing prediction models for clinical use using logistic regression: an overview," *Journal of thoracic disease*, vol. 11, no. Suppl 4, p. S574, 2019.

[34] P. Schober, T. R. Vetter, "Logistic regression in medical research," *Anesthesia & Analgesia*, vol. 132, no. 2, pp. 365–366, 2021.

[35] S. Joshi, B. Pandey, N. Joshi, "Comparative analysis of naive bayes and j48 classification algorithms," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 5, no. 12, pp. 813– 817, 2015.

[36] R. Rajeswari, K. Juliet, D. Aradhana, "Text classification for student data set using naive bayes classifier and knn classifier," *International Journal of Emerging Trends & Technology in Computer Science*, vol. 43, no. 1, pp. 8– 12, 2017.

[37] B. Bustami, "Penerapan algoritma naive bayes untuk mengklasifikasi data nasabah asuransi," *TECHSI- Jurnal Teknik Informatika*, vol. 5, no. 2, 2013.

[38] H. Muhamad, C. A. Prasojo, N. A. Sugianto, L. Surtiningsih, I. Cholissodin, "Optimasi naïve bayes classifier dengan menggunakan particle swarm optimization pada data iris," *Journal Teknologi Informasi dan Ilmu Komputer*, vol. 4, no. 3, p. 180, 2017.

[39] K. Sohn, D. Berthelot, N. Carlini, Z. Zhang, H. Zhang, C. A. Raffel, E. D. Cubuk, A. Kurakin, C.-L. Li, "Fixmatch: Simplifying semi-supervised learning with consistency and confidence," *Advances in neural information processing systems*, vol. 33, pp. 596–608, 2020.

[40] A. Blum, T. Mitchell, "Combining labeled and unlabeled data with co-training," in *Proceedings of the eleventh annual conference on Computational learning theory*, 1998, pp. 92–100.

[41] G. P. Redmond, "Thyroid dysfunction and women's reproductive health," *Thyroid*, vol. 14, no. 3, Supplement 1, pp. 5–15, 2004.

[42] R. Sharma, G. K. Mahanti, C. Chakraborty, G. Panda, A. Rath, "An iot and deep learning-based smart healthcare framework for thyroid cancer detection," *ACM Transactions on Internet Technology*, 2023.

[43] X. Zhang, V. C. Lee, J. Rong, J. C. Lee, F. Liu, "Deep convolutional neural networks in thyroid disease detection: a multi-classification comparison by ultrasonography and computed tomography," *Computer Methods and Programs in Biomedicine*, vol. 220, p. 106823, 2022.

[44] X. Zhang, V. C. Lee, J. Rong, J. C. Lee, J. Song, F. Liu, "A multi-channel deep convolutional neural network for multi-classifying thyroid diseases," *Computers in Biology and Medicine*, vol. 148, p. 105961, 2022.

[45] H. Lu, M. Wang, W. Zhao, T. Su, J. Yang, "Hyperthyroidism progress prediction with enhanced lstm," in *Web and Big Data: 4th International Joint Conference, APWeb-WAIM 2020, Tianjin, China, September 18-20, 2020, Proceedings, Part II 4*, 2020, pp. 209–217, Springer.

[46] A. Santillan, R. C. Tomas, R. Bangaoil, R. Lopez, M. H. Gomez, A. Fellizar, A. Lim, L. Abanilla, M. C. Ramos, L. Guevarra, *et al.*, "Discrimination of malignant from benign thyroid lesions through neural networks using ftir signals obtained from tissues," *Analytical and Bioanalytical Chemistry*, vol. 413, pp. 2163–2180, 2021.

Melih Agraz

Melih Agraz graduated from Dokuz Eylül University in Türkiye, majoring in Mathematics in the Department of Education. Following that, he obtained his Master's degree from Dokuz Eylül University in the Department of Statistics. Later, he earned his Ph.D. in Statistics from the Middle East Technical University in Türkiye. Subsequently, he embarked on a Fulbright postdoctoral research position at UC Berkeley in the USA. Afterward, he received an offer from Brown University and Beth Israel Hospital. Currently, he is working as an Assistant Professor in the Department of Statistics at Giresun University in Türkiye and has an affiliation with the Brown University Division of Applied Mathematics.

Enhancing Tennis Serve Scoring Efficiency: An AI Deep Learning Approach

Jing-Wei Liu*

Department of Sport Information and Communication, National Taiwan University of Sport, Taichung City, Taiwan (R.O.C.)

* Corresponding author: liujingwei.ntus@gmail.com

Received 7 June 2023 | Accepted 17 May 2024 | Early Access 23 July 2024

Abstract

The playing field of a tennis competition is a dynamic and complex formative environment given the following preliminary knowledge: (a) the basic technical, tactical, situational, and special types of shots used by the opponent; (b) the hitting area of the tennis player; (c) the place of service; (d) the ball drop position; and (d) batting efficiency and other related information that may improve the chances of victory. In this study, we propose an AI classification model for tennis serve scores. Using a deep learning algorithm, the model automatically tracks and classifies the serve scores of professional tennis players from video data. We first defined the players' techniques, volleys, and placements of strokes and serves. Subsequently, we defined the referee's tennis terms and the voice in deciding on a serve score. Finally, we developed a deep learning model to automatically classify the serving position, landing position, and use of tennis techniques. The methodology was applied in the context of 10 matches played by Roger Federer and Rafael Nadal. The proposed deep learning algorithm achieved a 98.27% accuracy in the automatic classification of serve scores, revealing that Nadal outscored Federer by 2.1% in terms of serve-scoring efficiency. These results are expected to facilitate the automatic comparison and classification of shots in future studies, enabling coaches to adjust tactics in a timely manner and thereby improve the chances of winning.

Keywords

Deep Learning, IoT, Markerless Motion Capture, Notational Analysis, Tennis Techniques And Tactics, Video Analysis.

DOI: 10.9781/ijimai.2024.07.003

I. Introduction

There are four major tennis Grand Slam tournaments played around the world: US Open, Australian Open, French Open, and Wimbledon. It is the ultimate dream of professional tennis players to win any of these championships, attaining the highest level of glory in the sport [1]. Unsurprisingly, many videos are generated during these tournaments. Whitson and Horne [2] compared the results of large-scale sports events in Canada and Japan, finding that the analysis of footage showing various sports competitions is associated with commercial interests, diverse entertainment effects, and a large audience base [3]. With developments in software and hardware for image analysis, Nhamo et al. and Keshkar et al. [4]–[5] considered the attendance restrictions for various sports competitions during the COVID-19 pandemic, and employed information technology to share footage analyzed in real time, allowing the audience to interact through the Metaverse technology. Today, athletes increasingly look toward the assistance of multimedia systems to obtain analyses of relevant factors such as athlete habits, movements, sports performance, basic and advanced data, and tactics [6]–[9]. To address the audience's perspective, sports video analyses tend to focus on scene classification or the capture of exciting moments. To achieve this, these analyses provide footage from multiple camera angles, allowing users to quickly receive the desired multimedia data and emotionally invest themselves in live matches [10]–[11]. During competitions, video assistant referees, players, coaches, and referees can request the use of computer technology to ensure a fair decision-making process when addressing disputed situations [12]–[14]. The aforementioned image analysis techniques primarily focus on the real-time transmission and analysis of individual matches, aiming to increase interaction with the audience. Rangasamy et al. [15] compiled deep learning methods and AI-based image analyses, many of which combine computer vision technologies such as human-computer interfaces, handwriting recognition, and speech recognition with big data to analyze an athlete's performance in single and multiple matches.

The analysis of video footage has a direct and profound impact on the outcomes of games, as the player's grip, technical movements, physical strength, and overall hitting style determine their performance on the court. Following an analysis, the player's posture and grip on the racket helps improve control, and the proficiency of technical movements directly relates to the accuracy and speed of hitting the ball [8]. In addition, the athlete's endurance and psyche are also key factors. Good endurance ensures sustained high-level performance in

Please cite this article as:
J. W. Liu. Enhancing Tennis Serve Scoring Efficiency: An AI Deep Learning Approach, International Journal of Interactive Multimedia and Artificial Intelligence, vol. 8, no. 7, pp. 48-54, 2024, http://dx.doi.org/10.9781/ijimai.2024.07.003

the game, while a strong mindset allows players to stay calm under pressure and adapt to changes in opponents. The quality of playing habits directly determines the outcome of a game and is crucial to the player's success on the court. The combat models of tennis players have been widely studied and classified at the academic level, with technical characteristics covering different tactical styles [7]. First, baseline golfers primarily stand at the baseline position and demonstrate superb baseline skills, including steady hitting and excellent endurance, to suppress their opponents. Second, net players focus on approaching the net area to win with fast and precise offensive tactics, demonstrating excellent reaction speed and volleying skills. Finally, all-court players combine the advantages of the baseline and net types by flexibly using the entire court and continuously adjusting tactics. This technical classification provides a theoretical basis for tennis matches that helps players tailor their strategies to combat their opponents.

In this study, we designed a two-dimensional virtual Internet-of-Things (IoT) tracking model by defining action images for commonly used tennis techniques [16]–[17]. The tennis court, represented virtually in two dimensions, is divided into 48 areas. We employed deep learning to analyze 10 matches played by Roger Federer and Rafael Nadal from 2007 to 2019. Each shot played by Federer and Nadal was recorded and classified automatically, enabling the analysis of tennis techniques and player positions. This allowed us to compile the playing styles and habits of both players.

The rest of the paper is structured as follows. In Section 2, we present a literature review of shot classification applied to motion analysis and applications of deep learning in sports. In Section 3, we describe the materials and algorithm used in this study. Our experimental results are presented and discussed in Section 4. Finally, Section 5 concludes the paper.

II. Literature Review

A. Shot Classification Applied to Motion Analysis

Currently, sensors and videos are used for the analysis of movements in tennis. Duan et al. [18] proposed a unified classification framework for sports video shots by extracting 5-10 shots from sports movies and defining semantic shot categories. Using supervised learning, they achieved accuracy rates of 85–95% in the classification of tennis, basketball, volleyball, and football events, videos, and catalogs. Dang et al. [19] proposed a court-line pixel detection method, using the RANSAC linear parameter estimation method to determine the sideline range and subsequently deploying an image tracking system to automatically identify video footage of the four major tennis tournaments, achieving accuracy rates of 96–99%. Connaghan et al. [20] examined the accuracy of sensor recognition and classification in tennis strokes, with seven players trained on the three actions of serves, backhands, and forehands. They achieved recognition rates of 82.5%, 86%, 88%, and 90% for configurations corresponding, respectively, to the accelerometer and gyroscope; accelerometer and magnetometer; gyroscope and magnetometer; and accelerometer, gyroscope, and magnetometer. Raymond et al. [21] employed wearable sensors to identify strike types and rotations. In their study, 17 college-age athletes completed 10 exercises for each action with 5–10 shots for each shot over six games. The statistical results show that the average error of shots and spins was 32.0%. Other studies have also employed sensors to collect data for the purpose of action recognition in tennis [22]–[26].

B. Deep Learning in Sports

Batting techniques in tennis can be classified into four primary categories: serve, draw, slice, and volley [27]. In addition to good technical, physical, tactical, psychological combat plans and strategies, players must have an excellent ability to control the ball in a variety of game scenarios. Furthermore, good playing strategies must rely on continuous simulation, training, and accumulated experience. To date, scholars have primarily integrated video features and design-related algorithms to achieve the content understanding, indexing, annotation, and retrieval of sports videos for the development of automatic referees, as well as technical and tactical analyses.

Voeikov et al. [28] proposed the TTNet neural network model, which employs a high-speed camera to provide online real-time automatic refereeing of table tennis matches, achieving an accuracy of 97.5%. Xu et al. [29] deployed a K-nearest neighbor algorithm to achieve human motion and gesture recognition in table tennis videos. The deep learning process was divided into two stages: semi-supervised video image feature learning, and the supervised optimization of video sequence features. The resulting image recognition method achieved a 1.9% improvement in accuracy over the conventional image capture method. Qiao [30] employed a long short-term memory model to instantly track images of table tennis games for action feature recognition, achieving a maximum recognition accuracy of 89% and a target tracking effect and trajectory prediction accuracy of 90%. Compared with a traditional convolutional neural network (CNN), the model proposed by Qiao achieved a 23.17% improvement in accuracy.

Deep learning has also been used for tennis image recognition. Reno [31] deployed a CNN to learn video data of tennis matches after filtering the background of the game environment to track the ball's landing, achieving an accuracy rate of 98.77%. Ganser et al. [32] developed an automatic classifier for tennis shots. Specifically, they equipped tennis players with wearable sensors, and subsequently analyzed and classified the collected signals using a CNN. Of the 5682 shots collected, 91% were successfully detected and classified. Bastanfard and Amirkhani [33] used a CNN to classify tennis videos with 92% accuracy. Huang [34] proposed the HyperNet CNN model to extract and analyze tennis videos through a loss function, achieving an orientation accuracy of 96.32% and a size accuracy of 91.05%. Sports video analyses are typically conducted with 2D images for identification. Ning and Na [35] employed a dynamic time normalization barycenter averaging algorithm, as well as a K-means clustering algorithm, to analyze 3D dynamic tennis videos with the objective of identifying batting actions, achieving an accuracy of 94.5%. Li et al. [36] proposed a 3D CNN architecture and constructed a 3D video analysis algorithm for tennis videos, reaching an identification accuracy of 94.8%.

Successful tennis techniques must combine power and speed as key factors to win the game. Accordingly, individual skill must be complemented by smart tactics to gain advantage. Therefore, we conducted video and annotation analyses, using deep learning to examine the tactics and habits of two professional players, including their positions, techniques, and shooting points. Our results can be used by players to optimize their strategies during matches and training sessions.

III. Proposed Framework

This research used a recurrent neural network (RNN) to develop a two-stage analysis model, analyze the tactics used by Roger Federer and Rafael Nadal from 2007 to 2019, construct a battle model, and adjust autonomous training and match strategies. To accurately identify the two players' positions, techniques, and scores from video data, we designed a two-stage deep learning algorithm. Prior to the analysis, we preprocessed the data by eliminating extraneous images from the footage. Subsequently, we deployed the RNN algorithm for classification. Through image recognition and deep learning with RNN, the algorithm identifies the ball placement when Nadal and Federer win points, their respective positions on the court, and the tennis techniques they employ. The research process is illustrated in Fig. 1, and the analysis procedures are detailed below.

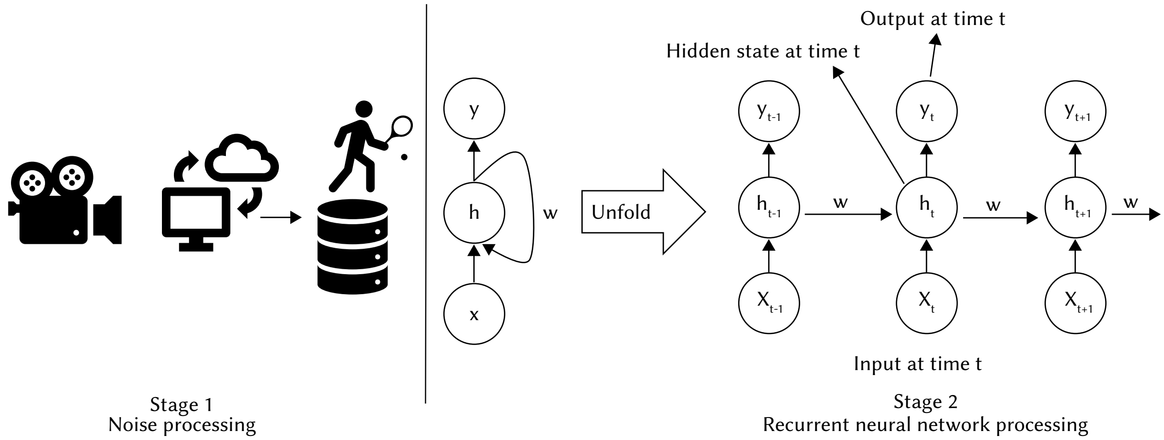


Fig. 1. The research process.

IV. Stage 1: Denoising

Step 1: Defining Noise

Certain scenes in the video data—such as audience reactions, ball retrieval breaks, court maintenance, shots with stationary balls, advertisements, reporter broadcasts, and graphic analyses—were considered noise, as they did not represent actual play. We eliminated these scenes prior to training the RNN to improve analytical accuracy. Noise was defined according to the following criteria: (1) scenes not focused on the tennis court, (2) scenes where the ball was out of bounds, and (3) scenes where neither player was engaged in receiving or serving actions. As shown in Fig. 2, we modeled the tennis court with 48 relative positions for both players during a match. The encoding of these positions corresponds to the players' on-court locations. Fig. 3 depicts the ball's landing positions. Table 1 lists common receiving and serving techniques in tennis [37]-[38]. Any scene that failed to meet the criteria defined in Fig. 2 was considered noise.

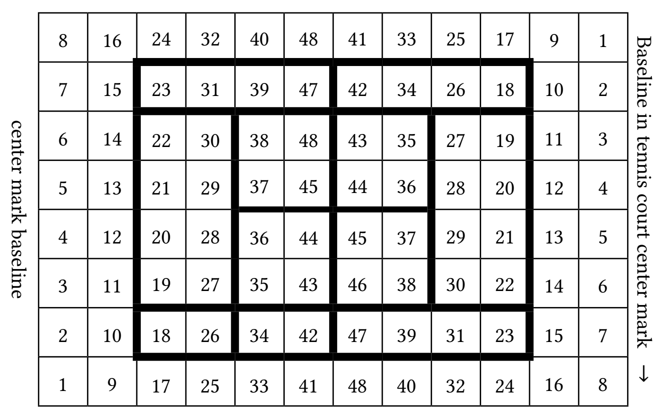


Fig. 2. Players' position on the tennis court.

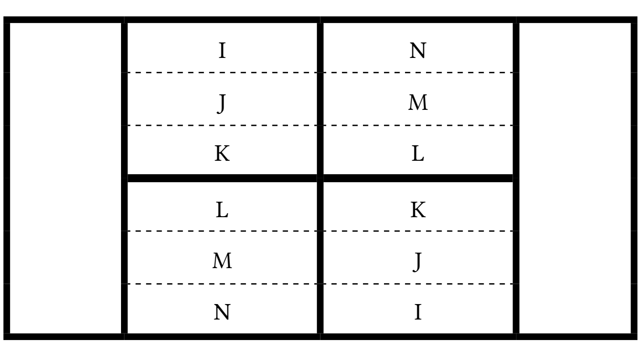


Fig. 3. Returner's impact location of player.

Step 2: Filtering Noise

The selected video footage was processed to eliminate noise as defined in Step 1.

Step 3: Marking Noise Timestamps

The timestamps of identified noise shots were marked for later verification and analysis.

V. Stage 2: Recurrent Neural Network Algorithm

We deployed the RNN algorithm according to the following steps:

Step 1: Defining Image Features for Analysis

Based on the definitions provided in Fig. 2, Fig. 3, and Table I, the positions of the players, the ball's landing locations, and the tennis techniques are defined and recognized in the images.

Step 2: Initialize

We define the dimensions of various parameters—including U, V, W, b, and c—to implement the basic RNN unit.

Input: At each time step t, the input $x(t)$ is fed into the network.

Hidden State: $h(t)$ represents the hidden state at time step t, serving as the "memory" of the network. $h(t)$ is computed based on the current input and hidden state from the previous time step, and $h(t)$ is defined as (1). This function is considered a non-linear transformation, such as tanh or *ReLU*.

$$h(\mathrm{t}) = f(Ux(\mathrm{t}) + W\mathrm{h}(\mathrm{t}-1)) \tag{1}$$

Weights: The RNN has input-hidden connections parameterized by the weight matrix U, recurrent hidden-hidden connections parameterized by the weight matrix W, and hidden-output connections parameterized by the weight matrix V. All the weights are shared across time.

Output: $o(t)$ represents the output of the network. This output is often subject to non-linear transformations, especially when the network contains more layers downstream.

Step 3: Forward pass

Based on our equations for each timestamp t, we compute the hidden state $h(t)$ and apply the softmax function to obtain the output $o(t)$, which represents the probability of the next character.

Calculating softmax and numerical stability:

The softmax function takes an N-dimensional real-valued vector and transforms it into a real-valued vector within the range of (0, 1), with the elements summing to 1. We performed this transformation using the formula as (2):

TABLE I. Tennis Techniques and Corresponding Algorithm Codes

Techniques	Code	Definition
Forehand	FH	A stroke in which the inner side of the palm of the dominant hand that is holding the racket faces forward. Essentially, the tennis forehand is made by swinging the racket across one's body in the direction one wants to land the ball.
Backhand	H	A shot in which one swings the racket around one's body with the back of the hand preceding the palm. In a backhand volley, the term backhand refers to a groundstroke.
Forehand Volley	FHV	A fairly simple movement, in which the player uses one arm to hit the ball by their dominant side without letting the ball touch the ground. The player usually hits volleys when standing close to the net and inside the service box. This move requires firm hands and fast reflexes.
Backhand Volley	BHV	The player's hitting arm is bent, and their elbow is centered between their shoulders. The backhand volley begins with a hip and shoulder turn. This volley can be shot while standing close to the back fence to test the size of one's backswing.
Forehand Half-volley	FH	A shot in which the player hits the ball just after it bounces off the ground and before it reaches the height of a normal volley: 1.Move quickly to get into position for the shot. The ideal position is slightly behind the baseline, but not too far back. 2.Keep your knees slightly bent and your weight on the balls of your feet to maintain balance, and be ready to move in any direction. 3.As the ball approaches, take a small step forward and to the side to get into the right position. 4.Keep your racket head up and swing forward with a smooth and controlled motion. Try to contact the ball just after it bounces off the ground. 5.Follow through with your swing and always keep your eye on the ball. 6.Recover quickly and be ready for the next shot.
Backhand Half-volley	BH	A shot where the player hits the ball just after it bounces off the ground and before it reaches the height of a normal volley, using their backhand stroke: 1.Move quickly to get into position for the shot. The ideal position is slightly behind the baseline, but not too far back. 2.Keep your knees slightly bent and your weight on the balls of your feet to maintain balance and be ready to move in any direction. 3.As the ball approaches, take a small step forward and to the side to get into the right position. 4.Keep your racket head up and swing forward with a smooth and controlled motion using your backhand stroke. Try to contact the ball just after it bounces off the ground. 5.Follow through with your swing, and always keep your eye on the ball. 6.Recover quickly and be ready for the next shot.
Forehand Spin	FHS	The ability of a player to apply spin to the ball while executing a forehand stroke. Forehand spin is achieved through a combination of racquet and stroke techniques. Topspin involves brushing the racquet from above the ball's center towards the bottom, generating a downward spin on the ball. This causes the ball to descend in an arched trajectory as its rotational speed increases.
Backhand Spin	BHS	The ability of a player to impart spin to the ball when hitting the backhand. Backhand spin is achieved through a combination of racket and stroke techniques. This can produce topspin or backspin depending on how the player uses the racket and body when hitting the ball. Spin is generated mainly from the angle, speed, and position of the hitting point of the racket.
First Serve	FS	The first serve a player makes at the beginning of a serving game. This serve is usually the strongest and fastest serve a player can deliver, and is intended to achieve a direct score or take advantage of the opponent's return.
Second Serve	SS	The player's second serve in the service game, occurring after the player misses the first serve or the receiver successfully returns the ball.
Smash	S	Usually occurs when the player is in front of the net. When the ball is in the air, the player quickly jumps and hits the ball down hard, so that the ball falls to the opponent's court with great speed and force.
Unsuccessful Shots	US	A shot that fails to achieve the intended goal or produce the desired effect. These shots may occur in a variety of situations. Examples of unsuccessful shots include the following: Errors: These include unforced and forced errors. An unforced error is the failure to complete a successful stroke without apparent stress, such as hitting the ball out of bounds or netting. A forced error may occur when the opponent applies pressure or creates difficulty, e.g., by making a hard return that leaves the other player unable to hit back. Hitting the ball too deep: When the player hits the ball beyond the bottom line, leading the ball to fall behind the backcourt, the shot is considered unsuccessful. This can result in missed scoring opportunities and give the opponent the opportunity to counterattack. Too-short shot: A shot is judged unsuccessful when it does not have enough power or height to keep the ball from clearing the net and landing in its own court. This usually gives the opponent the opportunity to attack, forcing the player to react. Net shot: A shot is deemed unsuccessful when the player does not hit the ball high enough for it to go over the net. This results in a lost scoring opportunity, allowing the opponent to score or restart the round. Unstable shot: A shot is considered unsuccessful when it lacks stability, i.e., when the player cannot accurately control the direction, height, or spin of the ball. This may cause the ball to go out of bounds or give the opponent a chance to score. Unsuccessful shots commonly occur in games, and present opportunities for players to improve and adjust their shot techniques and tactics. Through reflection and training, players can improve the stability and accuracy of their shots, thereby increasing their chances of success.
Others	O	Not defined above; all count as other tennis skills.

$$p_i = \frac{e^{a_i}}{\sum_{k=1}^{N} e_k^a} \quad (2)$$

Step 4: Compute Loss

In a text generation model, the next character can be any unique character from the given vocabulary. Accordingly, we implemented a cross-entropy loss. In multi-class classification, the logarithmic loss values are summed for each predicted class in the observation. The compute loss is also defined as (3).

$$\mathrm{CE} = -\sum_{c=1}^{M} y_{o,c} \log\left(p_{o,c}\right) \quad (3)$$

Step 5: Backpropagation

The gradients must propagate from the last cell to the first cell. The product of these gradients may become zero or grow exponentially. The latter case corresponds to the gradient explosion problem, where a significant increase in the gradient norm accumulates during the training process. The former case corresponds to the gradient vanishing problem, wherein long-term components reach a norm of zero, rendering the model unable to learn correlations between events that are far apart in time.

Step 6: Update Weights

The gradients for each model parameter are calculated and updated accordingly.

Step 7: Repeat Steps 2-6

To train the model and generate text from the data, it is necessary to train the model for a certain period of time and evaluate the loss following each iteration. If the loss exhibits an overall decrease, the model's learning is progressing as expected.

VI. Results and Discussions

Our analysis was conducted on 10 videos of matches played by Roger Federer and Rafael Nadal from 2007 to 2019. We recorded the techniques, player positions, shot landing points, and serve landing points for each ball until every ball was analyzed.

Our deep learning analysis yielded a total of 122 valid shots between the two players. Table II lists statistics pertaining to the techniques used by Federer and Nadal individually, and Fig. 4 and Fig. 5 depict the corresponding graphic representations. The five most used techniques by both players, in order, were: forehand groundstroke, backhand groundstroke, first serve, backhand slice, and second serve.

TABLE II. Analyzed Player Statistics

Tennis Techniques	Federer		Nadal	
Code	times	accuracy	times	accuracy
FH	2232	35.94%	2506	41.48%
H	1948	31.37%	1809	29.95%
FHV	770	12.40%	871	14.42%
BHV	477	7.68%	364	6.03%
FH	414	6.67%	341	5.64%
BH	118	1.90%	31	0.51%
FHS	87	1.40%	24	0.40%
BHS	88	1.42%	51	0.84%
FS	35	0.56%	19	0.31%
SS	23	0.37%	20	0.33%
S	11	0.18%	2	0.03%
US	7	0.11%	1	0.02%
O	0	0.00%	2	0.03%
Summary	6210	100%	6041	100.0%

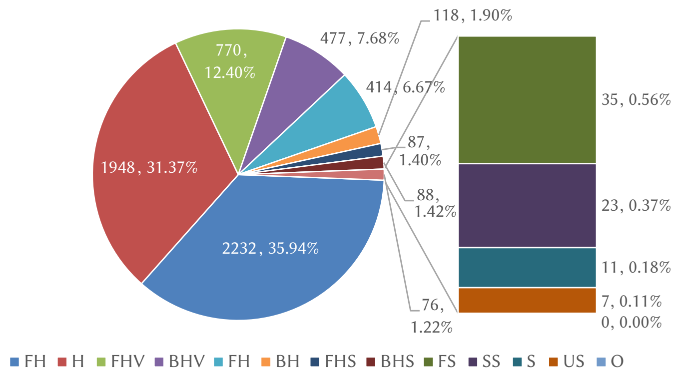


Fig. 4. Statistics of tennis techniques used by Federer.

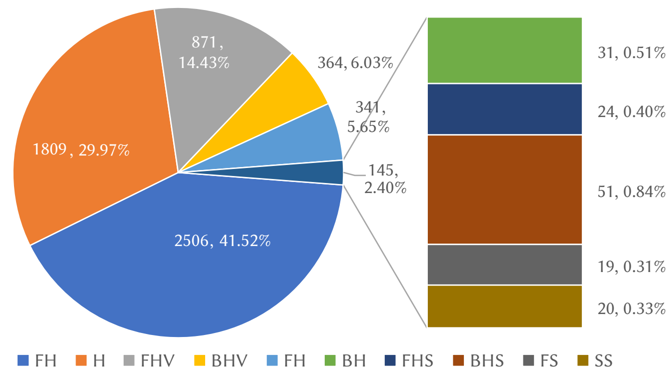


Fig. 5. Statistics of tennis techniques used by Nadal.

To illustrate the analytical results, we consider the first serves and winning shots used by Federer and Nadal. Federer hit a total of 770 first serves, of which 275 resulted in points and 495 led to second shots. Nadal hit a total of 871 first serves, with 243 yielding points and 628 leading to a second shot. Looking at the first serve points, Federer's scoring efficiency is 53.1%, whereas Nadal's is 46.9%. An analysis of player positions, techniques, and landing points corresponding to first serve scores is presented in Table III.

In Federer's case, 60.36% of the successful first serves were shot from position 12, with the remaining 39.64% shot from position 13.

In Nadal's case, 51.03% of the successful first serves were shot from position 12, with the remaining 48.97% shot from position 13.

Because Federer scored 275 points with his first serve, whereas Nadal scored 243 points, Federer has a lead of 32 points over Nadal. Both players were more successful when serving for position 12. For this position, Federer leads Nadal by 42 points.

When Federer served from position 12, most of the balls landed in position I, corresponding to 32.36% of his points. When Nadal served from position 12, most of the balls also landed in position I, corresponding to 23.46% of his points. We can observe that both players scored higher when they served from position 12, as well as when the ball landed in position I. However, these results indicate that Federer had an advantage over Nadal when serving.

In a comparative analysis, we employed existing models to examine the two players' first-serve performance. For Federer, Qiu's method [16] achieved a classification accuracy of 96.73%, the C5.0 algorithm by Chang and Qiu [17] achieved a classification accuracy of 98.36%, and our proposed method achieved a classification accuracy of 98.96%. For Nadal, Qiu's method achieved a classification accuracy of 98.36%, the C5.0 algorithm achieved a classification accuracy of 99.10%, and our proposed method achieved a classification accuracy of 99.37%. Thus, the proposed deep learning method outperformed the two existing methods for both players.

TABLE III. Analysis of First Serves

Plays \ Score Analysis	Batting Position	Ball Drop				
		Location	Times	Sum	Total	%
Federer	12	I	89	166	275	32.36%
		J	8			2.91%
		K	69			25.09%
	13	L	61	109		22.18%
		M	5			1.82%
		N	43			15.64%
Nadal	12	I	57	124	243	23.46%
		J	38			15.64%
		K	29			11.93%
	13	L	56	119		23.05%
		M	29			24.37%
		N	34			28.57%

Because tennis is an open-ended sport, players must not only exercise their own abilities, but also adjust their playstyle in response to their opponents. In this study, we defined and classified the variables of shot placement, technique usage, and landing point using a deep learning algorithm to analyze video footage. Our model was successfully used to interpret the playstyles, strengths, and weaknesses of two professional tennis players.

VII. Conclusions

In this study, we employed the techniques of annotation analysis and decision tree algorithms to categorize the scoring techniques of two prominent tennis champions. Accordingly, we constructed an adversarial model to identify the habitual scoring positions, stroke techniques, and shot placements of players during matches. The matchup model can be applied to analyze the strengths and weaknesses of any tennis player by simultaneously analyzing their opponent. Through this approach, we attained an understanding of the pivotal factors contributing to the victories of the two tennis champions.

The following observations were confirmed by our analysis:

(1) The results in Table 3 indicate that both Federer and Nadal aim for the inside corner K when serving from position 12, whereas they target the outside corner I when serving from position 13. This corresponds to the results shown in Figures 2 and 3, indicating that Federer's serves primarily attack Nadal's forehand, whereas Nadal focuses on targeting Federer's backhand.

(2) Although Nadal's serves predominantly target Federer's backhand, the probabilities of the six target areas are relatively even, and Nadal scores fewer points with his serves than Federer. In contrast, Federer's serves concentrate on the inside and outside corners irrespective of serving position, with lower probabilities in the middle. This tactical approach of aiming for the edges and creating wide angles not only increases the likelihood of scoring, but also forces the opponent to create openings, which explains why Federer scores more points.

(3) To prevent the opponent from scoring, data analyses can be conducted in advance to understand the opponent's preferred shot techniques and playstyle. In this study, we utilized the labeled analysis method to mark the landing points of shots, and integrated the serving positions, techniques, and landing points of both players using decision tree algorithms to develop a matchup model. This model can provide guidance to players during pre-match training and preparation, helping them overcome their limits, address weaknesses, enhance strengths, and maximize winning potential.

A limitation of this study is that we only considered 10 matches played by Federer and Nadal between 2007 and 2019. We did not account for variations in match venues, such as tennis court surfaces. In future studies, new deep learning methods can be developed to enhance classification performance.

Funding Author gratefully acknowledges the sponsorship of the National Science and Technology Council in Taiwan, ROC, under the project NSTC 112-2410-H-028-002-.

References

[1] J. Del Corral, "Competitive balance and match uncertainty in grand-slam tennis," *Journal of Sports Economics*, vol. 10, no. 6, pp. 563–581, 2009.

[2] D. Whitson and J. Horne, "Underestimated costs and overestimated benefits? Comparing the outcomes of sports mega-events in Canada and Japan," *The Sociological Review*, vol. 54, no. 2_suppl, pp. 73–89, 2006.

[3] D. Maynes-Aminzade, R. Pausch and S. Seitz, "Techniques for interactive audience participation," Proceedings of the Fourth IEEE International Conference on Multimodal Interfaces, Pittsburgh, PA, USA, 2002, pp. 15-20, doi: 10.1109/ICMI.2002.1166962.

[4] G. Nhamo, K. Dube, and D. Chikodzi, "Implications of COVID-19 on Gaming, Leisure and Entertainment Industry," in Counting the Cost of COVID-19 on the Global Tourism Industry, Springer, 2020, doi: 10.1007/978-3-030-56231-1_12.

[5] S. Keshkar, G. Dickson, A. Ahonen, K. Swart, F. Addesa, A. Epstein, A., et al., "The effects of coronavirus pandemic on the sports industry: an update," *Annals of Applied Sports Science*. vol. 9, no. 1, 2021.

[6] R. Rein and D. Memmert, "Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science," *SpringerPlus*, vol. 5, pp. 1410, 2016, doi: 10.1186/s40064-016-3108-2.

[7] D. He, L. Li, and L. An, "Study on sports volleyball tracking technology based on image processing and 3D space matching," *IEEE Access*, vol. 8, pp. 94258–94267, 2020.

[8] Z. Li, X. Ye, and H. Liang, "Sports video analysis system based on dynamic image analysis," *Neural Computing and Applications*, vol. 35, no. 6, pp. 4409–4420, 2022.

[9] B. Meng, X. Liu, and X. Wang, "Human action recognition based on quaternion spatial-temporal convolutional neural network and LSTM in RGB videos," *Multimedia Tools and Applications*, vol. 77, pp.26901–26918, 2018.

[10] G. Fele, and G. M. Campagnolo, "Expertise and the work of football match analysts in TV sport broadcasts," *Discourse Studies*, vol. 23, no. 5, pp. 616–635, 2021.

[11] G. Thonhauser and M. Wetzels, "Emotional sharing in football audiences," *Journal of the Philosophy of Sport*, vol. 46, no. 2, pp. 224–243, 2019.

[12] I. Errekagorri, J. Castellano, I. Echeazarra, and C. Lago-Peñas, "The effects of the video assistant referee system (VAR) on the playing time, technical-tactical and physical performance in elite soccer," *International Journal of Performance Analysis in Sport*, vol. 20, no. 5, pp. 808–817, 2020.

[13] C. Lago-Peñas, M. Gómez, and R. Pollard, "The effect of the video assistant referee on referee's decisions in the Spanish LaLiga," *International Journal of Sports Science & Coaching*, vol. 16, no. 3, pp. 824–829, 2020.

[14] I. Tamir and M. Bar-eli, "The Moral Gatekeeper: Soccer and Technology, the Case of Video Assistant Referee (VAR)," *Frontiers in Psychology,* vol. 11, pp. 613469, 2021, doi: 10.3389/fpsyg.2020.613469.

[15] K. Rangasamy, M. A. As'ari, N. A. Rahmad, N. F. Ghazali, and S. Ismail, "Deep learning in sport video analysis: a review," *Telkomnika*, vol. 18, no. 4, pp. 1926–1933, 2020.

[16] Y. R. Qiu, "Construct a model of tennis players' competition," Master Thesis, National Taiwan University of Sport, Department of Recreational & Graduate Institute of Recreational Sport Management, Taichung, Taiwan, 2020.

[17] C. W. Chang and Y. R. Qiu, "Constructing a Gaming Model for Professional Tennis Players Using the C5.0 Algorithm," *Applied Sciences,* vol. 12, no. 16, pp. 1–13, https://doi.org/10.3390/app12168222

[18] L. -Y. Duan, M. Xu, and Q. Tian, "Semantic shot classification in sports video," in *Storage and Retrieval for Media Databases 2003*, 2003, pp. 300–313.

[19] B. Dang, A. Tran, T. Dinh, and T. Dinh, "A real time player tracking system for broadcast tennis video," *Lecture Notes in Computer Science*, vol. 5991, pp. 105–113, 2010.

[20] D. Connaghan, P. Kelly, N. E. O'Connor, M. Gaffney, M. Walsh, and C. O'Mathuna, "Multi-sensor classification of tennis strokes," in *SENSORS*, 2011 IEEE, pp. 1437–1440, 2011.

[21] C. J. Raymond, T. J. Madar, and A. H. Montoye, "Accuracy of the Babolat Pop sensor for assessment of tennis strokes in structured and match play settings," *Journal of Sport and Human Performance*, vol. 7, no. 1, 2019.

[22] W. Pei, J. Wang, X. Xu, Z. Wu, and X. Du, "An embedded 6-axis sensor based recognition for tennis stroke," *2017 IEEE International Conference on Consumer Electronics* (ICCE), pp. 55–58, 2017.

[23] M. Sharma, R. Srivastava, A. Anand, D. Prakash, and L. Kaligounder, "Wearable motion sensor based phasic analysis of tennis serve for performance feedback," *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5945–5949, 2017.

[24] N. Myers, W. Kibler, A. Axtell, and T. Uhl, "The Sony Smart Tennis Sensor accurately measures external workload in junior tennis players," *International Journal of Sports Science & Coaching*, vol. 14, no. 1, pp. 24-31, 2019, doi: https://doi.org/10.1177/1747954118805278.

[25] C. J. Ebner and R. D. Findling, "Tennis Stroke Classification: Comparing Wrist and Racket as IMU Sensor Position," in *Proceedings of the 17th International Conference on Advances in Mobile Computing & Multimedia*, pp. 74-83, 2019.

[26] T. Perri, M. Reid, A. Murphy, K. Howle, and R. Duffield, "Validating an algorithm from a trunk-mounted wearable sensor for detecting stroke events in tennis," *Journal of Sports Sciences*, vol. 40, no. 10, pp. 1168–1174, 2022.

[27] R. J. Lake, Routledge Handbook of Tennis, Historical changes in playing styles and behavioural etiquette in tennis, England: Taylor & Francis, 2019.

[28] R. Voeikov, N. Falaleev, and R. Baikulov, "TTNet: real-time temporal and spatial video analysis of table tennis," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, pp. 884–885, 2020.

[29] S. Xu, L. Liang, and C. Ji, "Gesture recognition for human–machine interaction in table tennis video based on deep semantic understanding," *Signal Processing: Image Communication*, vol. 81, p. 115688, 2020.

[30] F. Qiao, "Application of deep learning in automatic detection of technical and tactical indicators of table tennis," *PLoS ONE,* vol. 16, no. 3, pp. 0245259, 2021.

[31] V. Reno, N. Mosca, R. Marani, M. Nitti, T. D'Orazio, and E. Stella, "Convolutional neural networks based ball detection in tennis games," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pp. 1758–1764, 2018.

[32] A. Ganser, B. Hollaus, and S. Stabinger, "Classification of tennis shots with a neural network approach," *Sensors*, vol. 21, no. 17, pp. 5703, 2021.

[33] A. Bastanfard, and D. Amirkhani, "Improving the Accuracy of the Annotation Algorithm in Pattern-Based Tennis Game Video," in *2021 29th Iranian Conference on Electrical Engineering (ICEE)*, pp. 493–497, 2021.

[34] H. Huang, "Object extraction of tennis video based on deep learning," *Wireless Communications and Mobile Computing*, vol. 2022, article ID 5402410, 2022.

[35] B. Ning, and L. Na, "Deep spatial/temporal-level feature engineering for tennis-based action recognition," *Future Generation Computer Systems*, vol. 125, pp. 188–193, 2021.

[36] H. Li, H. Guo, and H. Huang, "Analytical model of action fusion in sports tennis teaching by convolutional neural networks," *Computational Intelligence and Neuroscience*, vol. 2022, 2022.

[37] TennisCT, "THE TENNIS STROKES," Accessed: May 4, 2023. [Online]. Available: https://www.tennisct.com/strokes.

[38] My Tennis HQ, "Tennis Volleys: Definition, Steps & Tips," Accessed: May 4, 2023. [Online]. Available: https://mytennishq.com/.

Jing-Wei Liu

Dr. Jing-Wei Liu is an associate professor of Department of Sport Information & Communication, National Taiwan University of Sport. He specializes in Big Data, Deep learning, Fuzzy Time Series, Data Mining, Soft Computing, and Artificial Intelligence. His paper appeared in IEEE Access, Soft Computing, Journal of Ambient Intelligence and Humanized Computing, International Journal of Information Technology & Decision Making, Journal of Systems and Software, Journal of Computer Information Systems, Computers and Industrial Engineering, Computers & Education, Computers and Mathematics with Applications, Economic Modelling, Journal of Computer Information Systems, International Journal of Information and Management Sciences, Plant Systematics and Evolution, Expert Systems with Applications, Advanced Materials Research, Open Journal of Social Sciences, and others.

The Human Motion Behavior Recognition by Deep Learning Approach and the Internet of Things

Hui Li[1], Huayang Liu[2], Wei Zhao[3], Hao Liu[4]*

[1] Department of Physical Education, Luoyang Institute of Science and Technology, Luoyang, 471023 (China)
[2] College of Physical Education China West Normal University, Nanchong,637001 (China)
[3] College of Humanities and Social Sciences, Luoyang Institute of Science and Technology, Luoyang,471023 (China)
[4] College of Physical Education and Health, Guangxi Medical University.Nanning,530021 (China)

* Corresponding author: linyiliuhao@sina.com

Received 25 April 2023 | Accepted 17 May 2024 | Early Access 24 July 2024

unir
LA UNIVERSIDAD EN INTERNET

Abstract

This paper is dedicated to exploring the practical implementation of deep learning and Internet of Things (IoT) technology within systems designed for recognizing human motion behavior. It places a particular emphasis on evaluating performance in complex environments, aiming to mitigate challenges such as poor robustness and high computational workload encountered in conventional human motion behavior recognition approaches by employing Convolutional Neural Networks (CNN). The primary focus is on enhancing the performance of human motion behavior recognition systems for real-world scenarios, optimizing them for real-time accuracy, and enhancing their suitability for practical applications. Specifically, the paper investigates human motion behavior recognition using CNN, where the parameters of the CNN model are fine-tuned to improve recognition performance. The paper commences by delineating the process and methodology employed for human motion recognition, followed by an in-depth exploration of the CNN model's application in recognizing human motion behavior. To acquire data depicting human motion behavior in authentic settings, the Internet of Things (IoT) is utilized for extracting relevant information from the living environment. The dataset chosen for human motion behavior recognition is the Royal Institute of Technology (KTH) database. The analysis demonstrates that the network training loss function reaches a minimum value of 0.0001. Leveraging the trained CNN model, the recognition accuracy for human motion behavior achieves peak performance, registering an average accuracy of 94.41%. Notably, the recognition accuracy for static motion behavior generally exceeds that for dynamic motion behavior across different models. The CNN-based human motion behavior recognition method exhibits promising results in both static and dynamic behavior recognition scenarios. Furthermore, the paper advocates for the use of IoT in collecting human motion behavior data in real-world living environments, contributing to the advancement of human motion behavior recognition technology and its application in diverse domains such as intelligent surveillance and health management. The research findings carry significant implications for furthering the development of human motion behavior recognition technology and enhancing its applications in areas such as intelligent surveillance and health management.

Keywords

Behavior Recognition, Convolutional Neural Network, Human Body Movement, Internet of Things.

DOI: 10.9781/ijimai.2024.07.004

I. Introduction

In the domain of computer vision, human motion behavior recognition stands as a prominent and widely acknowledged subject of interest. This field holds considerable application value in various domains, including intelligent monitoring, robotics, human-computer interaction (HCI), virtual reality, smart home technologies, smart security systems, and athletic training assistance [1]. A practical application of human motion analysis is content-based video retrieval, allowing for efficient retrieval of specific athlete movements, such as those observed during horizontal bars competitions in sports events. This technology not only saves users significant time and effort in querying video data [2] but also facilitates the extraction of various technical parameters, such as joint position, angle, and angular speed, contributing valuable guidance and suggestions for athletes' training and overall improvement. Moreover, this application finds relevance in sports dance movement analysis and clinical orthopedic research. Exploring human motion-tracking research presents a range of

Please cite this article as:
H. Li, H. Liu, W. Zhao, H. Liu. The Human Motion Behavior Recognition by Deep Learning Approach and the Internet of Things, International Journal of Interactive Multimedia and Artificial Intelligence, vol. 8, no. 7, pp. 55-65, 2024, http://dx.doi.org/10.9781/ijimai.2024.07.004

theoretical and practical challenges within fields like computer vision, pattern recognition, and video image processing, especially when considering the non-rigid nature of the human body undergoing rotational motion of joints [3].

Previous research has demonstrated the significant success of convolutional neural network (CNN) models in image and video recognition tasks. However, for complex human motion behavior recognition, CNN faces challenges in terms of robustness. This paper proposes a method to enhance human motion behavior recognition by optimizing CNN model parameters. Additionally, it categorizes human motion behavior into static and dynamic actions, investigating the recognition results for different types of motion behaviors to improve accuracy. The application of the Internet of Things (IoT) introduces more diverse data scenarios for human motion behavior recognition, thereby enhancing recognition accuracy and robustness. Collecting diversified data via IoT better represents different aspects of human motion behavior, leading to a more comprehensive and accurate recognition [4]. Several challenges in human motion behavior recognition research include the immense variability in human movement patterns, resulting in identical movements manifesting in different behavioral performances. Additionally, the wide array of movements within human motion exhibits numerous distinct manifestations, posing challenges in accurately identifying human motion behaviors. Diverse viewpoints can yield various two-dimensional images of the same action, and occlusions between individuals and backgrounds present difficulties in early-action classification during feature extraction. The proposed system's insensitivity to video playback rates, exacerbated by dynamic and cluttered backgrounds, fluctuating ambient lighting conditions, and low image and video resolution, further complicates the recognition process [5]. Some researchers have addressed multi-vision and occlusion problems through the proposal of multi-camera fusion [6], a technique managed through three-dimensional (3D) reconstruction [7].

The incorporation of human behavior recognition technology in community management holds the potential to establish an efficient and secure intelligent service system. This technology enables real-time behavior recognition for individuals and groups within community monitoring. In situations where hazardous behaviors, such as illegal entry and high-altitude throwing, go unnoticed by video surveillance personnel, the automated system can promptly alert community managers, mitigating safety risks and streamlining the subsequent evidence-gathering process. The integration of CNN models, based on supervised learning, with human behavior recognition technology represents a promising research avenue deserving further exploration. Through additional research in this domain, this paper provides a robust theoretical framework and reliable technical support, laying the groundwork for future practical endeavors.

The paper aims to investigate the practical application of deep learning and IoT technology in human motion behavior recognition systems. It underscores the evaluation of performance in complex environments and addresses the challenges of poor robustness and high computational workload in traditional human motion behavior recognition methods using CNN. The focus is on evaluating the performance of human motion behavior recognition systems in real-world scenarios and optimizing them for real-time accuracy, enhancing their suitability for practical applications. This paper proposes an enhanced CNN-based human motion behavior recognition method by seamlessly integrating IoT technology and the CNN model. This method can accurately and efficiently recognize various types of human motion behaviors. The proposed approach in this paper holds significant practical relevance in real-world applications, with extensive utility in areas such as surveillance, health management, and intelligent transportation. It offers valuable technical support for real-time monitoring and recognition of human motion behavior. Moreover, the proposed method can adapt to different scales and complexities of application scenarios, laying a solid foundation for future research and applications in behavior recognition.

The paper is structured into five sections to provide a cohesive framework. Section 1 functions as an introduction, offering insights into the research background and the underlying motivation behind the paper. Section 2 presents a comprehensive review covering methodologies for target motion detection and the application of deep learning techniques in addressing challenges related to target recognition. Section 3 constitutes the paper's focal point, concentrating on the intricacies of human motion recognition. This section introduces an adaptive correlation learning module specifically based on traditional CNN, effectively calculating correlation weights between samples to enhance the recognition process. In Section 4, a series of meticulously designed experiments are conducted to empirically validate the performance of the proposed algorithm. Furthermore, the algorithm's practical significance and real-world applicability are extensively discussed. Finally, Section 5 serves as a succinct summary, encapsulating the essential findings and insights conveyed throughout the entirety of the paper.

II. Literature Review

In the field of behavior recognition technology, Guo et al. proposed a method for foreground target motion detection where the background of the target motion remains unchanged. This approach leverages the changing background as a foreground for discriminating targets. However, if the target remains stationary for a certain period, the static part is updated to the background, making it unidentifiable. Therefore, the construction of a robust background model capable of adaptive updates becomes crucial [8]. In a distinct context, Wei et al. introduced the fusion of Parametric Rectified Linear Units (ReLU) and robust initialization methods within a CNN to address applications in the ImageNet 2012 dataset. Their findings indicated that the recognition rate of the behavioral dataset surpassed that discernible by the human eye [9]. Acknowledging the complexity of the provided information, Bolanos et al. categorized behavior analysis methods into three classes: static gestures, motor behaviors, and recognition of complex processes. The static gesture recognition method primarily identifies the target's gesture within a static single-frame image, while motion behavior and complex process recognition focus on identifying video motion events [10]. Chebbout et al. performed human behavior recognition based on the spatial-temporal volume model of behavior recognition, which involves the projection of the human body onto the time axis and template matching [11]. Peng et al. engaged in feature extraction of 3D-Scale-invariant feature transform (3D-SIFT) points of interest and established a feature-based statistical histogram construction of video interest classes on the codebook. This eigenvector was then utilized for identification and training on Support Vector Machines (SVM). Lastly, leveraging the space-time trajectory method, key points in human motion were connected along the time axis to form a trajectory curve [12]. Chen et al. undertook the identification of multi-feature channels in a 3D-CNN, encompassing a grayscale image, vertical and horizontal gradients, and optical flow. Each input video sample comprised seven consecutive frames, ensuring effective utilization of time domain information. The experiments showcased the network's commendable recognition rate across real-world and Royal Institute of Technology public databases [13].

The video employs an innovative network architecture designed for human motion behavior recognition. Lin et al. formulated a 3D-CNN and conducted training on the University of Central Florida 101 (UCF101) dataset. During network training, eight convolution and four pooling operations were executed. The convolution core had dimensions of 3*3*3 with a stride of 1*1*1 [14]. In a separate study, Zhi et al. employed a 3D-CNN to extract spatial and motion features, integrating dense trajectory features into Long Short-Term Memory (LSTM) networks and embedding time series information. Subsequently, they utilized weighted averaging of the output from multiple LSTM units to obtain recognition results [15]. Jin et al. introduced a bidirectional CNN that incorporates both spatial and temporal information. This network employed two distinct paths to capture appearance information from a static frame and motion information between two frames, effectively enabling motion recognition [16]. Lu et al. developed a Trajectory-pooled CNN model, combining manual feature extraction with feature extraction derived from CNN models for motion recognition [17]. Furthermore, Guenzi et al. merged Deep Learning (DL) with Slow Feature Analysis (SFA), resulting in the construction of the Slow Feature Analysis-Deep Learning (SFA-DL) network specifically designed for behavior recognition [18]. Table I below illustrates the strategies, algorithms, and principal contributions adopted by different researchers in addressing behavior recognition problems.

Previous investigations have highlighted that targets exhibit not only spatial attributes but also temporal characteristics throughout their motion processes. The analysis of intricate behavioral processes places emphasis on human interactions and group behavior. In contrast to traditional human behavior recognition methods, CNN presents a considerable advantage by obviating the necessity for manual feature extraction. Instead, the network assimilates the characteristics that delineate the target's behavior and acts upon the target without prior experiential input. In the actual collection of human motion behavior data, challenges such as noise and incompleteness may arise, including image blurriness, occlusion, and missing data. These factors can influence the performance of the recognition model. In this study on human motion action recognition using CNN, the CNN model undergoes optimization to accommodate variations in complex environments and lighting conditions. This optimization reduces the model's parameter size and complexity, thereby mitigating computational workload, improving real-time performance and efficiency, and augmenting the CNN model's effectiveness in recognizing human motion behavior. For action behavior recognition in video sequences, this paper integrates multiple modalities of information, such as depth images and motion sensors, to capture temporal variations in action behavior, thereby enhancing recognition accuracy and robustness. Additionally, to address potential noise and incompleteness in the collected data from previous research, this paper utilizes the KTH database as the dataset for human motion behavior recognition experiments. The dataset undergoes preprocessing and enhancement, eliminating noise and supplementing missing data to fortify the model's robustness and accuracy.

III. Research Methodology

A. Research Approach

The principal aim of this paper is to proficiently utilize both labeled and unlabeled data, extracting valuable information to achieve optimal performance in semi-supervised behavior recognition. In order to fulfill this objective, this paper proposes a semi-supervised algorithm grounded in adaptive correlation learning. This algorithm capitalizes on the feature characteristics of samples to explore correlations between them and incorporates the acquired correlation information in the process of feature aggregation. Through the aggregation of features from neighboring samples for each sample, the algorithm generates more expressive and discriminative feature representations. The training process of the semi-supervised algorithm based on adaptive correlation learning is delineated in Fig. 1.

In Fig. 1, the initial step involves the preparation of both the training set and the unlabeled dataset. The training set comprises labeled samples designed for supervised learning, while the unlabeled

TABLE I. Strategies, Algorithms, and Main Contributions of Different Researchers in Behavior Recognition

Researcher	Strategies and Algorithms	Main Contributions
Guo et al. [8]	They proposed foreground target action detection method, the optimized background model	The research developed an adaptive and robust background model
Wei et al. [9]	They proposed CNN with the fusion of Parametric ReLU and robust initialization methods	The research achieved behavior recognition rates superior to human eyes on ImageNet 2012 dataset
Bolanos et al. [10]	They categorized behavior analysis into the static posture, motion behavior, and complex processes	The research classified recognition of video motion events and static postures
Chebbout et al. [11]	They used spatiotemporal volume model and template matching for human behavior recognition	The research introduced a novel approach to human behavior recognition
Peng et al. [12]	They employed 3D-SIFT and SVM for feature extraction and training of human actions	The research achieved favorable recognition results based on the spatial-temporal trajectory method
Chen et al. [13]	They utilized 3D-CNN multi-feature channels for human action recognition	The research demonstrated good recognition results in real-world and KTH public databases
Lin et al. [14]	They developed 3D-CNN for training the UCF101 dataset	The research employed specific network structure for human action recognition
Zhi et al. [15]	They used 3D-CNN to extract spatial and motion features, fused with LSTM for recognition	The research utilized LSTM's weighted average output for recognition results
Jin et al. [16]	They build bidirectional CNN for capturing spatiotemporal information	The research combined two pathways for motion and appearance recognition
Lu et al. [17]	They developed trajectory aggregation CNN model	The research integrated manual feature extraction and CNN model for motion recognition
Guenzi et al. [18]	They combined Deep Learning with Slow Feature Analysis	The research constructed SFA-DL network for behavior recognition

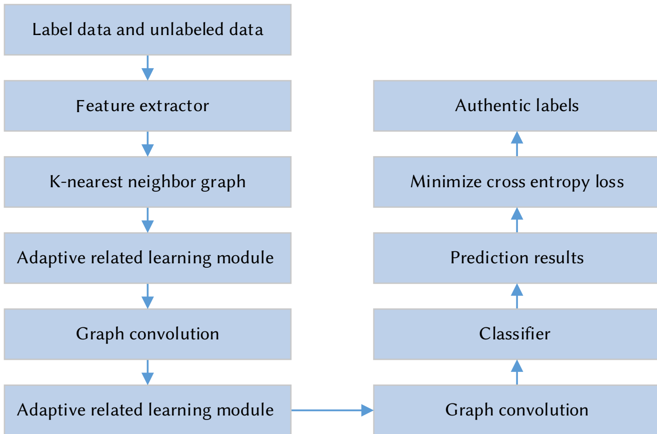


Fig. 1. Training process of semi-supervised algorithm based on adaptive correlation learning.

dataset includes samples devoid of labels, utilized in semi-supervised learning. Feature extraction is subsequently executed for both the training set and the unlabeled dataset, transforming raw data into more representative feature representations for subsequent learning. A pivotal stage in the semi-supervised algorithm based on adaptive correlation learning is correlation learning. The primary objective is to scrutinize the correlations between samples and incorporate this correlation information into the feature aggregation process. Unlabeled data is employed to learn these correlations. Following correlation learning, feature aggregation ensues to generate feature representations that are more expressive and discriminative. This is accomplished by aggregating the features of each sample with those of its neighboring samples, employing methods such as weighted averages or maximum pooling. Upon obtaining feature representations, a semi-supervised learning approach is employed to train the classifier. Semi-supervised learning combines both labeled and unlabeled samples for training, enhancing the classifier's generalization ability and accuracy by leveraging information from unlabeled samples. Ultimately, post-training, the model's performance on new samples is evaluated using either a validation set or cross-validation. The assessment results contribute to evaluating the model's effectiveness and generalization ability.

B. Recognition of Human Exercise Behavior

Human motion behavior recognition constitutes a pivotal research avenue within the realms of computer vision and pattern recognition, aiming to autonomously discern and comprehend diverse human motion actions through the implementation of computer algorithms and deep learning models. In the sphere of intelligent surveillance, the application of human motion behavior recognition technology in video surveillance systems facilitates the analysis and identification of pedestrian, vehicle, and other object behaviors. This integration enables functionalities like intelligent alerts and anomaly detection in behavior, significantly enhancing the efficiency and accuracy of surveillance systems. Such precise recognition contributes markedly to security personnel's ability to detect potential security risks. Within the domain of HCI, human motion behavior recognition technology finds utility in natural interaction, encompassing aspects such as posture recognition and gesture control. Identification of users' actions and postures enables computers to comprehend their intentions, thereby enhancing the convenience and intelligence of HCI. In the field of health management, the application of human motion behavior recognition technology extends to motion monitoring and rehabilitation assistance. Monitoring and analyzing human motion behavior allow for the assessment of individual movement status and the monitoring of movement performance. This, in turn, provides scientific evidence and personalized guidance for rehabilitation training, thereby augmenting rehabilitation outcomes. Despite the aforementioned applications, traditional human motion behavior recognition methods encounter challenges related to pose variations, complex backgrounds, and lighting changes, resulting in diminished recognition accuracy and weak robustness. The advent of deep learning, particularly the implementation of CNN, has substantially progressed human motion behavior recognition. CNN models demonstrate an inherent capacity to automatically learn features from data, exhibiting robust representational capabilities and adaptability. This advancement significantly elevates the accuracy and robustness of human motion behavior recognition.

The exhaustive analysis of human motion behavior encompasses several integral processes, including database construction, human motion detection, feature extraction, behavior comprehension, and recognition. The core emphasis in human motor behavior analysis lies in motion detection and feature extraction [19], as illustrated in the schematic representation of the recognition process for human motor behavior in Fig. 2.

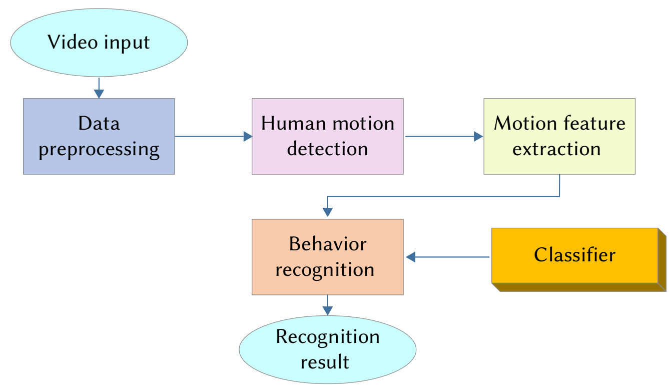


Fig. 2. Recognition process of human motion behavior.

As depicted in Fig. 2, the process of target classification detection entails extracting the region of interest from the foreground motion area identified by a moving object [20]. In intricate scenes, the foreground areas may encompass diverse targets, including pedestrians, vehicles, birds, clouds, swaying branches, among others. However, within the context of the human motion analysis system, the detection target is exclusively restricted to human movement. Hence, it becomes imperative to meticulously scrutinize, analyze, identify, and isolate human targets. The method utilized for detecting target classification is delineated in Fig. 3.

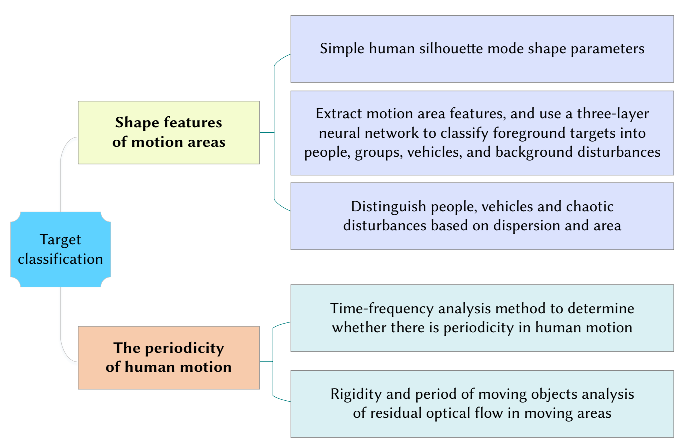


Fig. 3. Detection target classification method.

Fig. 3 entails the classification of detected targets based on the shape characteristics of the motion area and the periodicity of human movement. The assessment of residual light flow within the moving area enables the analysis of the rigidity and periodicity of the moving target. This approach proves effective due to the relatively higher residual light flow in non-rigid human movements compared to the movements of rigid vehicles, thereby facilitating the discrimination of human bodies. Currently, numerous video-based classification methods for moving objects are available, including those based on static features, dynamic features, or a combination of both. However, a single moving target feature often falls short of recognizing more than three targets or achieving satisfactory recognition accuracy. As a result, target classification studies typically select a minimum of two features. General features are indicative of characteristics applicable to all objects, while attribute characteristics represent the inherent qualities of a target, specifically reflecting its unique attributes.

Given the presence of multiple angles in the experimentally extracted foreground targets, including shadows and incomplete target area extraction, shape-based feature classification proves more suitable under such circumstances. The shape-based features of the target encompass target contour, area, aspect ratio, dispersion, centroid, and bounding rectangle [21]. The attribute characteristics of the moving target are delineated in Table II.

TABLE II. Attribute Characteristics of Moving Targets

Sports Goals	Attribute characteristics
People	Circular Periodicity of human motion (regular changes in human gait)
Automobile	Movement speed Variation in dispersion (variation of each target)
Bicycles	The attribute is somewhere between person and automobile.

In the present study, a wide array of target features is extensively employed, encompassing aspects such as aspect ratio, area information, dispersion (regional compactness), inertial principal axis direction, invariant moment, and other regional characteristics. For experimental purposes, several attributes are defined, including the ratio of target height to the width of the target area at approximately one-third of the height, the ratio of target height to width at about two-thirds of the target area height, and the duty cycle, defined as the ratio of the background area within the target boundary rectangle to the area of the boundary rectangle. Notably, the aspect ratio feature signifies the aspect ratio of the entire target. The analysis of moving object characteristics is demonstrated using the moving target within the scene, as depicted in Fig. 4.

Fig. 4 conducts classification on the extracted moving objects, distinguishing between "bicycle" and "automobile," "automobile" and "pedestrian," as well as "crowd" by detecting the moving objects present in the scene. Remarkably, the aspect ratio of the target "person" and "automobile" demonstrates significant differences.

C. CNN Modeling

CNN represents a variant of the Multilayer Perceptron (MLP) originating from early research conducted by biologists Hubel and Wiesel on the cat visual cortex [22]. The architecture of the CNN is delineated in Fig. 5.

In Fig. 5, the architecture of the CNN involves convolution, subsampling, and fully connected layers. Each level in the CNN comprises multiple feature maps, with each feature map extracting unique features from the input through a convolution filter, housing multiple neurons. Local features are extracted as the input image and

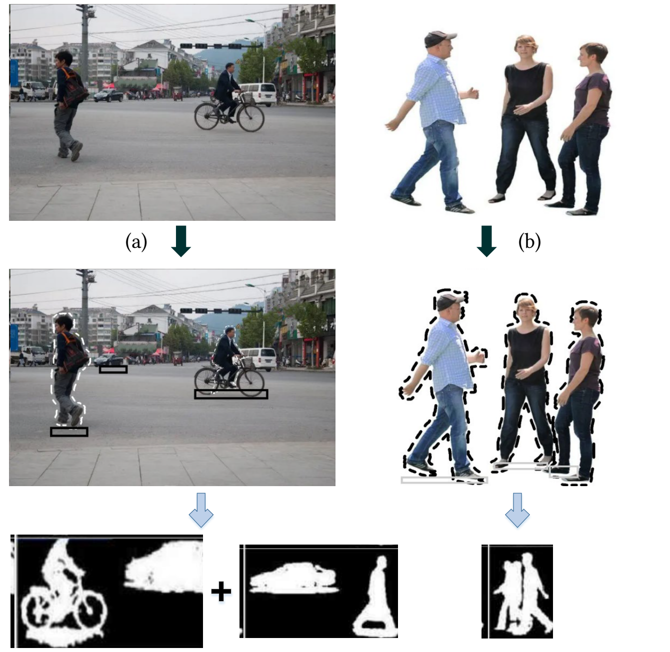


Fig. 4. Moving target in the scene (Data source: https://sucai.redocn.com/yixiang_6739023.html).

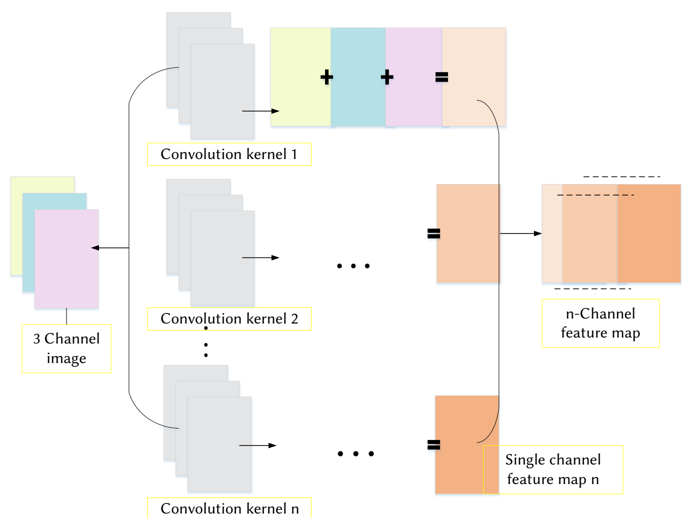


Fig. 5. Structure of CNN.

filters undergo convolution, determining their relationships with other features. The neurons in each layer receive the same input as the previous layer, establishing connected local receptive fields. The subsequent layer following each feature extraction layer is the computation layer responsible for local averaging and secondary extraction, also referred to as the feature mapping layer. This layer comprises multiple feature mapping planes with equal neuron weights. The mapping from the input layer to the hidden layer is commonly termed feature mapping. Consequently, the feature extraction layer is obtained through the convolution layer, while the feature mapping layer is achieved after downsampling [23]. The process of linking the convolution layer to the subsampling layer is illustrated in Fig. 6.

In Fig. 6, the input layer processes the normalized image, and each neuron within each layer receives input from a small local neighborhood of the previous layer. These neurons extract fundamental visual features, such as edges and corners, which are utilized by higher-level neurons. The CNN derives feature maps through the convolution operation, where cells at different locations acquire distinct features

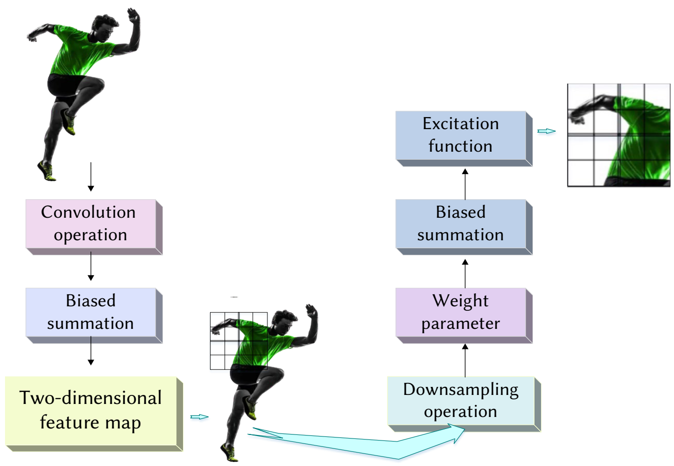


Fig. 6. Operation of connecting the convolution layer to the lower sampling layer.

from various feature maps. A convolution layer typically comprises multiple feature maps with different weight vectors, allowing for the retention of richer image features. Subsequently, the convolution layer is connected to the subsampling layer, serving dual purposes. Firstly, it reduces the image resolution and the number of parameters. Secondly, it fosters robustness to translation and deformation. The convolution and subsampling layers are interspersed throughout the network, progressively increasing the number of feature maps while decreasing the resolution [24], [25]. The calculation of the convolution layer is presented in Equation (1).

$$y_{mn} = f(\sum_{j=0}^{Q-1}\sum_{j=0}^{P-1} x_{m+i,n+j}w_{ij} + b), 0 \le m < M, 0 \le n < N \tag{1}$$

In Equation (1), $x_{(m+i,n+j)}$ represents the pixel value of the input data; $(m+i, n+j)$ indicates the two-dimensional coordinates of the point; y_{mn} represents the output data after convolution; b is offset; $P \times Q$ is the size of the convolution kernel; w_{ij} represents the value of the convolution kernel in (i, j). M and N are the input image in $P \times Q$; f is the excitation function. The excitation function is shown in Equations (2) and (3):

$$f(x) = \frac{1}{1+e^{-x}} \tag{2}$$

$$ReLU(x) = \max(0,x) = \begin{cases} 0, x < 0 \\ x, x \ge 0 \end{cases} \tag{3}$$

In Equations (2) and (3), within the Sigmoid function, the output of $f(x)$ ranges between [0, 1] when a real number is input. If the output approaches zero, the neuron exhibits no response; conversely, if the output tends to 1, the neuron is activated. However, when the input is excessively large or small, the output approaches 1 or 0, respectively. In such cases, the neuron becomes saturated, impeding weight updates, and resulting in vanishing gradients. Contrastingly, in the ReLU activation function, when $x \ge 0$, the output of $f(x)$ is x, leading to rapid network convergence, and the neurons remain unsaturated. Consequently, computational efficiency is enhanced. Following convolution, the calculation of image feature size is expressed in Equation (4):

$$N = \frac{W-F+2P}{S} + 1 \tag{4}$$

In Equation (4), $W \times W$ represents the input image size; F is the size of the convolution kernel; P indicates filling. The step size is S; $N \times N$ represents the output image size. The sampling calculation expression of the down-sampling layer in $S_1 \times S_2$ is shown in Equation (5):

$$y_{mn} = f(w\frac{1}{S_1S_2}\sum_{j=0}^{S_2-1}\sum_{j=0}^{S_1-1} x_{m\times S_1+i,n\times S_2+j} + b), 0 \le m < M, 0 \le n < N \tag{5}$$

In Equation (5), $x_{(m\times S_1+i,n\times S_2+j)}$ represents input data; $(m \times S_1 + i, n \times S_2 + j)$ represents the two-dimensional coordinates of the point; y_{mn} represents output data. b is offset; (S_1, S_2) is the pixel coordinate of the area; w is the weight value. Under the action of the common four excitation functions, a two-dimensional feature map is obtained with a resolution of ¼ of the original image, that is, S_{x+1} sample the layer for feature extraction again. The size of the feature map after downsampling is shown in Equation (6):

$$N = \frac{W-F}{S} + 1 \tag{6}$$

In Equation (6), $W \times W$ represents the input image size; F is the size of the downsampling template; the step size is S. $N \times N$ is the output image size. The network training process is divided into the forward and backward propagation stages. The calculation of forwarding propagation is shown in Equations (7), (8) and (9):

$$z^{(l)} = W^{(L)} \cdot a^{(l-1)} + b^{(l)} \tag{7}$$

$$a^{(l)} = f_l(z^{(l)}) \tag{8}$$

$$x = a^{(0)} \to z^{(1)} \to a^{(1)} \to z^{(2)} \to a^{(2)} \to \cdots \to z^{(l)} \to a^{(l)} \tag{9}$$

In Equations (7)-(9), l is the number of layers of CNN; $m^{(l)}$ is the number of neurons in l-layer; $W^{(L)}$ is the weight matrix; $b^{(l)}$ is offset; $a^{(l)}$ represents the output of l-layer neurons; $z^{(l)}$ is the input of l-layer neurons; $f_l(\cdot)$ is the activation function. Equation (9) represents the network prediction output $a^{(l)}$ of forward operation. The difference value of the backward propagation output layer is shown in Equation (10):

$$\delta^l = \frac{\partial J(W,b,x,y)}{\partial a^l} \odot \sigma'(z^l) \tag{10}$$

In Equation (10), δ^l represents the difference value of the output layer; $J(W, b, x, y)$ is the mean square error; z^l represents the input of layer l neurons; a^l is the output of l-layer neurons; $\sigma'(\cdot)$ indicates derivative operation. The calculation of δ^l is shown in Equation (11):

$$\delta^l = (W^{(l+1)})^T \cdot \delta^{(l-1)} \odot \sigma'(z^l) \tag{11}$$

All parameters (W, b) are updated, as shown in Equations (12) and (13).

$$\frac{\partial J(W,b,x,y)}{\partial W^l} = \delta^l (a^{l-1})^T \tag{12}$$

$$\frac{\partial J(W,b,x,y)}{\partial b^l} = \delta^l \tag{13}$$

D. IoT Technology

The IoT employs diverse connectivity technologies to meet connection requirements in various scenarios, including passive identification, short-distance wired, short-distance wireless, and long-distance wireless connections. The initial impetus for the IoT was driven by the emergence of Radio Frequency Identification (RFID) technology, although its passive reading nature limited its applicability in certain contexts. In the data collection and processing phase, the IoT integrates various sensor types into the network to capture different facets of human motion behavior. Visual sensors, such as cameras, are employed to obtain video data for recognizing human postures, actions, and motion trajectories. Motion sensors, like accelerometers, detect human movement status and acceleration changes. Concurrently, environmental sensors capture surrounding environmental information, such as light, temperature, and humidity, which may contribute to behavior recognition. Establishing real-time data transmission and communication mechanisms between sensors and the behavior recognition system is crucial. Through IoT technology, sensors can transmit collected data in real-time to the behavior recognition system, ensuring data timeliness and accuracy.

This facilitates real-time monitoring and recognition of motion behavior, thereby enhancing the system's real-time performance and efficiency [26],[27]. Additionally, the integration of edge computing into the IoT network allows local data processing and analysis, reducing the burden on centralized servers. In the behavior recognition process, tasks with high real-time requirements can be processed on edge devices where the sensors are located, mitigating the need to transmit all data to central servers for processing. This reduction in network latency improves response speed and lowers data transmission costs [28]. In order to enhance the energy efficiency of IoT devices, low-power sensors and energy-saving technologies are employed to extend sensor lifespan [29]. Energy harvesting techniques, such as solar charging or vibration energy harvesting, contribute to providing sustainable energy for sensors. By fully leveraging the potential of IoT, integrating various sensor types into the network, enabling real-time data transmission and communication, and incorporating edge computing technologies, an efficient, reliable, and real-time human motion behavior recognition system can be established [30]. Such a system is capable of addressing complex motion behavior recognition scenarios, enhancing accuracy and response speed, and providing comprehensive and reliable support for human motion behavior research and applications. The second wave of IoT was catalyzed by the maturity of short-range wireless networking technologies such as ZigBee and Wireless Fidelity (WIFI). Moreover, the ongoing evolution of cellular communication technology is anticipated to further facilitate the widespread adoption and advancement of IoT [31], [32]. In the context of ZigBee wireless networking technology, human motion behavior is collected in living environments, as illustrated in Fig. 7.

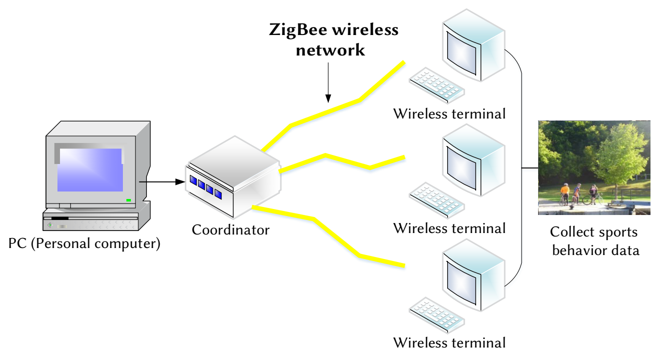


Fig. 7. Acquisition scenario under ZigBee wireless networking technology.

E. CNN Behavior Recognition Based on Adaptive Correlation Learning

In order to better represent the correlation between samples, this paper proposes an adaptive correlation learning module based on traditional CNN. This module can be used to calculate the correlation weights between samples. A shared learnable module is used to parameterize $corr(x_i, x_r)$ to obtain the specific value of w_{ir}, as shown in Equation (14):

$$w_{ir} = Z\left(Re\,L\,U\left(\frac{x_i \odot x_r}{\|x_i\|_2 \cdot \|x_r\|_2}\right)\right) \tag{14}$$

In Equation (14), Z represents a learnable weight vector and $\odot$ represents the Hadamard product. ReLU(*) represents an activation function, which can increase the sparsity of features and enhance the nonlinearity of the adaptive correlation learning module.

In the original feature space, the dimensions of features are usually relatively large. The input features are mapped to a low dimensional feature space, as shown in Equation (15):

$$\hat{x}_i = W x_i \tag{15}$$

In Equation (15), W represents a learnable linear transformation matrix. An offset term can also be added to the calculation process of feature mapping.

The operation of feature mapping can not only reduce the dimensionality of input features but also enhance the expression ability of features to a certain extent, as shown in Equation (16):

$$w_{ir} = Z\left(Re\,L\,U(*)\left(\frac{\hat{x}_i \odot \hat{x}_r}{\|\hat{x}_i\|_2 \cdot \|\hat{x}_r\|_2}\right)\right) \tag{16}$$

Initially, video samples are denoted as graph nodes, and a graph structure is established using the K-nearest neighbor method, integrating both labeled and unlabeled data. Within this graph structure, adaptive correlation learning is implemented, and the original feature X of each video sample serves as the input feature for the initial layer of graph convolution. The adaptive correlation learning module calculates the adjacency matrix for each layer of graph convolution, thereby capturing correlation information. In the course of feature aggregation within the graph convolutional networks, this correlation information is harnessed to generate more expressive features for the video samples. This is achieved by aggregating the features of neighboring samples within local neighborhoods.

IV. Experimental Design and Performance Evaluation

A. Datasets Collection

The KTH database has been selected as the dataset for human motion behavior recognition [33]. This dataset comprises six distinct actions (walk, jump, run, fist, wave, and clap) performed by 25 individuals across four scenes, totaling 599 videos. It is important to note that the background remains relatively static during the recording of these human motion behaviors. Although each video may have varying durations and camera shooting angles, the background remains relatively static, facilitating a more focused recognition of human motion behavior.

In order to ensure uniformity in size and resolution, video processing tools are employed for segmenting each video into individual frames, followed by cropping and scaling. Grayscale images, which contain only brightness information, are preferred over color images for human motion behavior recognition tasks, as they enhance the visibility of the human body's form and motion features. Employing data augmentation techniques, such as random rotations, flips, translations, and other operations, enhances the diversity and generalization ability of data samples, generating additional training samples. Prior to conducting experiments, the entire dataset is partitioned into training and testing sets with a ratio of 4:1, ensuring consistency in sample distribution and features between the two sets.

Recognition outcomes based on geometric shapes or motion information from various human motion behavior databases are presented in Table III. Notably, the KTH database demonstrates the highest recognition performance among the listed databases, achieving an impressive recognition rate of 95.77% based on geometric shape or motion information.

TABLE III. Recognition Results Based on Geometric Shape or Motion Information

Database	Accuracy of recognition
KTH [34]	95.77%
UCF [35]	86.5%
Hollywood 2 [36]	53.3%

B. The Setting of Experimental Environment and Parameters

The processor employed in this paper is the Intel (R) Core (TM) i5-7500 Central Processing Unit (CPU) @ 3.40GHz, while the operating system is Windows 10. The GTX1050, in conjunction with the Caffe framework, is utilized for GPU processing. The experimental dataset is divided into a training set and a test set with a ratio of 4:1, and each iteration encompasses 5000 generations. In order to optimize the network's recognition performance, certain network parameters undergo adjustment, including the size of the convolution kernel, the number of convolution layers, and the batch size. The optimization method involves maintaining the residual variable fixed and adjusting each individual variable until the optimal recognition rate is attained. Fig. 8 illustrates the specific parameters along with their corresponding value ranges.

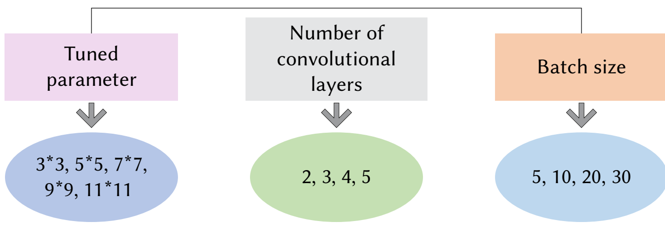


Fig. 8. Called parameters and their value ranges.

In Fig. 8, the CNN is configured to process single-channel grayscale images with a frame size of 80*80. The experiment centers on the KTH dataset and involves fine-tuning the CNN network architecture, specifically focusing on the convolution layers, kernel sizes, and other related parameters. Initially, the network configuration comprises three convolution and downsampling layers, two fully connected layers, and one Softmax layer responsible for generating identification results. The number of convolution kernels for the three convolution layers is set to 64, 128, and 128, respectively, while the two fully connected layers utilize 256 and 128 kernels. The initial learning rate is established at 0.005, and the training process is concluded after 20 Epochs.

C. Performance Evaluation

1. CNN Parameter Adjustment Results

When the batch size is defined as 10, the CNN network integrates three convolution layers and lower sampling layers. Fig. 9 illustrates the relationship between the CNN network parameters and the corresponding recognition accuracy.

In Fig. 9, the graphical representations illustrate the interdependencies between the convolution kernel size, the number of convolution layers, batch size, and their corresponding impact on recognition accuracy. Subfigure (A) elucidates the influence of the convolution kernel size on recognition accuracy, while Subfigure (B) delineates the effect of the number of convolution layers on recognition accuracy. Subfigure (C) provides insights into the relationship between batch size and recognition accuracy.

Upon meticulous examination of the findings, it is discerned that a convolution kernel size of 5*5 attains the highest recognition accuracy. Furthermore, when the number of convolution layers reaches 3, the network achieves its zenith recognition rate of 88.3%. Additionally, the network registers its peak recognition rate of 88.3% when the batch size is stipulated as 10. Consequently, the optimal configuration is ascertained to be a convolution kernel size of 5*5, three convolution layers, and a batch size of 10.

2. Analysis of Training Results Under the CNN Model

The training outcomes of the CNN model are visually represented in Fig. 10. Commencing at the initial iteration 0, the model's accuracy is documented at 0.0787. With successive iterations, there is a discernible enhancement in accuracy, coupled with a concurrent reduction in the value of the loss function. The loss function diminishes from its initial value of 1.7954 at iteration 0 to a minimal value of 0.0001 at iteration 5000. This decrease in the loss function signifies the progressive optimization of the model throughout the training process, resulting in a reduction of the disparity between predicted outcomes and actual labels. By the time the iteration count reaches 5000, the accuracy attains 92.59%. This data elucidates that the CNN model iteratively refines and assimilates knowledge during training, leading to a substantial improvement in classification accuracy on the test set.

Fig. 10 illustrates the categorization of human motion behavior into static and dynamic classifications. In order to evaluate the model's accuracy across distinct behavioral states, a comparative analysis is conducted, employing the proposed algorithm, CNN, SVM, and BPNN algorithm. The recognition accuracy of human motion behavior under these diverse models is depicted in Fig. 11.

In Fig. 11, the recognition accuracy of various models for static motion behavior generally exceeds that for dynamic motion behavior. Specifically, as depicted in Fig. 11(A), the CNN model attains the highest recognition accuracy for dynamic motion behavior, with an average accuracy of 93.61%. The proposed algorithm closely follows with an average recognition accuracy of 91.50%, while the SVM model achieves an average recognition accuracy of 83.83% for dynamic motion behavior recognition. The BPNN model records an average recognition accuracy of 90.44% for dynamic motion behavior recognition. In Fig. 11(B), concerning static motion behavior, the CNN model demonstrates the highest recognition accuracy, achieving an

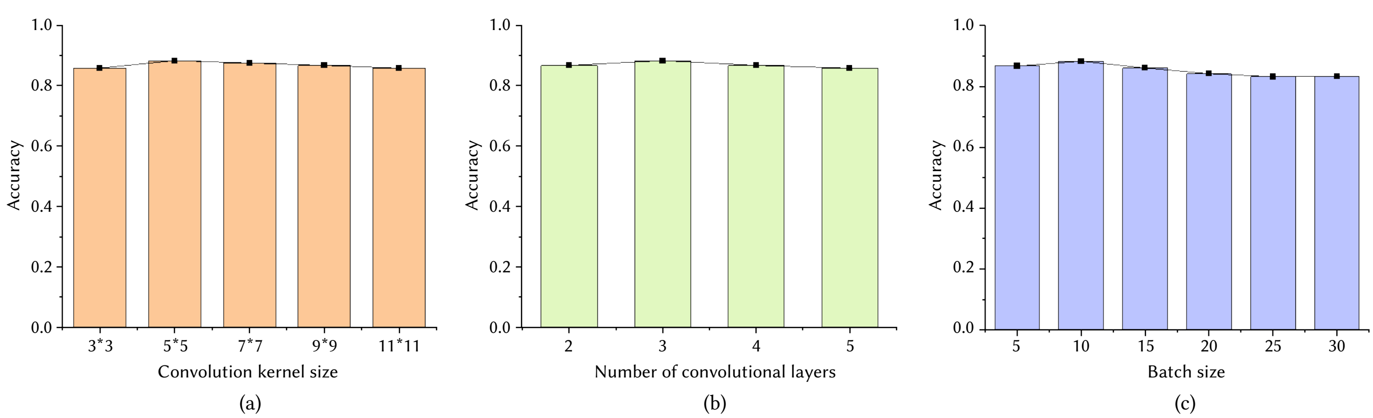


Fig. 9. Relationship between CNN network parameters and recognition accuracy.

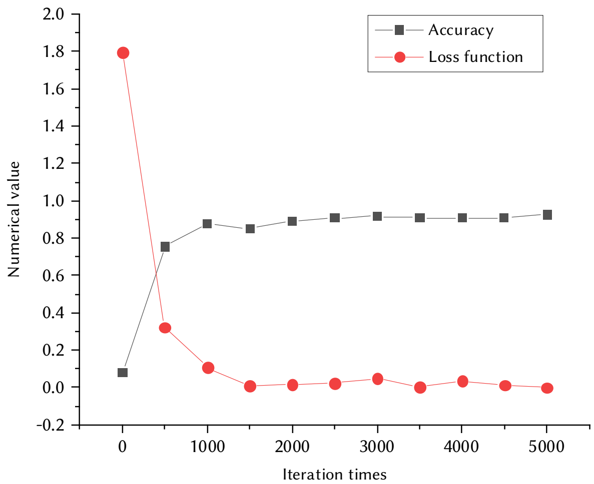


Fig. 10. Training results of the CNN model.

average accuracy of 94.41%. Following closely, the proposed algorithm achieves an average recognition accuracy of 92.76%. For dynamic motion behavior recognition, the SVM model attains an average recognition accuracy of 92.88%, and the BPNN model records an average recognition accuracy of 93.96%. Remarkably, the CNN model showcases the highest recognition accuracy overall for human motion behavior recognition.

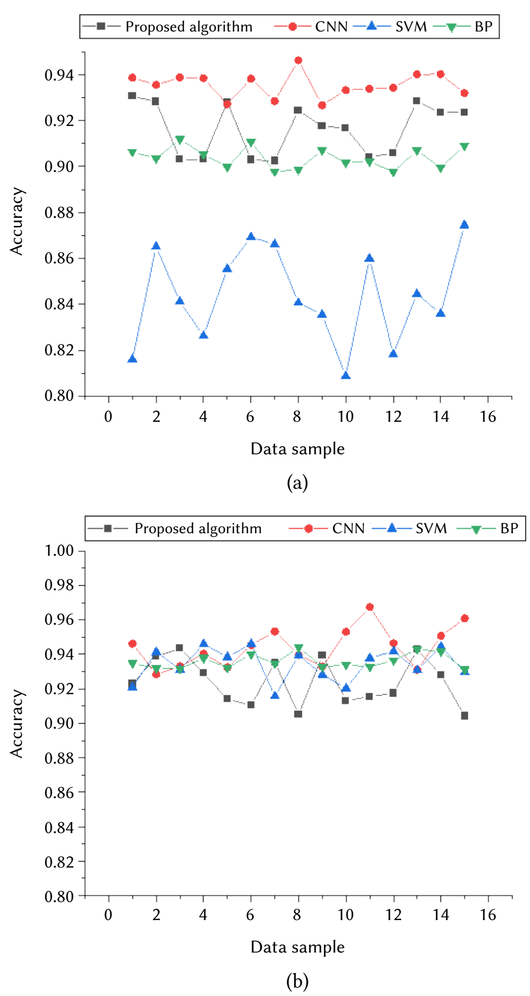


Fig. 11. Recognition accuracy of human motion behavior under different models (A: the dynamic motion behavior, B: the static motion behavior).

D. Discussion

The research findings presented in this paper underscore the substantial capability of CNN in recognizing human motion behavior. In direct comparison with alternative recognition algorithms, the CNN model emerges with the highest recognition accuracy for human motion behavior, boasting an average accuracy of 94.41%. For instance, Dong et al. delved into the application of 3D-CNN in human behavior recognition and achieved a recognition accuracy of 94.7%. In contrast, other recognition algorithms, such as the time-space domain depth CNN, recorded a comparatively lower recognition accuracy of 93.5%, underscoring the superior performance of the CNN model in this context [37]. In comparison to Dong et al.'s research outcomes, the 3D-CNN model employed in this paper showcases notably elevated recognition accuracy, reinforcing its superior performance among various recognition algorithms. The paper adopts a segmentation approach, dividing human motion behavior into static and dynamic categories and conducting separate network recognition, thereby yielding distinct recognition results for different motion behaviors. This segmentation strategy contributes to more accurate identification and differentiation of diverse types of motion behaviors, consequently enhancing recognition precision. Mahmoud conducted research on human behavior recognition utilizing LeNet-5CNN, revealing that with an increase in sample size, the recognition accuracy also improved, reaching a maximum accuracy of 78.54% [38]. In contrast to Mahmoud et al.'s research (2022), the CNN model employed in this paper demonstrates superior performance in human motion recognition. The 3D-CNN model adopted herein achieves a higher recognition accuracy in human motion behavior compared to the LeNet-5 CNN, indicating that the model utilized in this paper exhibits robust adaptability to complex motion recognition tasks.

V. Conclusion

This paper presents a methodology for extracting human motion behavior data scenes from the human living environment, leveraging the IoT framework. The primary objective is to investigate the recognition performance of human motion behavior using CNN. In order to achieve this, the KTH database is selected as the recognition dataset for human motion behavior. Rigorous parameter determination and analysis identify optimal settings for CNN recognition effectiveness, specifying a convolution core size of 5*5, three convolution layers, and a batch size of 10. The training loss function reaches a minimum value of 0.0001. Furthermore, the recognition accuracy of different models highlights CNN's superior performance in recognizing static motion behavior. While the paper introduces the concept of utilizing IoT to collect human motion behavior data characteristics, it acknowledges the challenge of processing and aggregating this data for network training due to its dispersed and complex nature. As a result, the KTH database is chosen as the training dataset instead of the collected data set. Experimental results demonstrate the recognition process and effectiveness of CNN in human motion behavior recognition using the KTH database as the training and testing dataset, yielding a commendable recognition accuracy. However, the limited number of sample data remains a consideration. In order to enhance the model's generalization ability and reliability, future research could explore the collection of more diverse and abundant human motion behavior data, validating it with other publicly available large-scale datasets. Additionally, despite utilizing an optimized CNN model to enhance human motion behavior recognition robustness, challenges may persist in complex scenarios, such as lighting variations, occlusions, and background interference, potentially affecting recognition accuracy. Future studies may delve into exploring more sophisticated network structures and data augmentation techniques to further improve the model's robustness in challenging scenarios.

References

[1] S. Wan, L. Qi, X. Xu, C. Tong, and Z. Gu, "Deep learning models for real-time human activity recognition with smartphones," *Mobile Networks and Applications*, vol. 25, no. 2, pp. 743-755, 2020.

[2] H. Yan, Y. Zhang, Y. Wang, and K. Xu, "WiAct: A passive WiFi-based human activity recognition system," *IEEE Sensors Journal*, vol. 20, no. 1, pp. 296-305,2020.

[3] A. Rudenk, L. Palmieri, M. Herman, K. M. Kitani, D. M. Gavrila, & K. O. Arras, "Human motion trajectory prediction: A survey," *The International Journal of Robotics Research*, vol. 8, no. 39, pp. 895-935, 2020.

[4] A. Jalal, N. Khalid, & K. Kim, "Automatic recognition of human interaction via hybrid descriptors and maximum entropy markov model using depth sensors," *Entropy*, vol. 22, no. 8, pp. 817, 2020.

[5] B. Sun, G. Cheng, Q. Dai, T. Chen, W. Liu, & X. Xu, "Human motion intention recognition based on EMG signal and angle signal," *Cognitive Computation and Systems*, vol. 1, no. 3, pp. 37-47, 2021.

[6] V. Bianchi, M. Bassoli, G. Lombardo, P. Fornacciari, M. Mordonini, & I. De Munari, "IoT wearable sensor and deep learning: An integrated approach for personalized human activity recognition in a smart home environment," *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 8553-8562, 2019.

[7] W. Sousa Lima, E. Souto, K. El-Khatib, R. Jalali, & J. Gama, "Human activity recognition using inertial sensors in a smartphone: An overview," *Sensors*, vol. 14, no. 19, pp. 3213, 2019.

[8] Y. Guo, Z. Xuan, & L. Song, "Foreground target extraction method based on neighbourhood pixel intensity correction," *Australian Journal of Mechanical Engineering*, vol. 3, no. 19, pp. 251-260, 2021.

[9] C. Wei, G. Xue-Bao, T. Feng, S. Ying, W. Wei-Hong, S. Hong-Ri, & K. Xuan, "Seismic velocity inversion based on CNN-LSTM fusion deep neural network," *Applied Geophysics*, vol. 4, no. 18, pp. 499-514, 2021.

[10] L. A. Bolaños, D. Xiao, N. L. Ford, J. M. LeDue, P. K. Gupta, C. Doebeli, & T. H. Murphy, "A three-dimensional virtual mouse generates synthetic training data for behavioral analysis," *Nature Methods*, vol. 4, no. 18, pp. 378-381, 2021.

[11] S. Chebbout, & H. F. Merouani, "A hybrid codebook model for object categorization using two-way clustering based codebook generation method," *International Journal of Computers and Applications*, vol. 2, no. 44, pp. 178-186, 2022.

[12] Y. Peng, H. Tao, W. Li, H. Yuan, & T. Li, "Dynamic gesture recognition based on feature fusion network and variant ConvLSTM," *IET Image Processing*, vol. 11, no. 14, pp. 2480-2486, 2020.

[13] H. Chen, C. Hu, F. Lee, C. Lin, W. Yao, L. Chen, & Q. Chen, "A supervised video hashing method based on a deep 3d convolutional neural network for large-scale video retrieval," *Sensors*, vol. 9, no. 21, pp. 3094, 2021.

[14] G. Lin, Y. Zhang, G. Xu, & Q. Zhang, "Smoke detection on video sequences using 3D convolutional neural networks," *Fire Technology*, vol. 5, no. 55, pp. 1827-1847, 2019.

[15] R. Zhi, H. Xu, M. Wan, & T. Li, "Combining 3D convolutional neural networks with transfer learning by supervised pre-training for facial micro-expression recognition," *IEICE Transactions on Information and Systems*, vol. 5, no. 102, pp. 1054-1064, 2019.

[16] H. Jin, J. Geng, Y. Yin, M. Hu, G. Yang, S. Xiang & H. Zhang, "Fully automated intracranial aneurysm detection and segmentation from digital subtraction angiography series using an end-to-end spatiotemporal deep neural network," *Journal of NeuroInterventional Surgery*, vol. 10, no. 12, pp. 1023-1027, 2020.

[17] X. Lu, H. Yao, S. Zhao, X. Sun, & S. Zhang, "Action recognition with multi-scale trajectory-pooled 3D convolutional descriptors," *Multimedia Tools and Applications*, vol. 1, no. 78, pp. 507-523, 2019.

[18] P. Guenzi, & E. J. Nijssen, "The impact of digital transformation on salespeople: an empirical investigation using the JD-R model," *Journal of Personal Selling & Sales ManageMent*, vol. 2, no. 41, pp. 130-149, 2021.

[19] A. Jalal, M. A. K. Quaid, & K. Kim, "A wrist worn acceleration based human motion analysis and classification for ambient smart home system," *Journal of Electrical Engineering & Technology*, vol. 4, no. 14, pp. 1733-1739, 2019.

[20] M. A. K. Quaid, & A. Jalal, "Wearable sensors based human behavioral pattern recognition using statistical features and reweighted genetic algorithm," *Multimedia Tools and Applications*, vol. 9, no. 79, pp. 6061-6083, 2020.

[21] L. Huang, X. Zhao, & K. Huang, "Got-10k: A large high-diversity benchmark for generic object tracking in the wild," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 5. no. 43, pp. 1562-1577, 2019.

[22] M. A. Khan, Y. D. Zhang, S. A. Khan, M. Attique, A. Rehman, & S. Seo, "A resource conscious human action recognition framework using 26-layered deep convolutional neural network," *Multimedia Tools and Applications*, vol. 28, no. 80, pp. 35827-35849, 2021.

[23] W. Qi, H. Su, C. Yang, G. Ferrigno, E. De Momi, & A. Aliverti, "A fast and robust deep convolutional neural networks for complex human activity recognition using smartphone," *Sensors*, vol. 17, no. 19, pp. 3731, 2019.

[24] Q. Teng, K. Wang, L. Zhang, & J. He, "The layer-wise training convolutional neural networks using local loss for sensor-based human activity recognition," *IEEE Sensors Journal*, vol. 13, no. 20, pp. 7265-7274, 2020.

[25] K. Wang, J. He, & L. Zhang, "Attention-based convolutional neural network for weakly labeled human activities' recognition with wearable sensors," *IEEE Sensors Journal*, vol. 17, no. 19, pp. 7598-7604, 2019.

[26] N. K. Benamara, E. Zigh, T. B. Stambouli, and M. Keche, "Towards a robust thermal-visible heterogeneous face recognition approach based on a cycle generative adversarial network," *International Journal of Interactive Multimedia and Artificial Intelligence,* vol. 7, no. 4, pp. 132-145, 2022.

[27] D. Dejene, B. Tiwari, and V. Tiwari, "TD2SecIoT: temporal, data-driven and dynamic network layer based security architecture for industrial IoT," *International Journal of Interactive Multimedia and Artificial Intelligence,* vol. 6, no. 4, pp. 146-156, 2020.

[28] A. J. F. García, J. C. P. Rodríguez, A. P. Ramos, F. S. Figueroa, and J. D. Gutiérrez, "CompareML: a novel approach to supporting preliminary data analysis decision making," *International Journal of Interactive Multimedia and Artificial Intelligence,* vol. 7, no. 4, pp. 225-238, 2022.

[29] S. M. Lezcano, F. López, and A. C. Bellot, "Data science techniques for COVID-19 in intensive care units," *International Journal of Interactive Multimedia and Artificial Intelligence,* vol. 6, no. 4, pp. 8-17, 2020.

[30] M. L. Ibáñez, A. R. Hernández, B. Manero, and M. G. Mata-García, "Computer entertainment technologies for the visually impaired: an overview," *International Journal of Interactive Multimedia and Artificial Intelligence,* vol. 7, no. 4, pp. 53-68, 2022.

[31] K. R. Gsangaya, S. S. H. Hajjaj, M. T. H. Sultan, & L. S. Hua, "Portable, wireless, and effective internet of things-based sensors for precision agriculture," *International Journal of Environmental Science and Technology*, vol. 9, no. 17, pp. 3901-3916, 2022.

[32] X. Gao, P. Pishdad-Bozorgi, D. R. Shelden, & S. Tang, "Internet of things enabled data acquisition framework for smart building applications," *Journal of Construction Engineering and Management*, vol. 2, no. 147, pp. 04020169, 2021.

[33] A. Abdelbaky, & S, Aly. "Human action recognition using short-time motion energy template images and PCANet features," *Neural Computing and Applications*, vol. 16, no. 32, pp. 12561-12574, 2020.

[34] J. Guo, P. V. Borges, C. Park, & A. Gawel, "Local descriptor for robust place recognition using lidar intensity," *IEEE Robotics and Automation Letters*, vol. 2, no. 4, pp. 1470-1477, 2019.

[35] Y. Liu, X. Yao, Z. Gu, Z. Zhou, X. Liu, X. Chen, & S. Wei."Study of the Automatic Recognition of Landslides by Using InSAR Images and the Improved Mask R-CNN Model in the Eastern Tibet Plateau," *Remote Sensing*, vol. 14, no. 14, pp. 3362, 2022.

[36] D. Jeong, B. G. Kim, & S. Y. Dong, "Deep joint spatiotemporal network (DJSTN) for efficient facial expression recognition," *Sensors*, vol. 7, no. 20, pp. 1936, 2020.

[37] M. Dong, Z. Fang, Y. Li, S. Bi, & J. Chen, "AR3D: attention residual 3D network for human action recognition," *Sensors*, vol. 5, no. 21, pp. 1656, 2021.

[38] H. Mahmoud, "Modern architectures convolutional neural networks in human activity recognition," *Advances in Computing and Engineering*, vol. 1, no. 2, pp. 1-16, 2022.

Hui Li

Hui Li was born in Luoyang, Henan, P.R. China, in 1988. He received his Ph.D. in Physical Education from Kunsan National University in Korea. Now he works in the Department of Physical Education of Luoyang Institute of Science and Technology. His research interest includes on motor learning and control, action-perception coupling, and human motor skill learning. E-mail: 200901701093@lit.edu.cn

Huayang Liu

Huayang Liu was born in Weihai, Shandong Province, People's Republic of China.He is currently a Ph.D. student at Gunsan National University of Korea, and will graduate in July 2024, and has now joined the College of Physical Education of China West Normal University in the field of sports and humanities sociology. E-mail: 690942317@qq.com

Wei Zhao

Wei Zhao was born in Xinxiang, Henan, P.R. China, in 1994. She received a "Master of Teaching Chinese to Speakers of Other Languages from Henan Normal University. She is currently working in the College of Humanities and Social Sciences of Luoyang Institute of Science and Technology. Her research interests are linguistics, psychology, and education. E-mail: 541741134@qq.com

Hao Liu

Hao Liu was born in 1995 in Linyi, Shandong Province, China. He received his doctorate from Gunsan National University . He is currently working in the School of Physical Education and Health of Guangxi Medical University, mainly studying sports biomechanics and sports statistics.Email: linyiliuhao@sina.com

Predicting Consumer Electronics E-Commerce: Technology Acceptance Model and Logistics Service Quality

Cheng-Feng Wu[1], Kunkun Zhang[2], Meng-Chen Lin[3], Chei-Chang Chiou[4]*

[1] Department of Business Administration, National Taipei University (Taiwan)
[2] School of Accounting, Hubei University of Economics, Wuhan (China)
[3] Department of Marketing and Distribution Management, Chien Hsin University of Science and Technology (Taiwan)
[4] Department of Accounting, National Changhua University of Education (Taiwan)

* Corresponding author: ccchiou@cc.ncue.edu.tw

Received 28 June 2023 | Accepted 6 June 2024 | Early Access 19 August 2024

Abstract

In online shopping for consumer electronics, information and physical flows are crucial determinants of consumer purchase intentions. This study examines these factors by integrating the Technology Acceptance Model with logistics service quality, analyzing the relationship between retailers and consumers in e-commerce. The focus is on how information and physical flows, as critical supply chain elements, affect consumers' decisions to purchase online. A structural model and machine learning algorithm with SHapley Additive exPlanations are employed to analyze the data, providing a comprehensive analysis of the Technology Acceptance Model in conjunction with logistics service quality. The findings reveal that attitude, perceived usefulness, and informativeness are the most influential factors affecting consumers' purchase intention. This study contributes to the understanding of consumer behavior in the context of e-commerce platforms for consumer electronic products by integrating the Technology Acceptance Model and logistics service quality theoretical perspectives and analyzing the data using innovative techniques, specifically, Shapley Additive Explanations. This research offers valuable insights into the significant role of various features in shaping consumers' purchase intention in the context of online e-commerce platforms for consumer electrical products.

Keywords

Consumer Behavior, Logistics Service Quality, Machine Learning, SHAP (SHapley Additive ExPlanation), Technology Acceptance Model.

DOI: 10.9781/ijimai.2024.08.001

I. Introduction

Due to the growth of the global economy and the prevalence of Electronic Commerce (EC), effective inter-organizational planning and implementation of value chain processes have become indispensable for the success of online retailers. Information and physical flows are crucial for the relationship between retailers and consumers in e-commerce, as they are integral components of the supply chain [1]. In consumer perspective, consumers have incentives to purchase consumer electronics products online rather than offline due to the dynamic nature of the products and their complexity. In the supply chain perspective, in response to the rapid growth of e-commerce, retailers in E-commerce platforms are continuously adapting their distribution network infrastructure [2]. Technological innovations, including wireless technologies such as Radio-frequency Identification [3] [4] and the Internet of Everything [5] [6], have significantly impacted the e-commerce landscape, emphasizing the paramount importance of logistics service quality in maintaining a competitive edge in the dynamic world of online retail. In this regard, the success of online retail heavily relies on various factors, and one crucial determinant is logistics service quality. The ability to excel in logistics operations has become paramount for online retailers to maintain a competitive advantage in the competitive business environment [7]. Furthermore, studies indicate that the quality of physical distribution services is a critical indicator of customer purchase satisfaction [8] and intention to shop online [9]. Nonetheless, limited studies have considered logistics service quality within the theoretical framework of the Technology Acceptance Model (TAM), where the former is related to the supply (firm) side and the latter is related to the demand (end consumer) side. It is surprising that retailers in e-commerce, as part of the supply chain, have not extensively considered the quality of logistics services, despite its potential impact

Please cite this article as:
C. F. Wu, K. Zhang, M. C. Lin, C. C. Chiou. Predicting Consumer Electronics E-Commerce: Technology Acceptance Model and Logistics Service Quality, International Journal of Interactive Multimedia and Artificial Intelligence, vol. 8, no. 7, pp. 66-85, 2024, http://dx.doi.org/10.9781/ijimai.2024.08.001

on individuals' perceptions and behaviors as end consumers [10]. Given the vital role of logistics service quality in the success of online retailers, it is imperative to explore its influence within established theoretical frameworks. The TAM is a widely recognized framework that explains individuals' acceptance and adoption of technology. Integrating logistics service quality as a factor within the TAM model can provide valuable insights into how it affects customer perceptions, attitudes, and behaviors in the context of online retail when purchasing consumer electronics products. Therefore, this study specifically aims to investigate the influence of logistics service quality as an external factor for consumers of consumer electrical products who use online e-commerce platforms for purchases.

With the ongoing process of economic globalization and the advancement of information technology, electronic commerce has increasingly impacted people's lives [11]. Online shopping, as a viable alternative to traditional shopping, offers numerous advantages. At the end-customer level, online shopping provides a wider selection of goods and services, enabling consumers to easily access product information and compare prices across different distributors [12] [13]. Additionally, consumers can make shopping decisions based on online reviews from other customers [14] [15]. Convenience [16] [17], time and cost savings [12] [18], and flexible transaction methods [17] [18] are among the key factors that drive customers to adopt online shopping. At the firm level, the Internet serves as an effective distribution channel, allowing businesses to reduce costs and overcome geographical barriers [12]. Many businesses have capitalized on this global distribution channel to expand their operations, recognizing it as a substantial growth opportunity [17]. Importantly, electronic commerce strengthens coordination between upstream and downstream supply chain members, fostering cooperation and facilitating technology innovation to enhance operational efficiency. This integration is crucial for businesses to gain a competitive advantage and increase profitability [11]. For instance, in the online environment, consumers conveniently obtain product and retailer performance information from e-commerce platforms, which subsequently influences their purchase intentions [15] [19]. In summary, online shopping provides customers with a wide array of choices, easy access to product information and price comparisons, and the ability to make purchase decisions based on online reviews. Simultaneously, businesses benefit from cost reduction, global expansion opportunities, and improved coordination in the supply chain, leading to enhanced competitive advantage and profitability.

Consumer electronic products, such as digital cameras, smartphones, and DVDs, are a significant and dynamic segment of the global economy [20]. These products are known for their complexity and specialized knowledge requirements for operation and maintenance [21]. In addition, these products undergo frequent updates and introductions of new models, making it challenging for regular consumers to keep up with their technical specifications. Consequently, consumers may end up purchasing electronics that do not meet their expectations in terms of quality [20]. To address this issue, consumers often engage in research activities to gather information about consumer electronics products [14].

When it comes to the pre-purchase stage of buying electrical products, online purchasing offers distinct advantages, particularly in the consumer electronics product category [14]. Online platforms, such as e-commerce websites, provide access to credible user-generated reviews that offer valuable insights into the quality, features, and performance of products. The abundance of online information allows consumers to access diverse opinions and perspectives, enhancing the credibility and reliability of the information they receive. Additionally, after purchasing electrical products, the complexity of these devices necessitates access to detailed product information and technical support. E-commerce platform retailers can provide immediate and non-distance chat-based technical support to assist customers in troubleshooting any product-related issues [21]. Given the significant investment often associated with consumer electronics, online reviews serve as a crucial source of information, helping consumers make informed purchase decisions and avoid potential pitfalls [14]. Therefore, online shopping in e-commerce platforms is the preferred choice for many consumers, especially college students who may lack experience, as it provides access to detailed product information, technical support, and objective reviews. This allows consumers to gather the necessary information to make informed decisions when purchasing consumer electronic products.

However, the supply chain perspective is also crucial in understanding consumers' purchase intentions, particularly when it comes to electrical products, which are physical goods delivered from the supply side to the demand side within the supply chain. While purchasing electronic products online offers numerous advantages, there are also challenges, such as the impact of logistics on consumers' purchase intentions [10]. For instance, during the product delivery stage, many electronic products contain delicate components that can be easily damaged if not handled properly, leading to potential disputes related to returns and exchanges. Effective customer service in online retail involves seamless integration between online ordering and offline delivery, with third-party logistics playing a key role [22]. The differentiation in logistics quality can influence shopping decisions and profitability [23] [24]. Thus, the quality of logistics is central in the supply chain, ensuring the delivery of functional electronic products to end customers [25].

In a theoretical context, research in information systems and e-commerce has explored how users come to accept and utilize new technologies through the Technology Acceptance Model from the perspective of end customers. Logistics service quality is a crucial factor that should be taken into account within the supply chain [10]. Particularly given the complexity and frequent updates of consumer electrical products, on one hand, from an information perspective, product information and technical support are more easily to access from online, facilitating informed purchase decisions in one place. On the other hand, the process of purchasing from the retailer does not end at placing the order on e-commerce platform, as transportation in the supply chain plays a vital role in delivering the product to the end customer. From a transportation perspective, it is important to consider the sensitivity and fragility of consumer electrical products, thus highlighting the need for a framework that incorporates the aspect where the last mile of product delivery is from the retailer to the end customer when investigating the purchasing behavior of consumer electrical products on e-commerce platform.

Prior studies used the methods such as Local Interpretable Model-Agnostic Explanations, Partial Dependence Plots, or ELI5 algorithms in the field of explainable artificial intelligence to interpret and explain machine learning models. However, SHAP's approach to local explanations using Shapley values from game theory is preferred by [26] as it is better than prior techniques. According to some studies [27] [28], Shapley Additive method (SHAP) is better than other statistical methods to interpret the output of machine learning models because it satisfies three key properties: local accuracy, missingness, and consistency. These three properties ensure that the feature attribution method accurately reflects the contribution of each feature to the model output, even when some features are missing or when the model's dependence on a certain feature changes. The SHAP framework is also aligned with human intuition and has a sound theoretic basis, making it suitable for regulated scenarios.

By satisfying these three key properties, the Shapley Additive method provides accurate, comprehensive, and reliable explanations for the prediction of consumer purchase behavior in questionnaire analysis. In the context of consumer purchase behavior in e-commerce platform, local accuracy ensures that the Shapley values correctly attribute the impact of each questionnaire feature on a particular consumer's purchase decision. On the other hand, consistency ensures that the Shapley values provide stable and reliable explanations across similar consumers with similar questionnaire responses. Thus, consistency helps in identifying general patterns and trends in purchase behavior, allowing businesses to make informed decisions based on the reliable interpretation of feature importance. The property of Missingness enables the method to provide meaningful explanations even when certain questionnaire features are not available, improving the applicability and robustness of the analysis. It enables businesses to gain insights into the relative importance of different questionnaire features and understand how they influence consumer decisions. Thus, this knowledge can guide marketing strategies, product design, and customer segmentation, ultimately leading to more effective and targeted approaches to meet consumer preferences and drive sales.

To address the aforementioned issues, this study incorporates the technology acceptance model with logistics service quality to analyze the purchase behavior in consumer electrical products. By building the technology acceptance model, the research examines beliefs, attitude, and purchase intention of end customers to investigate how they accept and use a technology on e-commerce platform. Additionally, logistics service quality was used to capture transportation issue that may affect the end customer accept and use of e-commerce platform for purchasing consumer electrical products. The research aimed to enhance the understanding of consumer behavior in online shopping, offering insights and managerial recommendations to improve logistics operations and customer relationship management in online retailing.

To the best of our knowledge, this is the first study that investigates the purchasing behavior of consumer electric products by incorporating the TAM with logistics service quality. This study contributes to the existing literature in the following aspects. First, the study highlights the inherent complexity and frequent updates of consumer electrical products. By considering the convenience and availability of detailed consumer electrical product information and technical support online, the study emphasizes how individuals' acceptance of technology (online platforms) influences their inclination to purchase consumer electrical products online. Second, the study emphasizes the significance of logistics service quality within the supply chain when investigating consumer purchase behavior on e-commerce platforms. By considering the last mile of product delivery from the retailer to the end customer, the study underscores the importance of transportation and the sensitivity/fragility of consumer electrical products in the purchasing process. This study applies two methods, one for parametric methods and another for non-parametric method. The former analysis confirms the structural model, where the proposed theoretical framework is taken into account, and provides a summary of the overall relationship between the variables. On the other hand, the later one as a complement and robust provides a detailed explanation of how each variable contributes to the prediction. Shapley Additive Explanation method is to interpret the output of the machine learning models and assess the contribution of each feature to the value produced by the model. Third, the research offers valuable insights for both academia and practitioners by providing managerial recommendations. These recommendations enable online retailers to maintain a competitive advantage, promote successful online consumer behavior, and foster profitability in the e-commerce sector.

The following sections are organized in the following manner: in Section II, the related literature is reviewed. Introduction to the method is covered in Section III. The results are included in Section IV. A discussion and conclusion are included in Section V.

II. Literature

This paper aims to investigate the factors that influence consumers' intention to purchase electronic products online. In order to expand upon the TAM, we have introduced logistics service quality as an external variable and incorporated it into the questionnaire design. Through the utilization of a machine learning algorithm, we analyzed the data and applied the SHAP method to interpret the impact of each variable on consumers' ultimate purchase intention. Previous studies have made significant contributions in this field, and in this section, we review the relevant literature encompassing TAM, logistics service quality, and machine learning.

A. TAM Studies

Among the various theories used to predict consumer purchase intention, the TAM has emerged as one of the most widely and successfully employed frameworks in the realm of online consumption. This model has been used in diverse contexts, including clothing [29] [30] [31], luxury products [32], and virtual goods [33] [34] [35]. A study [32] affirms the applicability of TAM in the luxury domain and enriches TAM theory by applying it within the context of online consumer behavior, particularly in the luxury domain. Another research [35] contributes to theory by integrating prospect theory with the TAM to elucidate how perceptions of gains and losses influence the behavioral tendencies of older adults, as well as the role of perceived risk as a barrier to technology adoption. Zhang et al. [31] apply the TAM framework to examine the role of Virtual Try-On (VTO) technology in the online purchase decision-making process of consumers. They explore the relationships between consumers' perceived usefulness, perceived ease of use, perceived enjoyment, and perceived privacy risk with their attitude towards VTO technology, which subsequently influences on their online purchase intention. Although each study has its unique focus and context, their common contribution lies in shedding light on consumer behavior and understanding the factors that influence consumer intentions in various domains. In addition, by incorporating the theoretical foundation of the TAM, researchers can establish a robust and comprehensive framework for studying consumer behavior across various contexts. This framework can effectively elucidate the underlying factors and mechanisms that influence consumer decision-making processes and provide valuable insights for enhancing technology adoption and purchase intentions strategies. The original TAM incorporates key criteria, namely perceived ease of use and perceived usefulness, to gauge the adoption of new technologies. Researchers have expanded the scope of the TAM over time to encompass diverse contexts and concepts, necessitating the inclusion of external components. Table I shows the prior studies where the theory is based on TAM.

On the other hand, some studies have combined the TAM with other issues to predict consumer intentions, such as the Theory of Planned Behavior model [34] [36] [37], task-technology fit theory [38], information adoption model [39], and prospect theory [39]. For example, in their study, Vafaei-Zadeh et al. [37] expanded the combined Theory of Planned Behavior and Technology Acceptance Model by introducing three additional variables to enhance the understanding of purchase intention for electric cars among Generation Y consumers. These variables include price value, perceived risk, and environmental self-image. The results of the study highlight the significant influence of perceived usefulness and perceived ease of use on attitude formation. Furthermore, attitude, subjective norms, perceived behavioral control, price value, and environmental self-image all exhibit positive effects on purchase intention

In summary, e-commerce platforms have become a popular choice for online shopping due to their provision of detailed product

TABLE I. The Summary of TAM-based Studies

Author	Construct	Analysis Method	Theory model	Empirical results	Tools	respondents	context
Liang et al. [66]	Perceived Usefulness (PU), Perceived Ease of Use (PEOU), Performance Risk, Technology Attitudes (TechAtt), Attitude Toward AI (Att), Fashion Involvement (FI), Purchase Intention (PI)	Exploratory Factor Analysis (EFA) Confirmatory Factor Analysis (CFA) Structural Equation Modeling (SEM)	Technology Acceptance Model (TAM)	PU→Att; PEOU→Att; Performance Risk→Att; TechAtt→PI	SPSS 25; Amos 25	313 subjects from the top 10 metropolitan areas in the United States	AI device
Xu et al.[33]	PU, PEOU, Attitude (ATT), Perceived Risk (PR), Group Conformity (GC)	Demographics and Descriptive Statistics; SEM	Extended TAM	PEOU→PU; PEOU→ATT; ATT→PI; PU→PI; PR→PU→PI; GC→PEOU→ATT→PI	IBM SPSS 22.0 and AMOS 22.0	405 young Chinese participants	online paid knowledge
Nguyen et al.[36]	PU, PEOU, ATT, Subjective Norm(SN), PR, Perceived Risk of COVID-19(PRC); Online Shopping Intention(OSI); Online Shopping Behavior(OSB)	Scale reliability test using Cronbach's Alpha, Discriminant and Convergence test using EFA, Pearson's correlation test, Hierarchical regression	Combined Theory of Acceptance Model (TAM) and Theory of Planned Behaviors Model (TPB)	PR→OSB; OSI→OSB	SPSS	638 Vietnamese Internet shoppers	-
Lee and Wong[34]	Personal innovativeness; Subjective Norm(SN); Environmental Consciousness(EC); Price Consciousness(PC);PI; PU; PEOU; Perceived Safety Risk(PSR); Perceived Privacy Security(PPS); Perceived Value(PV); Word-of-Mouth(WoM);	SEM	TAM and TPB	PC→WoM; PU→WoM; PEOU→WoM; PSR→WoM; WoM→PI	SPSS 22.0	277 respondents from social media platforms	on-demand ride-hailing
Wang et al. [38]	Interface Visual Complexity(IVC), Visual Search Efficiency(VSE), Mobile Search(MS), Mobile Payment(MP), Security Precautions(SP), User Experience(UE), Implementation Intentions of Online Shopping(PI)	SEM; Fuzzy-set Qualitative Comparative Analysis (fsQCA)	Task-Technology Fit Theory (TTF) and Technology Acceptance Model (TAM)	VSE→PI; UE→PI; VSE→UE; IVC→VSE; MS/MP/SP→UE	AMOS	College students	-
Rahaman et al.[39]	eWOM Information Quality (IQ); eWOM Information Credibility (IC); eWOM Ease of Use(EOU); eWOM Usefulness (USE); eWOM Information Adoption (INAD); Purchase Intention (PI)	PLS-SEM	Information Adoption Model (IAM) and TAM	USE→PI; EOU→PI; IQ→PI; IC→PI	SPSS 23 SMARTPLS 3.3	College students from Chattogram City of Bangladesh	-----
Lee et al. [99]	Perceived Number of Users(PNOU); Perceived Number of Friends((PNOF)); Perceived Enjoyment(PE); PU; Perceived Desire for Jackpot(PDFJ); PI; Intention to Use(ITU)	CFA; SEM	TAM	ITU→PI; PDFJ→PI; PNOU/PNOF/PE/PU→ITU; PE/PNOF→PDFJ	AMOS 21.0 and SPSS Statistics 21.0	Users of online communities	probability-based items in mobile social network games
Mazzù et al.[69]	PEOU; PU; ATT; PI; Trust Towards the Label (TTL)	CFA; SEM	Front-of-Pack Acceptance Model (FOPAM)	PU→PI; ATT→TTL→PI; PEOU/PU→ATT→PI; PEOU→PU	-	Primary grocery shoppers on Prolific	processed foods
Wong et al.[35]	Perceived Enjoyment(PE); Perceived Effectiveness of Gamification(PEG); Perceived Risks(PR); Adopyion Intention(AI); ATT; PU; PEOU	EFA; CFA; SEM	TAM and Prospect theory	PE→PEG; PU→ATT→AI; PR→AI	SPSS 25.0	Elderly users in residential areas of Suzhou, China	mobile payment.
Wang and Wang[29]	PU, PEOU, ATT, PI, Perceived Performance Risk (PR), Functionality (FUN), Aesthetic (AES); Compatibility (COM)	CFA; SEM	TAM	COM→PU/PEOU/PR; PU/FUN→ATT/PI; PE→PU; PE/AES→ATT; PR→ATT	SPSS; AMOS	-	parent-child smart clothing
Jain [32]	PU, PEOU, ATT, PI, PE, PR, Price Consciousness (PC), Web Atmospherics (WA)	CFA and Hayes Process macro	TAM	PU/PEOU/PE/PC→PI/ATT; ATT→PI	SPSS; AMOS	Luxury fashion consumers in India.	luxury
Vafaei-Zadeh et al.[37]	PU, PEOU, ATT, PR, PI, Subjective Norms (SN), Perceived Behavioural Control (PBC), Price Value (PV), Environmental Self-Image (ESI), Infrastructure Barrier (IB)	PLS (Partial Least Square)-SEM	Combined Theory of Planned Behavior and Technology Acceptance Model (C-TAM-TPB)	PU/PEOU→ATT; ATT/SN/PBC/PV/PR/ESI→PI	-	Generation Y consumers in Malaysia	electric vehicles
Chidambaram et al.[30]	Virtual Try-On (VTO), PU, PEOU, PE, PR; PI	Hayes's Process macros	TAM	Attitude towards VTO mediated the relationship between PU and PI; PR negatively moderated the relationship between PU and Attitude towards VTO; PE positively moderated the relationship between PU and PR and PI mediated through Attitude towards VTO.	-	Millennial respondents in the Southern part of India	online apparel
Zhang et al. [31]	PU, PEOU, PE, PI, Perceived Socialization (PS) Perceived Product Risk (PROR)	PLS-SEM	TAM	PEOU→PU/PE; PU/PE/PRIR→ATT; ATT→PI	-	Online consumers	garment

TABLE II. The Summary of Logistics Service Quality Studies

Author	Factors of logistics service quality	Research variables	Method	finding	respondents
Choi et al.[44]	Quality of Information(QI); Quality of Delivery(QD); Quality of Order(QO); Price of Delivery(PD); Customer Service(CSer)	Service Quality, contain: QI; QO; CSer; PD; Order Procedure (OP); Accuracy(ACC); Order Accuracy(OA); Logistic Service Quality (LSQ); Customer Satisfaction(CSat); Repurchase Intention(RPI)	Validity test; Reliability test; Correlation analysis	QO→CSat; QI→CSat; QD→CSat; PD→CSat; CSer→CSat; CSat→RPI;	Young Chinese customers with experience purchasing products online
Zheng et al. [25]	Order Quality (OQ), Customization Service Quality (CSQ), Response Quality (RQ), Delivery Quality (DQ), Order Discrepancy Handling Quality (ODHQ)	Customer Satisfaction (ECS), Customer Trust (ECT), Customer Loyalty (ECL); OQ; CSQ; DQ; RQ; ODHQ	Exploratory Factor Analysis (EFA); Confirmatory Factor Analysis (CFA); Structural Equation Modeling (SEM)	OQ/CSQ/DQ/ODHQ→ECS; CSQ/RQ/ODHQ→ECT; ECS/ECT→ECL	An online research firm in China
Jiang et al.[45]	Personnel Contact Quality (PCQ); Delivery Quality (DLQ); Infomation Quality (IMQ); Timeliness Quality (TLQ); Empathy Quality (EPQ)	PCQ; DLQ; IMQ; TLQ; EPQ; Satisfaction (SAT); Perceived Importance (PIM)	Hierarchical regression analysis; Importance-Performance Analysis (IPA); SEM	PCQ/EPQ/TLQ/PIM→SAT	Online consumers of fresh food
Hu et al.[100]	Customized Logistics Services (CLS)	Customized Logistics Services (CLS); Satisfaction Level (SAT); Product Type(PT)	EFA; two-way ANOVA	The results indicate that CLS positively impacts SAT. PT does not have moderate effect on the relationship between CLS and SAT.	Tmall.com in China
Oh et al.[42]	Delivery Service Quality(DSQ); Delivery Information Service(DIS); Return Logistics Service(RLS); Delivery Stability(DS); Eco-Friendliness(EF);	DSQ; DIS; RLS; DS; EF; Customer Satisfaction(CS); Intention to Reuse(IRU)	Statistically analyzed	Logistics service quality positively influences CS and IRU (the most significant factor was DS); DSQ/DS/DIS→IRU; CS→IRU	Korean consumers
Dong [43]	Integrity of Delivered(INT); Accuracy of Delivery Time(ADT); The Correctness of the Delivered Goods(CORD); Service Attitude of Delivery Staff(SATT); Delivery Speed(DSP); Whether the Logistics Information is Updated in Time(INF); The Outer Packaging of the Goods is Reasonable and in Good Condition(PAC)	INT; ADT; CORD; SATT; DSP; INF; PAC	Smart sensor technology	Under JD's selfoperated logistics distribution model, users pay the most attention to the INT, ACC, and SATT of the delivery personnel. Under the third-party logistics distribution model of Taobao, the main influencing factors are INT, ACC, PAC	Online evaluation surveys form JD and Taobao

information, technical support, and objective reviews. These features empower consumers to make well-informed decisions when purchasing products. However, the existing literature on consumer behavior in the context of online shopping for consumer electrical products is limited. Given the complex nature and continuous advancements in consumer electronic products, consumers often need to invest more time and effort in evaluating how well these products meet their specific needs and preferences. This includes not only assessing the features and specifications of the products but also considering the post-sale support and functionality. Offline decision-making for consumer electrical products within a short time frame can be challenging. Therefore, this study aims to investigate consumer behavior specifically in the context of online shopping for consumer electrical products.

B. Logistics Service Quality Studies

Online shopping differs from offline shopping in that they are electronic retail markets where goods are transported between individual customers and businesses through logistics. The perception of logistics service quality directly impacts consumers' online purchase intentions. Consequently, researchers recognize logistics service quality as a significant factor impacting consumers' online shopping intention, leading to the continuous development of study on the relationship between logistics service quality factors and consumer purchase behavior. Table II shows the prior studies where the theory is based on logistics service quality.

Some studies examine logistics service quality as an independent variable, along with other factors, to identify consumers' adoption of online purchase behavior [1] [40]. Gao [1] divided the quality of the blockchain system in cross-border e-commerce into three dimensions: commodity information quality, logistics service quality, and payment security. They studied the influence mechanism of blockchain technology application on consumers' willingness to purchase in cross-border e-commerce. Cang and Wang [40] explored the key variables influencing the online purchase intentions of fresh agricultural goods across different customer segments. The findings from hypothesis testing revealed that product quality, online word of mouth, and logistics service quality exert significant influences on potential consumers. However, the study did not find a significant impact of website information quality on potential consumers.

Other studies focus on establishing the key factors determining the quality of logistics services related to online shopping [25] [41] [42] [43] [44] [45], such as timeliness, empathy, information quality, and delivery stability. Through various empirical analysis methods, Oh et al. [42] determined that logistics service quality plays a constructive role in shaping customer satisfaction and intention to engage in future transactions within the context of overseas direct purchases. Specifically, among the various dimensions of logistics service quality, delivery stability emerged as the most influential factor. In another study [43], an examination of e-commerce data and online assessment surveys was undertaken to investigate and evaluate the significance of factors influencing the quality of logistics services and customer satisfaction levels across various distribution models. The findings indicated distinct patterns for different logistics models. Specifically, under JD's self-operated logistics distribution model, users placed high importance on the integrity of delivered goods, the accuracy of delivery time, and the service attitude of the delivery personnel. On the other hand, for Taobao's third-party logistics distribution model, the primary influencing factors included the integrity of delivered goods, the accuracy of delivery time, the importance of outer packaging, and the significance of product integrity. These findings highlight the varying considerations and priorities of customers under different distribution models in terms of logistics quality.

In sum, while numerous studies have explored the impact of logistics service quality on consumer behavior, research specifically focused on consumer electrical products is limited. Nevertheless, it is crucial to take into account the delicate nature of consumer electrical products during transportation. In operation and supply chain perspective, effective product development hinges on customer experience, as the identification and prioritization of pertinent factors contribute to a comprehensive understanding and successful product outcomes [46]. Furthermore, the significance of customer experience aspects varies across distinct product categories [47], suggesting that comprehending these variations would inform product development strategies that align with operation and supply chain strategy. Thus, it becomes necessary to include the final stage of product delivery, known as the last mile, which spans from the retailer to the ultimate customer, when examining the buying patterns of consumer electrical products on e-commerce platforms. This emphasizes the significant role of the perception of logistics service quality in influencing consumers' intention to make online purchases. Integrating the Technology Acceptance Model with logistics service quality helps predict consumers' online purchase intention for consumer electronic products, highlighting the significance of a seamless customer experience in e-commerce.

C. Machine Learning Studies

Machine learning (ML) constitutes a significant and relatively nascent facet of artificial intelligence, involving the training of computer programs to execute tasks and acquire knowledge from the gained experience. As these programs accumulate additional experience, their practical performance in these tasks is enhanced. Consequently, machines can derive decisions and predictions based on data [48]. Table III displays previous studies that utilize questionnaire surveys to predict outcomes in diverse scenarios.

Several investigations have utilized questionnaire data and applied machine learning techniques to predict consumers' propensity to purchase electric vehicles [49] [50] [51] [52]. Additionally, studies have examined purchase behavior or intention for other products, such as "holiday homes" [53], self-defense tools [54], and organic products [55]. Commonly, these studies assess the performance of diverse machine learning algorithms, including Random Forest (RF), Logistic Regression, Decision Trees, Support Vector Machine (SVM), Gradient Boosted Trees (XGBoost, CatBoost, etc.), and Neural Network (KNN, ANN, etc.), to ascertain the most effective one. The choice of optimal algorithms varies depending on the specific context. For instance, Taghikhah et al. [55] applied four machine learning algorithms (SVM, LR, DT, and RF) to analyze consumers' wine preferences, and the RF algorithm yielded the highest accuracy of 89%. Conversely, in another study [56] concerning the application of machine learning in differentiating dampness-heat patterns in patients with type 2 diabetes mellitus in Chinese medicine, SVM outperformed RF.

In real-world machine learning applications, interpretability of models can at times outweigh accuracy [57]. The SHAP method is employed to interpret predictions made by the most effective machine learning models by quantifying and ranking the significance of each variable to the target variables. Within the medical domain, Ballester et al. [58] utilized the XGBoost model to assess predictors of suicide risk and employed graphical representations of SHAP values to interpret the associations of each variable with the outcome, revealing whether it acts as a protective factor or a risk factor. Similarly, Huang and Huang [59] employed Shapely Additive Explanations to visualize the relationships between continuous covariates and the risk of sleep disorders utilizing the National Health and Nutrition Examination Survey dataset. Fan et al. [60] developed various machine-learning models based on tryptophan hydroxylase-2 methylation and environmental stress to identify patients with major depressive disorder. SHAP values were utilized to demonstrate the differential effects of each feature on the outputs of the BPNN model. In their study, higher SHAP values corresponded to a higher probability of patients having a major depressive disorder.

To summarize, SHAP provides a visual, intuitive, and comprehensive approach to augment the interpretability of ensemble models. It aids in comprehending and interpreting the entire model, as well as visualizing feature attributions at the individual observation level for any machine learning model [57]. This paper represents the initial attempt to incorporate the interpretability of machine learning models in predicting consumers' purchase intentions. Whereas preceding studies have predominantly focused on enhancing the accuracy of purchase intention models, this paper presents the initial endeavor to employ SHAP values and associated visualizations in the quest to enhance the explainability of the TAM with respect to logistics service quality. By exploring the underlying factors that drive consumers' purchasing behavior, this research aims to enhance the understanding and applicability of the TAM with the logistics service quality framework. Through the utilization of SHAP values, researchers can gain valuable insights into the relative importance of different variables, thereby improving the transparency and interpretability of machine learning models in predicting and explaining consumers' purchase behavior.

III. Hypotheses

A. Hypotheses

Logistic Service Quality is defined as "logistics services relating to all the problems in the process of shipping goods"[44]. Logistics plays a crucial role in facilitating the transfer of goods from suppliers to consumers. In the context of e-commerce, where transactions occur remotely without face-to-face interactions between consumers and salespeople, the quality of logistics service assumes a role similar to that of sales staff in traditional retail settings [42]. As a result, the quality of logistics service directly influences consumers' perception of their online shopping experience. It has been confirmed that logistics fulfillment quality has a significant influence on the perceived usefulness of e-procurement services [10]. Fu [61] found that management service, platform technology, and the application effect of an intelligent logistics information platform have significantly positive influences on the user's perceived ease of use. Furthermore, the security of using delivery drones by logistics service providers will significantly influence consumers' perceived ease of use of this new technology [62]. Prior research has established a significant relationship between users' perceived ease of use of IT tools and logistics process quality [63]. Similarly, Jain et al. [64] found a positive correlation between mobile service quality, encompassing both forward and reverse logistics, and consumers' perceived usefulness of mobile shopping. In light of this, we have introduced logistics service quality as an innovative external variable within the TAM framework. Based on this premise, we propose the following hypothesis:

Hypothesis 1: The influence of consumers' perception of logistics service quality on perceived usefulness in the context of online e-commerce for purchasing consumer electronic products is significant.

Hypothesis 2: The influence of consumers' perception of logistics service quality on perceived ease of use in the context of online e-commerce for purchasing consumer electronic products is significant.

Perceived ease of use in TAM is defined as the "degree to which an individual believes that the usage of a particular technology does not

TABLE III. Attributes, Algorithms, and Data Mining Techniques Frequently Used to Predict Questionnaire Surveys in Various Scenarios

Author	Attributes/Variables used	Algorithm	Performance	Data mining technique	Tools	Preprocessing Technique
Li et al. [53]	Enduring Involvement,Destination Familiarity,Place Attachment,Purchase Intention,Air Quality, Air Quality, Age, Gender, Education, Income, Package Tour, Family, Friend, Colleague, Travel Experience	Decision Trees (DT), Support Vector Machine(SVM), AKMC	AKMC=82%; KNN=53%; DT=58%; SVM=53%	Classification, Clustering	MATLAB; R2019b	Correlation analysis
Borres et al.[54]	Understanding Safety, Perceived Risk, Self-Effificacy, Perceived Severity, Perceived Behavioral Control, Subjective Norm, Attitude, Perceived Safety, Purchase Intention, Buying Impulse	DT, Random Forest Classifier (RFC), and Deep Learning Neural Network (DLNN)	DT<60%; RFC=96%; DLNN=97.7%	Classification, Neural Network	SPSS 25, Python	Normalization, Correlation analysis
Jia et al.[50]	Electric Vehicle,Household income,Home own,Household size,Young child,Household vehicle,Urban rural,Population density,Price,Place,Age,Gender,Education,Race,Multi-job,Occupation,Car sharing, Time to work,Year mile	DT, Random Forest (RF), Logistic Regression (LR), Naïve Bayes (NB)	NB=0.878; LR=0.795; RF=0.999; SVM=0.993; DT=0.999	Classification, Regression	Python	Synthetic Minority Over-sampling Technique (SMOTE)
Jia [49]	Household income, Home own, Houschold size, Young child, Household vehicle, Urban rural, Population density, Price, Place, Age, Gender, Education, Race, Multi-job, Occupation, Car sharing, Time to work, Year mile	LR, NB, SVM, DT, RF	NB=0.650; LR=0.661; RF=0.924; SVM=0.888; DT=0.908	Classification, Regression	Python	Factor analysis; Regression analysis; SMOTE
Shu et al. [51]	Range Anxity, Climatic Conditions, Technical Maturity, Radation Injury, Physical Discomfort, Accident, Cost-in-use, Acquisition Cost, Maintenance of Value, EV Charging Facilities, Charging Time, Charging Convenience , Social Needs, Preference and Trust Rank, Environmental Conservation	BERT-Att-BiLSTM BERT-TextCNN	BERT-TextCNN (MaF1=0.92); BERT-Att-BiLSTM(F1=0.90)	NLP, Classification	Python	Labels classification, Semantic identification, Emotion analysis.
Sobiech-Grabka et al. [52]	Period of availability, Available amount of subsidy, Limit of car price, Eligible cars, Weakness	Classification and Regression Trees (CART); KNN; SVM; RF	CART=0.870 KNN=0.830 SVM=0.943 RF=0.9867	Classification, Regression, Neural Network	Python	Descriptive statistics
Christidis and Focas [101]	Age, gender, living area, availability of cars and public transport, frequency of trips, duration, distance, inter-modality, Long-distance trips, Attitude	Gradient Boosting	AUC≈0.80	Classification	Python	Descriptive statistics
Lu et al. [48]	Perceived usefulness, Perceived ease of use, Consumer factors, Cross-border e-commerce platform factors	ML	-	Deep Learning, Neural Networks	SPSS, Python	Descriptive statistics; Regression analysis
Carreón et al.[102]	Advert Viewing Time, Purchase Intention, Demographics	SVM, XGBoost, LR	-	Classification, Gradient Boosted Regression Tree, Regression	SPSS, Python	t-test
Taghikhah et al. [55]	Gender, Age, income, Average household size, Education level, Attitude, Perceived Behavioral Control, Habit, Hedonic goals, Gain goals, Normative goals, Social norms, Emotions, Spontaneous urge	SVM, LR, DT, RF	3 classes:(Acc) SVM=78%, LR=78%, DT=86%, RF=89%	Classification, Regression, Density-based clustering	Python	Standardization descriptive analysis, Correlation analysis
Ballester et al. [58]	General Health, SocloeconomIc Status, SRQ-20 Total Score, Bodily Paln, Physlcal Functloning, currently Studying, Sex, Wtalty, Mental Health, Age	XGBoost	XGBoost (AUC=0.71)	Gradient Boosted Trees; SHAP	R	Statistical analysis
Ghorbany et al.[103]	20 KPI indicators, such as Financing Cost, Value for Money, Construction Period, etc.	Copula Bayesian Network (CBN)	CBN=91%	Neural Network, SHAP	SPSS, Python	Statistical analysis; Correlation analysis
Liu et al. [56]	Slimy yellow tongue fur, Slippery pulse or rapid-slippery pulse, Sticky stool with ungratifying defecation, Red tongue, Bitter taste in mouth, Obesity, Thick tongue fur, Halitosis, Dry mouth and thirst, Sticky and greasy in mouth, Heavy body, Constipation, Deep-colored urine, Heavy sensation of head	ANN, KNN, NB, SVM, XGBoost, RF	AUC: XGBoost=0.951; SVM=0.945; ANN=0.947; KNN=0.922; NB=0.922; RF= 0.941	Neural Network; Gradient Boosting; Classification; SHAP	python	-
Fan et al. [60]	gender, CTQ and NLES scores, 25 TPH2 CpG sites (TPH2–11–86, TPH2–11–121, TPH2–11–154, etc.)	BPNN, RF, RBF-SVM, POLY SVM	AUC: RBF-SVM=0.864; POLY SVM=0.832 BPNN=0.988 RF=0.906	Neural Network; Classification; SHAP	Python, SAS, R	Normalization

TABLE III. Attributes, Algorithms, and Data Mining Techniques Frequently Used to Predict Questionnaire Surveys in Various Scenarios (Cont.)

Author	Attributes/Variables used	Algorithm	Performance	Data mining technique	Tools	Preprocessing Technique
Yao et al. [90]	Gender, Age, Marriage status, Education, Working years of psychotherapy, Practice qualification, Licensed psychiatrist, Have professional supervisor, Have professional personal experience, Professional background, Assessment of possible side effects in psychotherapy, Possible causes of side effects in psychotherapy.	RF, XGBoost, CatBoost, LR, AdaBoost, SVM	AUC: RF=0.717, XGBoost=0.689, CatBoost=0.694, LR=0.675, AdaBoost=0.653, SVM=0.629	Classification; Gradient Boosting; SHAP	Python	SMOTE
Huang and Huang [59]	Variables from The National Health and Nutrition Examination Survey	XGBoost, RF, AdaBoost, ANN	XGBoost=0.87, RF=0.82, ANN=0.83, AdaBoost=0.84	Neural Network; Classification; Gradient Boosting; SHAP	Python	Statistical analysis
Ramkumar et al.[104]	Age, Weight, Height, Body mass index, Baseline data (SF-36 pain, KOS-ADL, IKDC Subjective etc.)	GNB; XGBoost; RF; LR; isotonically calibrated XGBoost; sigmoid calibrated XGBoost; and an ensemble soft-voting classifier composed of LR, RF, and XGBoost.	Model for MCID (AUC) GNB=0.72, LR=0.88, isotonic=0.74, RF=0.86, sigmoid=0.94, Ensemble=0.81,	SHAP; Gradient Boosting; Neural Network	Python; R	All ordinal variables were converted to continuous variables.
Wang and Xu [105]	Basic data features, user features, product features, and user product features, totaling 30 features	Fuzzy Support Vector Machine (FSVM); AdaBoost-FSVM; AdaBoost-SVM; SVM; LR; RF; XGBoost	ACC: AdaBoost-FSVM =0.849; AdaBoost-SVM = 0.8112; FSVM=0.7912; SVM=0.7704; LR=0.7616; RF=0.7899; XGBoost=0.8080	Classification; Regression	Python	Select valuable features

require extra effort", On the other hand, perceived usefulness is "the degree to which a person believes that using a particular system would enhance his or her job performance" [65].

Considering the complexity inherent in consumer electronic products, online shopping offers consumers the advantage of intuitive parameter comparisons and access to word-of-mouth recommendations from other users. Furthermore, online shopping enables consumers to make purchases anytime and anywhere, eliminating the need for a substantial continuous time allocation. This enhanced convenience facilitates efficient selection of the most suitable electronics. As a result, consumers are more likely to develop a favorable attitude towards purchasing consumer electronic products online. Notably, a consistent pattern of findings has emerged from prior research employing the TAM framework across diverse domains, highlighting the profound influence of individuals' perceptions on consumer attitudes. This robust association has been observed in various contexts, including electric vehicles [37] [66], mobile food ordering apps [67], ride-hailing services [34], and mobile payment platforms [35]. Building upon the aforementioned discussion, we propose the following hypothesis:

Hypothesis 3: Perceived usefulness positively influences consumers' attitude towards purchasing electronics online.

Hypothesis 4: Perceived ease of use positively influences consumers' attitude towards purchasing electronics online.

Attitude refers to "an individual's positive or negative feeling regarding performing the target behavior". According to the theory of reasoned action, an individual's behavioral intention is contingent upon their attitude towards the behavior [68]. Previous research has consistently demonstrated the impact of consumers' attitudes on purchase intentions across various online domains, including online grocery shopping [69], online luxury products [32], and online garment retailing [31]. Hence, we propose the following hypothesis:

Hypothesis 5: Consumers' attitude towards online purchasing positively influences their intention to use online channels for purchasing consumer electronic products.

B. Research Model

Drawing on the theoretical underpinnings of the TAM, this research endeavors to construct a comprehensive framework (see Fig. 1) that incorporates the dimension of logistics service quality. By delineating a set of constructs and their associated hypotheses, the study seeks to investigate the determinants that influence the purchase intention of online consumers in the context of consumer electronic products.

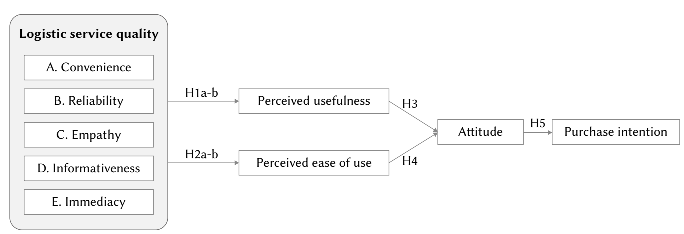


Fig. 1. Conceptual framework.

IV. Method

Fig. 2 presents a comprehensive flowchart that illustrates our methodological process. This flowchart provides a detailed step-by-step description of the methodological process, including survey design, data collection, data processing, model selection, and interpretation.

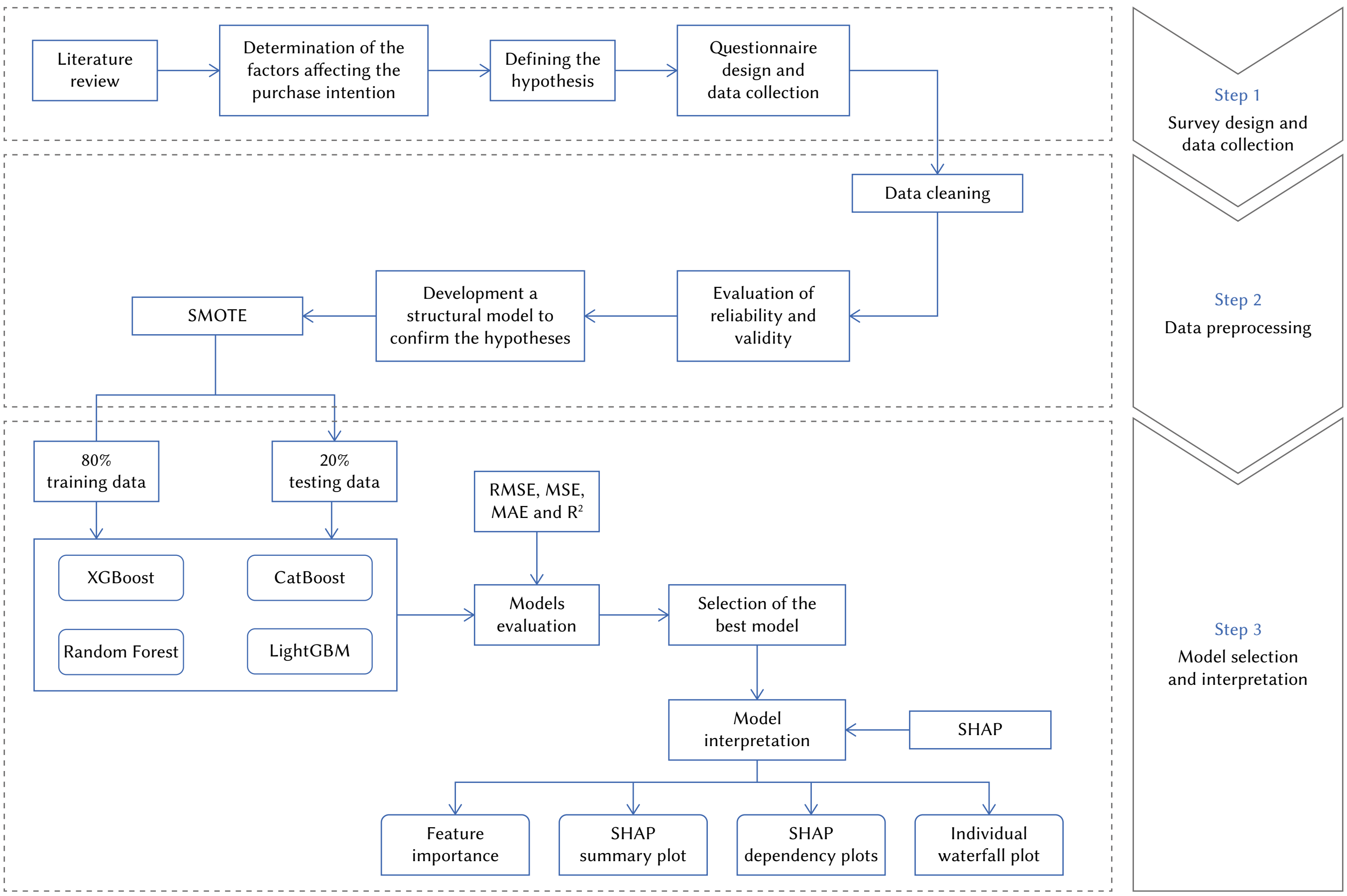


Fig. 2. Methodological process.

A. Techniques

1. Structural Model

SPSS and AMOS, a widely acclaimed and extensively employed statistical software package, provides a diverse array of statistical analysis tools that are instrumental in investigating the TAM with logistics service quality. In the specific context of this study, the utilization of SPSS empowers us to undertake meticulous statistical analyses. AMOS enables the examination of the structural model, which encompasses the relationships between variables and hypothesis testing within the framework of the TAM with a focus on logistics service quality. The employment of SPSS and AMOS ensures the rigor and dependability of the analyses conducted in this study, consequently enhancing the credibility of the research outcomes.

2. Machine Learning Algorithms & Interpreter

For our data analysis, we have employed Random Forest [70], which is widely recognized and has shown excellent predictive performance in recent years. In addition to Random Forest, we have also utilized three other algorithms known for their robustness and effectiveness in prediction: LightGBM, CatBoost [71], and XGBoost [72].

These algorithms possess distinct principles and characteristics. RF is an ensemble learning model that combines multiple Decision Trees (DTs) to achieve predictions with higher accuracy and robustness compared to an individual DT [73].

The computation can be performed using the equation (1) [70]:

$$\text{GiniInde} = \sum \sum_{j \neq i} \left(\frac{f(Y_i\ ,T)}{|T|}\right)\left(\frac{f(T_j,T)}{|T|}\right) \quad (1)$$

where, *T* presents the training dataset. $\frac{f(Y_i\ ,T)}{|T|}$ presents the probability of belonging to category Y_i.

Prokhorenkova et al. [71] introduced CatBoost, a novel gradient boosting technique that effectively handles category features while minimizing information loss. Distinguished from other gradient boosting methods, CatBoost initially applies ordered boosting, which is a modified and efficient gradient boosting technique. This approach proves advantageous for small datasets and effectively handles category features. The underlying base predictor in CatBoost consists of binary decision trees. The estimated output can be calculated as shown in equation (2) [71] [74]:

$$Z = F(x_i) = \sum_{j=1}^{j} b_j\, 1_{\{x \in R_j\}} \quad (2)$$

where $F(x_i)$ is the function of the decision tree of the independent variables x_i, and b_j is the disjoint region that corresponds to the tree's leaves.

The XGBoost algorithm [72] has gained significant popularity across various domains in recent years. Built upon the concept of "boosting," XGBoost combines the predictions of multiple weak learners using additive training techniques to construct a robust learner. To mitigate overfitting and enhance performance, XGBoost incorporates a regularized method formulation [75]. The integrated framework employs random sampling to reduce variance and improve the predictive capabilities of the final model. The estimated output can be calculated as indicated in equation (3) [72]:

$$Z = H(x_i) = \sum_{t=1}^{T} f_t\,(x_i) \quad (3)$$

where x_i represents the independent variables, and $f_t(x_i)$ is each tree output function.

TABLE IV. Questionnaire for Technology Acceptance Model and Logistics Service Quality

When I choose e-commerce way to purchase consumer electronic products online,

Convenience	Qa1: The convenience of payment enhances my online shopping experience.
	Qa2: Setting the pick-up time from logistics is convenient.
	Qa3: Returning goods is convenient.
Reliability	Qa4: The logistics service provider consistently delivers services as promised.
	Qa5: The order delivery process is smooth and hassle-free.
	Qa6: The logistics service providers accurately process my orders according to my requirements.
Empathy	Qa7: The logistics services provided are flexible and customizable.
	Qa8: The transportation and delivery time is appropriate.
	Qa9: The logistics service provider understands my demands well.
Informativeness	Qa10: I can easily access timely and accurate logistics distribution information.
	Qa11: It is easy for me to check the logistics distribution information.
	Qa12: I receive complete and sufficient feedback regarding logistics distribution information.
Immediacy	Qa13: The time between placing an order and receiving the delivery is short.
	Qa14: Deliveries arrive on the promised date.
	Qa15: The back-order time for requisitions is minimal.
Perceived usefulness	Qd1: Buying consumer electronic products online improves my shopping efficiency.
	Qd2: Shopping for consumer electronic products online makes the shopping process easier for me.
	Qd3: Buying consumer electronic products online enhances my shopping ability.
Perceived ease of use	Qb1: It is convenient to purchase consumer electronic products using an e-commerce platform.
	Qb2: It is easy to understand how to buy consumer electronic products using an e-commerce platform.
	Qb3: Learning to purchase consumer electronic products online is effortless for me.
	Qb4: Buying consumer electronic products through an e-commerce platform does not require much mental effort.
Attitude	Qc1: I have a positive attitude on purchasing consumer electronic products online.
	Qc2: Using an e-commerce platform to buy consumer electronic products is a good idea.
	Qc3: It makes sense to purchase consumer electronic products online.
Purchase intention	Qf1: I am likely to buy consumer electronic products online.
	Qf2: I am inclined to consider purchasing consumer electronic products online.
	Qf3: It is certain that I will explore buying consumer electronic products online.

LightGBM, a gradient-boosting decision tree algorithm, has been proposed by Microsoft Research. Acknowledged for its rapidity and exceptional performance, LightGBM finds applications in a variety of machine learning tasks, including ranking, regression, and classification. The primary objective of this algorithm is to enhance computational efficiency in resolving challenges related to predictive analysis on large-scale datasets [76]. The mathematical expression representing LightGBM is provided in equation (4) [77]:

$$F_M(x) = \sum_{m=1}^{T} \gamma_m h_m(x) \tag{4}$$

where M represents the maximum number of iterations and h_m (x) denotes the base decision tree.

The SHAP algorithm [26] offers a method to determine the impact of features in tree-based models, addressing the absence of a direct prediction equation in Decision Trees (DTs) and their derivatives. SHAP values quantify each feature's average marginal contribution, enhancing the understanding of model predictions. This interpretation can be achieved using the shap.explainers.tree function within the SHAP package, which analyzes the trained model and test data predictions [78]. Unlike traditional feature importance analysis, SHAP provides detailed insights into how features affect individual predictions, highlighting their positive and negative impacts on the outcome. This approach, based on the Shapley value from game theory, allocates "credit" to features, allowing for a nuanced understanding of model behavior [26] [79]. SHAP's methodology, applying game-theoretic principles to model interpretation, offers a comprehensive framework for analyzing feature contributions across various machine learning techniques [80]. Notably, in this particular study, algorithms such as CatBoost were employed for conducting the SHAP analysis. Based on several axioms to help fairly allocate the contribution of each feature, shapely values are represented by equation (5) [26] [79]:

$$\emptyset_i = \frac{1}{|N|!} \sum_{S \in N(i)} \frac{|S|!(|N|-|S|-1)!}{N} [f(S \cup \{i\}) - f(S)] \tag{5}$$

where f(S) corresponds to the output of the CatBoost model, S to the set of features, and N represents the whole set of entire features. The ultimate contributions or Shapley value of feature i($\emptyset_i$) is calculated as the average of its contributions over all permutations of a feature set.

Consequently, the inclusion of features into the set is performed individually, and the resulting change in the model's output serves as an indicator of their significance. This approach leverages the utilization of feature orderings, which play a pivotal role in influencing the observed variations in the model's output, particularly in cases where correlated features are present.

B. Research Methodology

1. Data Collection

The questionnaire was developed based on the conceptual framework and hypotheses delineated above, aiming to investigate the factors influencing consumers' online purchase intention of consumer electronic products. The items measuring Convenience and Reliability were adapted from the work by Jiang et al. [81]. Empathy was adapted from other works [82] [83]. Informativeness was adapted from the study by Jiang et al. [45]. Immediacy was adapted from the study by Huang et al. [84]. Perceived usefulness was adapted from the work by Jain [32]. Perceived ease of use and Purchase intention were adapted from the study by Lee and Wong [34]. Attitude was adapted from the work by Mazzù et al. [69]. Table IV displays the questionnaires used in this study, which were translated from Chinese to English. In total, 32

TABLE V. The Results of Reliability & Validity

Constructs	KMO	Cronbach's Alpha	CR	AVE
Convenience	0.68	0.73	0.85	0.65
Reliability	0.65	0.66	0.82	0.60
Empathy	0.69	0.74	0.85	0.66
Informativeness	0.63	0.63	0.80	0.58
Immediacy	0.67	0.69	0.83	0.61
Perceived usefulness	0.78	0.78	0.86	0.60
Perceived ease of use	0.78	0.81	0.87	0.63
Attitude	0.68	0.74	0.85	0.66
Purchase intention	0.70	0.78	0.87	0.69

items were included in the questionnaire, addressing these influential factors. The response options for all items utilized a five-point Likert scale, ranging from 1 = "strongly disagree" to 5 = "strongly agree." Sampling error refers to the statistical variation that arises when a sample does not perfectly represent the entire population. To collect the data for this study, an online survey was conducted from March 26, 2022, to September 7, 2022, employing the SO JUMP platform, a renowned Chinese Internet-based survey platform akin to Amazon Mechanical Turk. This survey leveraged prominent Chinese social media conduits, including WeChat, QQ, and Weibo, to disseminate the instrument. The respondent pool consisted of a random sampling of users across these platforms, ensuring a broad demographic representation. A total of 1323 questionnaires were collected, and after removing any invalid responses, 1069 questionnaires were deemed suitable for further analysis. The sample profile showed that 45.5% of the respondents were female and 55.5% were male. Among them, 2.5% were freshmen, 34.9% were sophomores, 23.8% were juniors, 16% were seniors, 22.5% were graduate students, and 0.3% were doctoral students. This approach was designed to minimize sampling error, ensuring the representativeness and validity of our findings. To enhance the accuracy of the research findings, a reliability analysis and a validity analysis were performed on the questionnaire. These analyses aimed to ensure the robustness and accuracy of the collected data.

2. Reliability & Validity

An initial evaluation of reliability was performed using Cronbach's alpha and composite reliability measures, adhering to the suggested thresholds of 0.60 [85] [86] and 0.7 [87], respectively. The reliability results for the constructs are presented in Table V, indicating that all constructs surpassed the minimum values of 0.6 and 0.7 for Cronbach's alpha and composite reliability, respectively. These outcomes provide robust evidence supporting the instrument's reliability.

In terms of construct validity, the assessment focused on evaluating convergent validity [87]. Convergent validity is deemed to be established when the Average Variance Extracted (AVE) of a construct surpasses the threshold of 0.50 [87]. The findings of the convergent validity analysis, presented in Table V, reveal that the AVE values for all constructs exceed 0.50, substantiating the model's constructs in terms of convergent validity.

C. Preprocessing

Following the reliability test and data cleaning process, a total of 1069 remaining data points were selected for implementation in the machine learning algorithm. To represent the feature values of each factor, we calculated the mean of the question scores associated with that particular factor. Subsequent analysis revealed that a majority of the samples exhibited high scores across all feature values. For instance, considering the target variable of purchase intention, it was observed that only approximately 5% of the samples possessed scores below 3. This significant data imbalance can potentially lead to issues such as overfitting or substantial prediction errors when employing machine learning algorithms. To address this situation, we opted to focus solely on the feature values of the target variable. Each value of the target variable was treated as a distinct category, and the Synthetic Minority Over-sampling Technique (SMOTE) [88] was applied to generate new samples based on the original data. In their study, the generated samples can be denoted as follows:

$$S_{new} = S_i + \omega(S' - S_i) \tag{6}$$

where, S_i presents the samples belonging to minority category. S' is the selected sample close to S_i. ω defines the weight.

Oversampling techniques increase the representation of minority category samples by duplicating them, risking overfitting in models. In contrast, undersampling reduces the sample count by removing random samples, potentially wasting valuable data [50]. SMOTE stands out by generating new, unique samples, thus better supporting prediction models and avoiding the drawbacks of simple duplication [50]. Beyond its common use in addressing class imbalance, SMOTE has been applied in survey research, enhancing data quality and analysis [89] [90]. In this study, using SMOTE resulted in a balanced dataset of 3344 samples, supporting in the effective evaluation and comparison of predictive models.

V. Results

A. Structural Model

The aim of this study is to examine the impact of logistics service quality, as an external factor, on consumers of consumer electrical products who utilize online e-commerce platforms for their purchases. First, to validate the theoretical model established, this study employed a structural model to confirm the conceptual framework. Based on the conceptual framework of this study in Fig. 1, the path coefficients and their corresponding levels of significance in the structural model are presented as follows.

Convenience (β = 0.227, p = 0.000), reliability (β = 0.143, p = 0.000), informativeness (β = 0.13, p = 0.000), and immediacy (β = 0.388, p = 0.000) were found to have positive influences on perceived usefulness. However, empathy (β = -0.013, p = 0.652) did not have a significant effect on perceived usefulness. Therefore, H1a, H1b, H1d, and H1e were supported, while H1c was not supported. Additionally, convenience (β = 0.041, p = 0.309), reliability (β = -0.064, p = 0.144), and informativeness (β = -0.017, p = 0.713) did not have a significant effect on perceived ease of use. However, empathy (β = 0.115, p = 0.006) had a positive impact, and immediacy (β = -0.093, p = 0.035) had a negative impact on perceived ease of use. Thus, H2a, H2b, and H2d were not supported, but H2c and H2e were supported. Furthermore, perceived usefulness (β = 0.676, p = 0.000) positively influenced attitude, and attitude (β = 0.073, p = 0.000) positively influenced purchase intention. Therefore, H3 and H5 were supported. However, perceived ease of use (β = -0.02, p = 0.448) did not have an effect on attitude. Thus, H4 was not supported.

In sum, convenience, reliability, informativeness, and immediacy positively impact perceived usefulness, but empathy does not significantly influence it. Only empathy has a positive influence, and immediacy has a negative impact on perceived ease of use. Perceived usefulness positively affects attitude, which in turn positively impacts purchase intention. However, perceived ease of use does not significantly influence attitude.

B. EML Models Training and Testing

This research aims to compare the predictive performance of various EML models, such as XGBoost, CatBoost, LightGBM, and RF, in forecasting consumer purchase behavior on e-commerce platforms. Utilizing questionnaire data enhanced with SMOTE, the study involves supervised EML training and testing, dividing the dataset into an 80% training set and a 20% testing set, with analysis conducted using Python. Parameter optimization for each EML model is crucial for predictive accuracy, involving a grid search via scikit-learn to fine-tune parameters like 'learning_rate', 'n_estimators', and 'max_depth' to enhance model performance and prevent overfitting. These parameters are detailed in Table VI.

TABLE VI. EML Methods Tuning Parameters

Regression Method	Leaning_rate	n_estimators	max_depth
XGBoost	0.2	100	12
CatBoost	0.1	300	10
LightGBM	0.2	500	05
RF	-	800	10

C. Comparison of EML Model Performance

In this study, EML algorithms (RF, CatBoost, LightGBM, and XGBoost) were evaluated using questionnaire data, with their performance compared across metrics like RMSE, MSE, MAE, and R2, detailed in Table VII. The results showed that the training set consistently yielded higher R2 and lower error metrics than the test set for all models, indicating no overfitting. Table VII shows the comparison of the models' predictive accuracy. Notably, CatBoost emerged as the most effective model, achieving the highest R2 value of 0.889 and the lowest values in RMSE (0.386), MSE (0.149), and MAE (0.235), emphasizing its superior predictive performance in this context.

D. Feature Analysis

The interpretability of most ML algorithms has been subject to criticism due to the challenge of comprehending the importance of features and the contribution of individual predictor variables to the final model outcome. However, the development of an accurate online purchase intention prediction model is crucial, as more precise models can effectively capture the relationships between explanatory and response variables. Moreover, it is essential to interpret the model results and translate them into actionable insights. To address these concerns, this study employed a feature importance analysis based on CatBoost to establish the relative ranking of input variables. Furthermore, partial dependence plots using SHAP analysis were employed to gain additional insights into model interpretability, along with an examination of individual sample observations through a waterfall plot.

1. Feature Importance Analysis

The findings depicted in Fig. 3 provide valuable insights into the significant role of various features in shaping consumers' purchase intention within the realm of online e-commerce platforms for consumer electrical products. Notably, attitude emerges as the most influential factor, highlighting the crucial role of consumers' overall attitude towards platform usage in their purchase decisions.

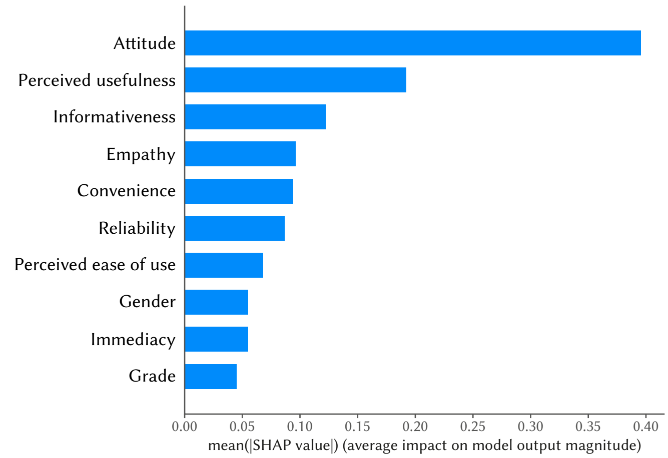


Fig. 3. Features importance using SHAP for CatBoost.

Following attitude, perceived usefulness, and informativeness demonstrate substantial impact on consumers' purchase intention. This underscores the significance consumers place on the practical benefits and logistics-related information provided by the platform during product delivery. The platform's ability to effectively convey useful information and ensure efficient product delivery significantly influences consumers' decision-making process.

Among the moderately important features, empathy exerts its influence on consumers' purchase intention. This finding suggests that consumers' perception of a logistic company's understanding and consideration of their needs plays a role in influencing their decision to engage in online purchases. Additionally, convenience, reliability, and ease of use are identified as contributing factors, albeit to a lesser extent.

Conversely, gender, immediacy, and grade exhibit low importance, indicating their minimal impact on consumers' purchase intention in the context of online purchases for consumer electrical products. This finding implies that consumers' gender, the immediacy of their purchase decision, and their academic grade have negligible influence on their decision-making process in this specific domain.

2. SHAP Summary Analysis

The SHAP summary plot for the CatBoost model, as seen in Fig. 4, visually communicates the significance and impact of various features on purchase intention. The y-axis arranges features based on their mean absolute SHAP values, illustrating their importance, while the x-axis displays the SHAP values themselves. Each feature is represented as a row in the plot, where the color signifies the feature's effect on purchase intention. A secondary color scale on the y-axis

TABLE VII. Accuracy Metrics of EML Methods for Train and Test Set

Models	Train set result				Test set result			
	R^2	MAE	RMSE	MSE	R^2	MAE	RMSE	MSE
RF	0.935	0.201	0.293	0.086	0.835	0.313	0.220	0.469
CatBoost	0.995	0.043	0.078	0.006	0.889	0.235	0.386	0.149
LightGBM	0.977	0.116	0.175	0.031	0.863	0.280	0.428	0.183
XGBoost	0.998	0.022	0.051	0.003	0.866	0.241	0.423	0.179

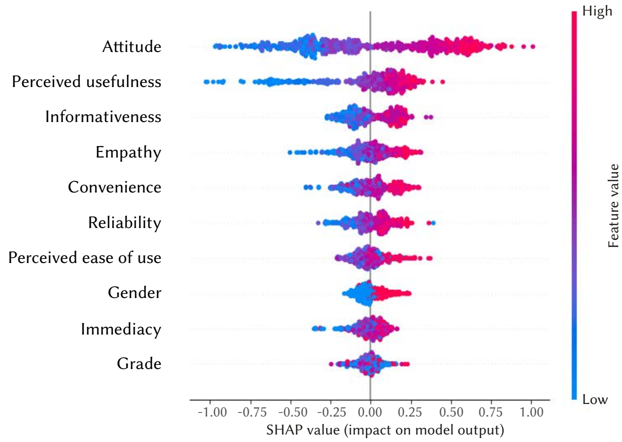


Fig. 4. SHAP summary plot of the CatBoost model
(The higher SHAP value of a feature, the higher purchase intention levels).

denotes the relative importance of each feature, with shades ranging from blue for less importance to red for greater importance, providing a clear, intuitive understanding of how each feature influences the model's predictions regarding purchase intention.

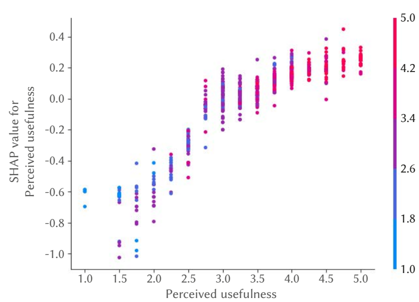
The range of negative SHAP values associated with perceived usefulness is broader, but the maximum positive SHAP value reaches only 0.5. This suggests that high perceived usefulness has a moderate positive impact on purchase intention, while low perceived usefulness significantly and negatively influences purchase intention. The impact of the remaining input variables is relatively narrow, and for the variables ranked lower, the color boundaries appear less distinct, indicating less clarity in their influence on purchase intention.

3. Variables Association Analysis

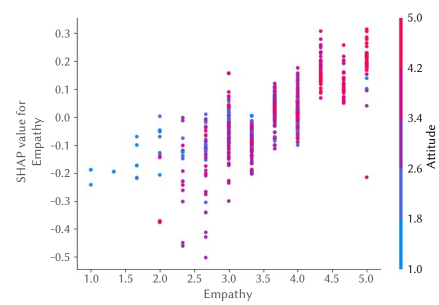
The SHAP summary plot in Fig. 4 provides a comprehensive overview of the relationship between purchase intention and the explanatory variables. To explore deeper into these relationships and their impact on purchase intention, SHAP dependency plots are employed, as illustrated in Fig. 5.

These plots detail how individual features affect the CatBoost model's predictions, with the primary y-axis showing the SHAP value of a feature and the x-axis its actual value. The secondary y-axis's color bar indicates the influence of another feature, highlighting interaction effects. These dependency plots reveal both main and interaction effects, demonstrating how features interact to influence the model's output. Notably, they explore the interactions involving the attitude feature, displaying how it interacts with other variables to impact purchase intention, with color variations representing different levels of attitude's influence.

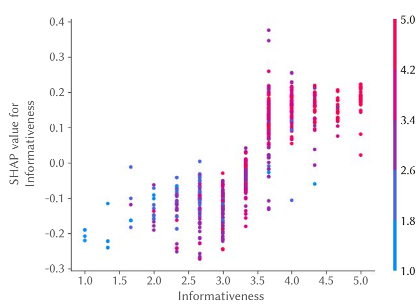
Fig. 5(a) presents the interaction between attitude and perceived usefulness. A positive SHAP value indicates high perceived usefulness and corresponds to a high attitude, displaying a clear upward trend



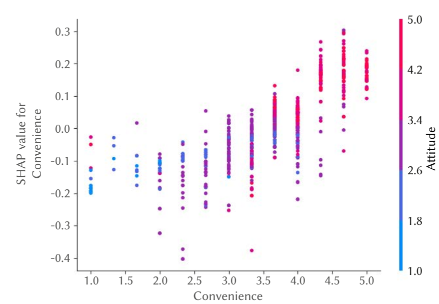
(a) Interaction between attitude and perceived usefulness



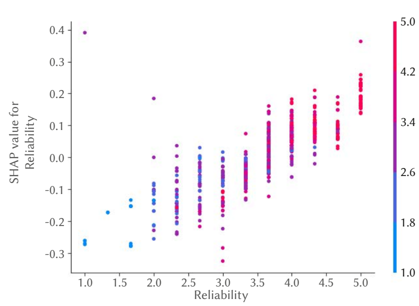
(b) Interaction between attitude and empathy



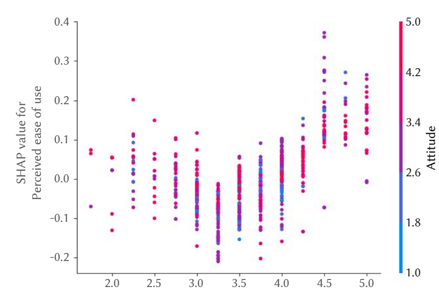
(c) Interaction between attitude and informativeness



(d) Interaction between attitude and convenience

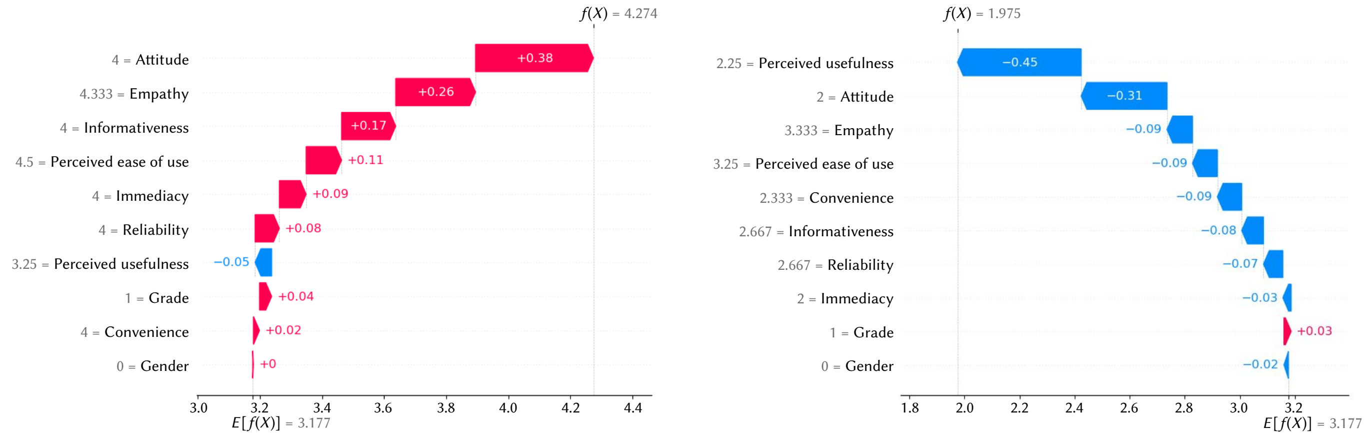


(e) Interaction between attitude and reliability



(f) Interaction between attitude and ease of use

Fig. 5. SHAP dependence plots.



(a) The sample with a high purchase intention

(b) The sample with a low purchase intention

Fig. 6. Explanation of the prediction generated by the CatBoost model using tree SHAP.

in the SHAP value as both attitude and perceived usefulness increase. Conversely, a negative SHAP value is associated with a low attitude and empathy. Figs. 5(b-e) demonstrate the interactions of attitude with empathy, informativeness, convenience, and reliability, respectively. These figures exhibit similar interaction trends to Fig. 5(a).

Fig. 5(f) illustrates the interaction between attitude and ease of use, revealing a U-shaped curve in the relationship between the SHAP value and ease of use. When the ease of use is below 3.33, the SHAP value decreases with an increase in ease of use. However, after surpassing this threshold, the SHAP value increases as the ease of use feature value rises. Notably, the distribution of attitude values lacks a distinct color boundary.

4. Individual Observation Analysis

The conventional attribute importance algorithm provides a global importance value for an attribute across the entire dataset, whereas the SHAP value offers importance values for each individual observation. Local interpretability allows us to assess how the feature values contribute to the prediction score of each observation within the sample.

Fig. 6 presents a waterfall plot based on a single observation of an individual. Fig. 6(a) illustrates the observed values of an individual from a sample with high purchase intention, while Fig. 6(b) represents one of the samples with low purchase intention. The vertical axis on the left represents the input feature and its actual value. E[f(X)] denotes the mean value of the predicted value for the target feature, while f(x) represents the predicted value of the target feature for this specific instance. In the figure, red indicates that the feature increases the purchase intention, while blue indicates that the feature decreases the purchase intention of this sample. The numbers within the arrows indicate the magnitude of influence, and the input features on the left vertical axis are ranked based on the absolute value of the magnitude of influence.

As depicted in Fig. 6(a), the mean value of the predicted purchase intention is 3.177, while the actual purchase intention value for this individual is 4.274. The difference between these two values signifies the impact of various input variables on the purchase intention. Specifically, the variables of grade (+0.04), convenience (+0.02), and gender (0) have a negligible influence on purchase intention. Conversely, perceived usefulness (-0.05), reliability (+0.08), immediacy (+0.09), and perceived ease of use (+0.11) demonstrate a slight positive impact. Moreover, informativeness (+0.17), empathy (+0.26), and attitude (+0.38) significantly contribute to the positive purchase intention.

In Fig. 6(b), the purchase intention prediction value is 1.975 for the depicted individual. For this person, grade (+0.03) demonstrates a weak positive impact. On the other hand, perceived usefulness (-0.45) and attitude (-0.31) significantly negatively impact her purchase intention. The remaining input variables exhibit negative effects with absolute values below 0.1.

Fig. 7 displays the strength of relationships among the variables. The correlation matrix highlights strong correlations between empathy and informativeness, attitude and perceived usefulness, perceived usefulness and immediacy, as well as informativeness and reliability.

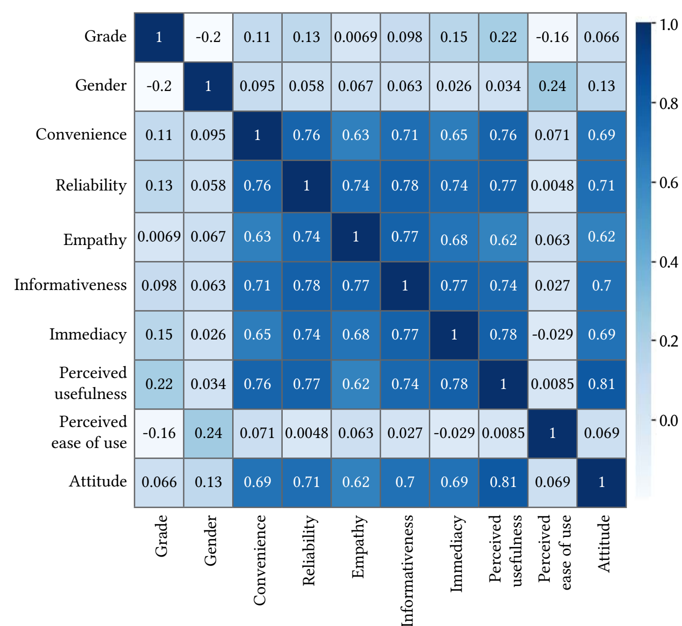


Fig. 7. Correlation matrix for the different variables.

VI. Discussion

The rapid evolution of science and technology has made consumer electrical products important for daily life, impacting communication, information access, and shopping. Consumers typically invest time in researching and comparing electronics, reflecting the tendency of buyers to conduct research before making purchase decisions [14]. With the growth of e-commerce, more consumers are choosing online platforms for their convenience and the ease of comparing products

[12]. This study focuses on identifying the main factors that influence consumer acceptance and use of e-commerce platforms to purchase consumer electrical products.

Although previous studies have extensively investigated consumers' purchase intentions on e-commerce platforms using the TAM, its application in the context of consumer electronics has been limited. Moreover, most prior research has overlooked the influence of logistics service quality (LSQ) in addressing this research question. Considering that logistics plays a crucial role in the transportation of goods between individual customers and businesses, LSQ directly affects consumers' experiences and intentions to buy online [23] [24] [91] [92].

Our findings indicate that consumers' attitudes and perceived usefulness towards online shopping, as well as the informativeness and empathy of LSQ, strongly influence their intention to purchase consumer electrical products online. Additionally, convenience, reliability, and ease of use are identified as contributing factors, although to a lesser extent. Conversely, the low importance of gender, immediacy, and grade suggests their minimal impact on consumers' purchase intention in the context of online purchases for consumer electrical products. Based on these findings, we discuss relevant theoretical and practical implications below.

A. Theoretical Implications

The findings of this study demonstrate that the majority of proposed main effects related to LSQ are relevant to purchase intention, confirming the suitability of incorporating LSQ into the TAM. Among the various drivers influencing purchase intention, attitude emerges as the most influential factor in the intention to purchase consumer electrical products online, as supported by its higher SHAP value (please refer to Fig. 3).

Attitude represents an individual's positive or negative feelings towards performing a specific behavior [68]. It is a psychological process that shapes an individual's preference or aversion towards a particular item [93]. When consumers hold positive attitudes towards a behavior, they are more likely to engage in that behavior [94]. Previous research has consistently reported positive effects of attitude on purchase intention in relation to various products (e.g., [31] [32] [33] [35] [69]).

According to Davis [65], perceived usefulness refers to "the degree to which a person believes that using a particular system would enhance his or her job performance." This factor has also been identified as a key driver of purchase intention in the context of consumer electrical products, aligning with previous studies that employed the TAM (e.g., [29] [32] [33]). However, the influence of perceived usefulness on intention can vary across different contexts. For instance, in the context of online food shopping in Vietnam, perceived usefulness strongly affects attitudes toward online food purchasing, but it does not significantly predict consumers' intention to purchase food products online [36].

In contrast, our study focuses on consumer electrical products, where perceived usefulness directly influences purchase intention. This divergence in findings may be attributed to the distinct characteristics and purposes of the products involved. Unlike electronic products, food items do not require extensive parameter comparisons, and consumers can easily gather relevant information from product packaging in offline purchases, resulting in a faster decision-making process. Moreover, consumers can find various food brands in a single physical store, whereas comparing different electronic brands is often more challenging within a single store. In summary, perceived usefulness plays a significant role in shaping consumers' purchase intention, both directly and indirectly. The higher SHAP values obtained in this experiment further support the importance of perceived usefulness in driving purchase intention.

Informativeness refers to the timeliness of logistics information [43]. Similarly, Oh et al. [42] define information quality as the extent to which overseas direct purchase platforms provide various logistics information, such as product delivery location details, to consumers. In our study, informativeness pertains to the information related to logistics provided by the platform during product delivery. Previous research has consistently highlighted the pivotal role of informativeness in online shopping decisions [95], specifically emphasizing its significant impact on consumers' online purchase behavior within the realm of logistics service quality (e.g., [42] [44]). These studies have predominantly focused on online shopping scenarios, where the merchant delivers the goods to the buyer through logistics after the buyer's payment. During this process, consumers are often concerned about the real-time location of the product, the estimated delivery time, and the convenience of the delivery schedule. Consequently, consumers of electrical products require detailed delivery information that is promptly updated.

Empathy quality encompasses the perspective of a company and its employees in providing personalized services and safeguarding customer safety and rights [45]. In our study, empathy emerges as a moderately influential factor and exerts a significant impact on consumers' purchase intention. This finding reinforces the notion that consumers' perception of a logistic company's understanding and consideration of their needs influences their decision to make an online purchase. Prior research, which regards empathy as one of the dimensions of SERVQUAL within the logistics industry, supports this conclusion [96]. Given that consumer electrical products are susceptible to damage during transportation and may be at risk of loss when stored at a pick-up point, customers expect more intimate and meticulous logistics services to mitigate these hidden risks. Thus, logistics workers need to adopt a customer-centric approach and enhance empathy in transportation and delivery processes.

Convenience and reliability are identified as contributing factors to consumers' purchase intention, albeit to a lesser extent. Informativeness, empathy, convenience, and reliability are all components associated with logistics service quality, indicating that the quality of logistics services significantly impacts consumers' willingness to purchase electrical products online. This perspective is supported by previous studies examining consumer online purchase intention [1] [25] [40] [41] [42] [43] [44] [45]. Since consumers' online purchases of electronic products rely on logistics for delivery, whether through self-operated logistics or third-party logistics, the quality of the delivery service becomes an integral part of consumers' evaluation of their online shopping experience. Considering the significant investment often associated with consumer electrical products, customers seek a reliable delivery service that ensures the protection of their valuable items from damage or loss. Consequently, a high-quality delivery service not only contributes to a satisfactory customer experience but also fosters customer retention [23]. Therefore, our findings suggest that logistics service quality should be considered as a new external variable within the framework of TAM when examining online purchases of consumer electrical products.

On the other hand, the factor of immediacy, another component of logistics service quality, has minimal impact on consumers' purchase intention in the context of online purchases for consumer electrical products. This finding contrasts with other studies that emphasize the importance of timely product receipt (e.g., [44] [45]). The discrepancy can be attributed to the specific research contexts. In the case of fresh food, consumers prioritize the freshness of the products, leading to higher quality requirements and a shorter product shelf life. As a result, consumers place greater emphasis on delivery time, the merchant's delivery capacity, and the merchant's responsiveness to return requests. In contrast, consumer electrical products do not possess stringent "shelf life" conditions, reducing consumers' urgency for timely delivery.

Furthermore, perceived ease of use has a slight effect on consumers' purchase intention online in this study. Typically, when consumers consider buying a product or adopting a technology, they prioritize its usefulness rather than its simplicity and ease of use, as suitability to their needs is paramount. Moreover, previous TAM-related studies have revealed that perceived ease of use indirectly affects purchase intention by influencing attitude [29] [32] [33] [37] or perceived usefulness [31] [33] [69] [94]. In the realm of online e-commerce, consumer behavior indicates a higher inclination towards utilizing an e-commerce platform when they perceive it as a facilitator in locating desired consumer electrical products, conducting price comparisons, and completing purchase transactions. Notwithstanding any potential complexities in platform usability, consumers exhibit a willingness to expend effort if they perceive the platform to be beneficial in addressing their needs. This may explain the limited significance of perceived ease of use for purchase intention in our study.

In conclusion, our study demonstrates that perceived usefulness, informativeness, empathy, convenience, and reliability are significant factors influencing online purchase intentions for consumer electrical products. These findings highlight the crucial role of logistics service quality in shaping consumer behavior and provide valuable insights for both academics and practitioners.

B. Practical Implications

Our findings have significant practical implications as they can provide guidelines for online electronics sellers to improve the services provided to consumers and enhance their willingness to purchase electronics online. Among the drivers of consumers' intention to purchase electronic products online, attitude and perceived usefulness emerge as two key factors. Consumers with a more positive attitude toward buying consumer electronic products (CEP) online exhibit a stronger purchase intention. To cultivate positive attitudes, e-commerce platforms need to rigorously assess the business qualifications of electronic product stores to ensure that consumers can purchase authentic products. Merchants should also exercise careful control over product quality and refrain from delivering poor-quality products to consumers.

Furthermore, consumers are more inclined to purchase CEP online when they perceive it as more useful and efficient compared to offline purchases. To facilitate this perception, platforms should provide comprehensive product parameters and explanations regarding the impact of these parameters on device functionality. Merchants should offer timely and professional pre-sales and after-sales customer service, capable of recommending suitable product models based on consumers' functional needs and budget considerations. User comments and word-of-mouth play a significant role in influencing consumers' online purchasing decisions, representing an advantage of online shopping over offline alternatives. Consumer electronics are particularly sensitive to external word-of-mouth effects, given the limitations on consumers' ability to directly experience the products. Hence, consumers often rely on reviews to avoid making erroneous purchase decisions [14]. Consequently, platforms must combat "fake reviews" to ensure the authenticity and credibility of comments, further enhancing consumers' perceived usefulness of online CEP purchases and their intention to engage in online shopping.

Moreover, e-commerce platforms should strive to improve consumers' perceived ease of use, even though its direct impact on purchase intention may be relatively modest. The individuals' perceived ease of use, as an expression of user experience, can influence their decision to accept a product or platform. This highlights the importance for online store operators and e-commerce platforms to emphasize user experience in the development of effective and user-centered platforms [46] [47]. Online store operators should optimize the display of product information and corresponding keywords to facilitate consumers' search for products aligned with their needs. The platform could incorporate a parameter comparison function module, enabling consumers to compare multiple products selected from different brand stores within the same interface.

Several LSQ factors in our study exhibited high SHAP values for purchase intention, particularly informativeness and empathy. Therefore, third-party logistics enterprises and self-operated logistics involved in fulfilling online orders for electronic products are encouraged to leverage or develop digital technologies and artificial intelligence tools, such as real-time courier positioning, to provide timely updates on delivery information and help customers understand the delivery locations. Logistics service providers should prioritize empathy by employing a larger number of employees assigned to specific clients, facilitating greater individualization and strengthening customer relationships [97]. Furthermore, service providers must adopt a customer-centric perspective and provide friendly services throughout transportation, delivery, and other stages. To enhance reliability, logistics and distribution processes associated with CEP orders should consider appropriate packaging measures to minimize the loss of electronic products during transit [98]. Additionally, proactive communication with buyers prior to delivery, confirming convenient arrangements for personal package receipt, is crucial to ensure reliable delivery and prevent loss of goods. Regarding convenience, optimizing the package layout and pickup process of post stations or express delivery cabinets, and ideally offering door-to-door delivery services, can enhance the convenience of the overall delivery experience.

In sum, this research emphasizes the pivotal roles of consumers' attitudes and perceived usefulness in shaping their intentions to purchase consumer electronic products online. For bolstering positive purchase intentions, e-commerce platforms must ensure product authenticity and maintain rigorous quality standards. Comprehensive product details and robust customer service enhance the perceived usefulness of online shopping. Additionally, specific LSQ factors, particularly informativeness and empathy, play substantial roles in influencing purchase intentions. The emphasis on empathetic and reliable logistics is further reinforced by findings in other studies [97] [98].

VII. Conclusion

Prior studies have given limited attention to examining logistics service quality within the Technology Acceptance Model (TAM) framework, particularly from a demand-side perspective. This research aims to address the gap in the literature by investigating the influence of logistics service quality on consumers' online purchases of consumer electronics products. Our study adopts an integrative approach combining TAM and logistics service quality dimensions to provide a comprehensive assessment of consumer behavior from both informational and transportation perspectives.

This study makes several contributions. First, the study integrates the TAM with logistics service quality. This research bridges the TAM framework with logistics service quality to provide a more comprehensive analysis of consumers' purchase behavior. This study highlights how individuals' acceptance of online platforms shapes their propensity to purchase consumer electronics online, given the convenience and availability of detailed product information and technical support on these channels. By considering both the information perspective and transportation perspective in purchasing consumer electronic products, the research highlights the significance of logistics service quality within the supply chain and its impact on consumers' purchase intention. Second, the research identifies influential determinants. The research identifies attitude, perceived usefulness, and informativeness as the most influential factors affecting consumers' purchase intention. By examining the impact of these factors, the research provides valuable insights into the determinants of consumers' intention to purchase electronic products online. Third, this study underscores the vital role of logistics service quality. The research emphasizes the importance of logistics service quality in the success of online retailers. It highlights the need to consider logistics service quality within the framework of TAM and discusses the challenges related to logistics service quality in the delivery of electrical products and its impact on consumers' purchase intentions. Fourth, the study employs both parametric and non-parametric analytical techniques to facilitate robust analysis of the research model. The research adopts the SHAP machine learning algorithm to analyze and interpret the impact of each variable on consumers' purchase intention. This approach enhances the transparency and interpretability of machine learning models in predicting and explaining consumers' purchase behavior. Lastly, the research offers valuable insights for both academia and practitioners by providing managerial recommendations. These recommendations enable online retailers to maintain a competitive advantage, promote successful online consumer behavior, and foster profitability in the e-commerce sector. In summary, this is the first study to synthesize the TAM and logistics service quality factors to examine consumer electronics purchase decisions in online retail, utilizing a mixed methods approach. The integration of technology acceptance and supply chain considerations provides a holistic perspective on consumer decision making, while the empirical findings provide actionable guidelines for practitioners. This research advances the field by providing a comprehensive analysis of the factors influencing consumers' purchase behavior in the context of online e-commerce platforms for consumer electronic products.

Although this study has several contributions, it also presents limitations and future work. Firstly, the generalizability of our findings to other shopping contexts may be limited as our research is specifically tailored to the online purchase of consumer electronics. To address this limitation, future studies can employ our research framework to examine the dynamics of online shopping in relation to different types of products. Secondly, while this study examines consumers' online purchase intention of consumer electrical products based on the TAM and incorporates LSQ as an external variable, we exclusively focus on the main effects proposed by TAM and exclude the examination of moderators. To enrich our understanding, future studies should explore the potential moderating effects on the proposed relationships. For instance, Wong et al. [35] discovered that perceived risk mediates the relationship between perceived usefulness and purchase intention. Conversely, Jain [32] found no moderating role of perceived risk in the relationship between attitude and intention, but identified web atmospherics as a moderator between attitude toward online shopping and online purchase intention. Future research would investigate further into complex interrelations among these factors and introduce other variables such as virtual reality shopping experiences or the role of social media influencers, thereby deepening the insight into consumer behavior concerning online purchases of electronics.

References

[1] T. Gao, "Study on the intention of foreign trade driven by cross-border E-commerce based on blockchain technology," *Security and Communication Networks*, pp.1-10, 2021.

[2] Y. Qi, X. Wang, M. Zhang, Q. Wang, "Developing supply chain resilience through integration: An empirical study on an e-commerce platform," *Journal of Operations Management*, vol. 69, no. 3, pp. 477-496, 2023.

[3] I. Masudin, E. Lau, N. T. Safitri, D. P. Restuputri, D. I. Handayani, "The impact of the traceability of the information systems on humanitarian logistics performance: Case study of Indonesian relief logistics services," *Cogent Business & Management*, vol. 8, no. 1, p. 1906052,2021.

[4] P. Kgobe, P. Ozor, "Integration of radio frequency identification technology in supply chain management: A critical review," *Operations and Supply Chain Management: An International Journal*, vol. 14, no. 3, pp. 289-300, 2021.

[5] J. Zhan, S. Dong, W. Hu, "IoE-supported smart logistics network communication with optimization and security," *Sustainable Energy Technologies and Assessments*, vol. 52, p. 102052, 2022.

[6] W. Wu, L. Shen, Z. Zhao, A. R. Harish, R. Y. Zhong, G. Q. Huang, "Internet of Everything and Digital Twin Enabled Service Platform for Cold Chain Logistics," *Journal of Industrial Information Integration*, vol. 33, p. 100443, 2023.

[7] M. Nilashi, A. M. Baabdullah, R. A. Abumalloh, K. B. Ooi, G. W. H. Tan, M. Giannakis, Y. K. Dwivedi, "How can big data and predictive analytics impact the performance and competitive advantage of the food waste and recycling industry?" *Annals of Operations Research*, pp. 1-42, 2023.

[8] M. Cotarelo, H. Calderón, T. Fayos, "A further approach in omnichannel LSQ, satisfaction and customer loyalty," *International Journal of Retail & Distribution Management,* vol. 49, no. 8, pp. 1133-1153, 2021.

[9] P. Raman, "Understanding female consumers' intention to shop online: The role of trust, convenience and customer service," *Asia Pacific Journal of Marketing and Logistics*, vol. 31, no. 4, pp. 1138-1160, 2019.

[10] M. Ramkumar, T. Schoenherr, S. M. Wagner, M. Jenamani, "Q-TAM: A quality technology acceptance model for predicting organizational buyers' continuance intentions for e-procurement services," *International Journal of Production Economics*, vol. 216, pp. 333-348, 2019.

[11] Q. Chen, N. Zhang, "Does E-Commerce Provide a Sustained Competitive Advantage? An Investigation of Survival and Sustainability in Growth-Oriented Enterprises," *Sustainability*, vol. 7, no. 2, pp. 1411–1428, 2015.

[12] Q. T. Pham, X. P. Tran, S. Misra, R. Maskeliūnas, R. Damaševičius, "Relationship between convenience, perceived value, and repurchase intention in online shopping in Vietnam," *Sustainability*, vol. 10, no. 1, p. 156, 2018.

[13] N. W. Masri, A. Ruangkanjanases, S. C. Chen, "The effects of product monetary value, product evaluation cost, and customer enjoyment on customer intention to purchase and reuse vendors: institutional trust-based mechanisms," *Sustainability*, vol. 13, no. 1, p. 172, 2020.

[14] R. Barbado, O. Araque, C. A. Iglesias, "A framework for fake review detection in online consumer electronics retailers," *Information Processing & Management*, vol. 56, no. 4, pp. 1234-1244, 2019.

[15] U. Chakraborty, S. Bhat, "The effects of credible online reviews on brand equity dimensions and its consequence on consumer behavior," *Journal of promotion management*, vol. 24, no. 1, pp. 57-82, 2018.

[16] P. Duarte, S. C. e Silva, M. B. Ferreira, "How convenient is it? Delivering online shopping convenience to enhance customer satisfaction and encourage e-WOM," *Journal of Retailing and Consumer Services*, vol. 44, pp. 161-169, 2018.

[17] M. W. A. Ghouri, L. Tong, M. A. Hussain, "Does Online Ratings Matter? An Integrated Framework to Explain Gratifications Needed for Continuance Shopping Intention in Pakistan," *Sustainability*, vol. 13, no. 17, p. 9538, 2021.

[18] U. Tandon, R. Kiran, A. N. Sah, "The influence of website functionality, drivers and perceived risk on customer satisfaction in online shopping: an emerging economy case," *Information Systems and e-Business Management*, vol. 16, pp. 57-91, 2018.

[19] D. A. Aaker, "Managing brand equity," *New York: Simon and Schuster,* 2009.

[20] Q. C. He, Y. J. Chen, "Dynamic pricing of electronic products with consumer reviews," *Omega*, vol. 80, pp. 123-134, 2018.

[21] P. Utomo, T. F. SUTRISNO, "What Influences the Online Purchasing of Electrical Equipment Products?," *KnE Social Sciences*, pp. 143-156, 2021.

[22] X. Qin, Q. Su, S. H. Huang, U. J. Wiersma, M. Liu, "Service quality coordination contracts for online shopping service supply chain with competing service providers: integrating fairness and individual rationality," *Operational Research*, vol. 19, pp. 269-296, 2017.

[23] R. Cui, M. Li, Q. Li, "Value of High-Quality Logistics: Evidence from a

Clash Between SF Express and Alibaba," *Management Science*, vol. 66, no. 9, pp. 3879–3902, 2020. doi:10.1287/mnsc.2019.3411

[24] D. Li, Y. Liu, C. Fan, J. Hu, X. Chen, "Logistics service strategies under different selling modes," *Computers & Industrial Engineering*, vol. 162, p. 107684, 2021.

[25] B. Zheng, H. Wang, A. M. Golmohammadi, A. Goli, "Impacts of logistics service quality and energy service of Business to Consumer (B2C) online retailing on customer loyalty in a circular economy," *Sustainable Energy Technologies and Assessments*, vol. 52, p. 102333, 2022.

[26] S. M. Lundberg, S. I. Lee, "A unified approach to interpreting model predictions," *Advances in neural information processing systems*, vol. 30, 2017.

[27] S. B. Jabeur, H. Ballouk, W. B. Arfi, R. Khalfaoui, "Machine learning-based modeling of the environmental degradation, institutional quality, and economic growth," *Environmental Modeling & Assessment*, pp. 1-14, 2021.

[28] L. Antwarg, R. M. Miller, B. Shapira, L. Rokach, "Explaining anomalies detected by autoencoders using Shapley Additive Explanations," *Expert systems with applications*, vol. 186, p. 115736, 2021.

[29] W. Wang, S. Wang, "Toward parent-child smart clothing: Purchase intention and design elements," *Journal of Engineered Fibers and Fabrics*, vol. 16, p. 1558925021991843, 2021.

[30] V. Chidambaram, N. P. Rana, S. Parayitam, "Antecedents of consumers' online apparel purchase intention through Virtual Try On technology: A moderated moderated-mediation model," *Journal of Consumer Behaviour*, vol. 23, no. 1, pp. 107-125, 2024.

[31] T. Zhang, W. Y. C. Wang, L. Cao, Y. Wang, "The role of virtual try-on technology in online purchase decision from consumers' aspect," *Internet Research*, vol. 29, no. 3, pp. 529-551, 2019.

[32] S. Jain, "Examining the moderating role of perceived risk and web atmospherics in online luxury purchase intention," *Journal of Fashion Marketing and Management: An International Journal*, vol. 25, no. 4, pp. 585-605, 2021.

[33] A. Xu, W. Li, Z. Chen, S. Zeng, L. A. Carlos, Y. Zhu, "A Study of Young Chinese Intentions to Purchase "Online Paid Knowledge": An Extended Technological Acceptance Model," *Frontiers in Psychology*, vol. 12, p. 695600, 2021.

[34] C. K. H. Lee, A. O. M. Wong, "Antecedents of consumer loyalty in ride-hailing," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 80, pp. 14-33, 2021.

[35] D. Wong, H. Liu, Y. Meng-Lewis, Y. Sun, Y. Zhang, "Gamified money: exploring the effectiveness of gamification in mobile payment adoption among the silver generation in China," *Information Technology & People*, vol. 35, no. 1, pp. 281-315, 2022.

[36] T. M. A. Nguyen, T. H. Nguyen, H. H. Le, "Online Shopping in Relationship with Perception, Attitude, and Subjective Norm during COVID-19 Outbreak: The Case of Vietnam," *Sustainability*, vol. 14, no. 22, p. 15009, 2022.

[37] A. Vafaei-Zadeh, T. K. Wong, H. Hanifah, A. P. Teoh, K. Nawaser, "Modelling electric vehicle purchase intention among generation Y consumers in Malaysia," *Research in Transportation Business & Management*, vol. 43, p. 100784, 2022.

[38] L. Wang, Z. Wang, X. Wang, Y. Zhao, "Explaining consumer implementation intentions in mobile shopping with SEM and fsQCA: Roles of visual and technical perceptions," *Electronic Commerce Research and Applications*, vol. 49, p. 101080, 2021.

[39] M. A. Rahaman, H. K. Hassan, A. A. Asheq, K. A. Islam, "The interplay between eWOM information and purchase intention on social media: Through the lens of IAM and TAM theory," *PloS one*, vol. 17, no. 9, p. e0272926, 2022.

[40] Y. M. Cang, D. C. Wang, "A comparative study on the online shopping willingness of fresh agricultural products between experienced consumers and potential consumers," *Sustainable Computing: Informatics and Systems*, vol. 30, p. 100493, 2021.

[41] A. H. Ali, T. Gruchmann, A. Melkonyan, "Assessing the impact of sustainable logistics service quality on relationship quality: Survey-based evidence in Egypt," *Cleaner Logistics and Supply Chain*, vol. 4, p. 100036, 2022.

[42] K. Y. Oh, S. Y. Kang, Y. G. Oh, "The Moderating Effects of Eco-Friendliness between Logistics Service Quality and Customer Satisfaction in Cross-Border e-Commerce: Evidence from Overseas Direct Purchasers in Korea," *Sustainability*, vol. 14, no. 22, p. 15084, 2022.

[43] Z. Dong, "Construction of mobile E-commerce platform and analysis of its impact on E-commerce logistics customer satisfaction," *Complexity*, vol. 2021, pp. 1-13, 2021.

[44] D. Choi, C. Y. Chung, J. Young, "Sustainable online shopping logistics for customer satisfaction and repeat purchasing behavior: Evidence from China," *Sustainability*, vol. 11, no. 20, p. 5626, 2019.

[45] Y. Jiang, P. Lai, C. H. Chang, K. F. Yuen, S. Li, X. Wang, "Sustainable management for fresh food E-commerce logistics services," *Sustainability*, vol. 13, no. 6, p. 3456, 2021.

[46] D. Winter, C. Hausmann, A. Hinderks, J. Thomaschewski, "Development of a shared UX vision based on ux factors ascertained through attribution," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 8, no. 2, pp. 247-254, 2023.

[47] M. Schrepp, J. Kollmorgen, A. L. Meiners, A. Hinderks, D. Winter, H. B. Santoso, J. Thomaschewski, "On the importance of UX quality aspects for different product categories," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 8, no. 2, pp. 232-246, 2023.

[48] C. W. Lu, G. H. Lin, T. J. Wu, I. H. Hu, Y. C. Chang, "Influencing factors of cross-border e-commerce consumer purchase intention based on wireless network and machine learning," *Security and Communication Networks*, vol. 2021, no. 1, pp. 1-9, 2021.

[49] J. Jia, "Analysis of alternative fuel vehicle (AFV) adoption utilizing different machine learning methods: a case study of 2017 NHTS," *IEEE Access*, vol. 7, pp. 112726-112735, 2019.

[50] J. Jia, B. Shi, F. Che, H. Zhang, "Predicting the regional adoption of electric vehicle (EV) with comprehensive models," *IEEE Access*, vol. 8, pp. 147275-147285, 2020.

[51] T. Shu, Z. Wang, L. Lin, H. Jia, J. Zhou, "Customer perceived risk measurement with NLP method in electric vehicles consumption market: empirical study from China," *Energies*, vol. 8, no. 5, p. 1637, 2022.

[52] K. Sobiech-Grabka, A. Stankowska, K. Jerzak, "Determinants of electric cars purchase intention in Poland: personal attitudes v. economic arguments," *Energies*, vol. 15, no.9, p.3078, 2022.

[53] F. Li, S. Katsumata, C. H. Lee, Q. Ye, W. D. Dahana, R. Tu, X. Li, "Autoencoder-enabled potential buyer identification and purchase intention model of vacation homes," *IEEE Access*, vol. 8, pp. 212383-212395, 2020.

[54] R. D. Borres, A. K. S. Ong, T. W. O. Arceno, A. R. Padagdag, W. R. L. B. Sarsagat, H. R. M. S. Zuñiga, J. D. German, "Analysis of Factors Affecting Purchase of Self-Defense Tools among Women: A Machine Learning Ensemble Approach," *Applied Sciences*, vol. 13, no. 5, pp. 3003, 2023.

[55] F. Taghikhah, A. Voinov, N. Shukla, T. Filatova, "Shifts in consumer behavior towards organic products: Theory-driven data analytics," *Journal of Retailing and Consumer Services*, vol. 61, p. 102516, 2021.

[56] X. Liu, X. Huang, J. Zhao, Y. Su, L. Shen, Y. Duan,... J. Guo,"Application of machine learning in Chinese medicine differentiation of dampness-heat pattern in patients with type 2 diabetes mellitus," *Heliyon*, vol. 9, no.2, 2023.

[57] H. Sahlaoui, A. Nayyar, S. Agoujil, M. M. Jaber, "Predicting and interpreting student performance using ensemble models and shapley additive explanations," *IEEE Access*, vol. 9, pp. 152688-152703, 2021.

[58] P. L. Ballester, T. D. A. Cardoso, F. P. Moreira, R. A. da Silva, T. C. Mondin, R. M. Araujo,... L. D. de Mattos Souza, "5-year incidence of suicide-risk in youth: A gradient tree boosting and SHAP study," *Journal of affective disorders*, vol. 295, pp. 1049-1056, 2021.

[59] A. A. Huang, S. Y. Huang, "Use of machine learning to identify risk factors for insomnia," *Plos one*, vol. 18, no. 4, p. e0282622, 2023.

[60] R. Fan, T. Hua, T. Shen, Z. Jiao, Q. Yue, B. Chen, Z. Xu, "Identifying patients with major depressive disorder based on tryptophan hydroxylase-2 methylation using machine learning algorithms," *Psychiatry Research*, vol. 306, p. 114258, 2021.

[61] H. Fu, "Factors influencing user usage intention on intelligent logistics information platform," *Journal of Intelligent & Fuzzy Systems*, vol. 35, no. 3, pp. 2711-2720, 2018.

[62] D. Edwards, N. Subramanian, A. Chaudhuri, P. Morlacchi, W. Zeng, "Use of delivery drones for humanitarian operations: analysis of adoption barriers among logistics service providers from the technology acceptance model perspective," *Annals of Operations Research*, pp. 1-23, 2023.

[63] C. C. Bienstock, M. B. Royne, D. Sherrell, T. F. Stafford, "An expanded model of logistics service quality: Incorporating logistics information technology," *International Journal of Production Economics*, vol. 113, no. 1, pp. 205-222, 2008.

[64] N. K. Jain, D. Kaul, P. Sanyal, "What drives customers towards mobile shopping? An integrative technology continuance theory perspective," *Asia Pacific Journal of Marketing and Logistics*, vol. 34, no. 5, pp. 922-943, 2022.

[65] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quarterly*, vol. 13, no. 3, pp. 319-339, 1989, doi: 10.2307/249008.

[66] Y. Liang, S. H. Lee, J. E. Workman, "Implementation of artificial intelligence in fashion: Are consumers ready?," *Clothing and Textiles Research Journal*, vol. 38, no. 1, pp. 3-18, 2020.

[67] X. Wang, W. Zhang, T. Zhang, Y. Wang, S. Na, "A study of Chinese consumers' consistent use of mobile food ordering apps," *Sustainability*, vol. 14, no. 19, p. 12589, 2022.

[68] M. Fishbein, I. Ajzen, "Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research," *Addison-Wesley*, Reading, MA, 1975.

[69] M. F. Mazzù, A. Baccelloni, S. Romani, A. Andria, "The role of trust and algorithms in consumers' front-of-pack labels acceptance: a cross-country investigation," *European Journal of Marketing*, (ahead-of-print), 2022.

[70] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[71] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, A. Gulin, "CatBoost: unbiased boosting with categorical features," arXiv preprint arXiv:1706.09516, 2017.

[72] T. Chen, C. Guestrin, "Xgboost: A scalable tree boosting system," Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, pp. 785–794, 2016.

[73] A. Dasgupta, Y. V. Sun, I. R. König, J. E. Bailey-Wilson, J. D. Malley, "Brief review of regression-based and machine learning methods in genetic epidemiology: the Genetic Analysis Workshop 17 experience," *Genetic epidemiology*, vol. 35, no. S1, pp. S5-S11, 2011.

[74] A. V. Dorogush, V. Ershov, A. Gulin, "CatBoost: gradient boosting with categorical features support," arXiv preprint arXiv:1810.11363, 2018.

[75] A. Jamal, M. Zahid, M. Tauhidur Rahman, H. M. Al-Ahmadi, M. Almoshaogeh, D. Farooq, M. Ahmad, "Injury severity prediction of traffic crashes with ensemble machine learning techniques: a comparative study," *International journal of injury control and safety promotion*, pp. 1–20, 2021.

[76] W. Liang, S. Luo, G. Zhao, H. Wu, "Predicting Hard Rock Pillar Stability Using GBDT, XGBoost, and LightGBM Algorithms," *Mathematics*, vol. 8, no. 5, p. 765, 2020.

[77] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T. Y. Liu, "Lightgbm: a highly efficient gradient boosting decision tree," *Adv. Neural Information Processing Systems,* vol. 30, pp. 3146–3154, 2017.

[78] P. Pornaroontham, K. Kim, S. Kulprathipanja, P. Rangsunvigit, "Water-soluble organic former selection for methane hydrates by supervised machine learning," *Energy Reports,* vol. 9, pp. 2935-2946, 2023.

[79] L. S. Shapley, "A value for n-person games," *Contribution to the Theory of games II*,vol. 28, pp. 307–317, 1953.

[80] C. Yang, M. Chen, Q. Yuan, "The application of XGBoost and SHAP to examining the factors in freight truck-related crashes: An exploratory analysis," *Accident Analysis & Prevention,* vol. 158, p. 106153, 2021.

[81] X. Jiang, H. Wang, X. Guo, X. Gong, "Using the FAHP, ISM, and MICMAC approaches to study the sustainability influencing factors of the last mile delivery of rural E-commerce logistics," *Sustainability*, vol. 11, no. 14, p. 3937, 2019.

[82] Ž. Stević, I. Tanackov, A. Puška, G. Jovanov, J. Vasiljević, D. Lojaničić, "Development of modified SERVQUAL–MCDM model for quality determination in reverse logistics," *Sustainability*, vol. 13, no. 10, p. 5734, 2021.

[83] L. A. Luyen, N. V. Thanh, "Logistics service provider evaluation and selection: Hybrid servqual–fahp–topsis model," *Processes*, vol. 10, no. 5, p. 1024, 2022.

[84] Y. K. Huang, Y. W. Kuo, S. W. Xu, "Applying Importance-performance Analysis to Evaluate Logistics Service Quality for Online Shopping among Retailing Delivery," *International Journal of Electronic Business Management*, vol. 7, no. 2, 2009.

[85] C. Fornell, D. F Larcker, "Evaluating structural equation models with unobservable variables and measurement error", *Journal of marketing research*, vol. 18, no. 1, pp. 39-50, 1981.

[86] T. H. Tsai, H. T. Chang, Y. C. Chang, Y. S. Chang, "Personality disclosure on social network sites: An empirical examination of differences in Facebook usage behavior, profile contents and privacy settings," *Computers in Human Behavior*, vol. 76, pp. 469-482, 2017.

[87] J. F. Hair, W. C. Black, B. J. Babin, R. E. Anderson, "Multivariate Data Analysis," *Pearson*, London, 2010.

[88] N. V. Chawla, K. W. Bowyer, L. O. Hall, W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of artificial intelligence research*, vol. 16, pp. 321-357, 2002.

[89] M. M. Liu, L. Wen, Y. J. Liu, Q. Cai, L. T. Li, Y. M. Cai, "Application of data mining methods to improve screening for the risk of early gastric cancer," *BMC medical informatics and decision making*, vol. 18, no. 5, pp. 23-32, 2018.

[90] L. Yao, Z. Xu, X. Zhao, Y. Chen, L. Liu, X. Fu, F. Chen, "Therapists and psychotherapy side effects in China: A machine learning-based study," *Heliyon*, vol. 8, no. 11, p. e11821, 2022.

[91] R. Hammami, Y. Frein, A.S. Albana, "Delivery time quotation and pricing in two-stage supply chains: Centralized decision-making with global and local managerial approaches," *European Journal of Operational Research*, vol. 286, no. 1, pp. 164–177, 2020.

[92] P. He, S. Zhang, C. He, "Impacts of logistics resource sharing on B2C E-commerce companies and customers,"*Electronic Commerce Research and Applications*, vol. 34, p. 100820, 2019.

[93] N. Sreen, S. Purbey, P. Sadarangani, "Impact of culture, behavior and gender on green purchase intention,"*Journal of Retailing and Consumer Services*, vol. 41, pp. 177–189, 2018.

[94] C. Hwang, T. L. Chung, E. A. Sanders, "Attitudes and purchase intentions for smart clothing: Examining US consumers' functional, expressive, and aesthetic needs for solar-powered clothing," *Clothing and Textiles Research Journal*, vol. 34, no. 3, pp. 207-222, 2016.

[95] D. Dwidienawati, D. Tjahjana, S.B. Abdinagoro, D. Gandasari, "Customer review or influencer endorsement: which one influences purchase intention more?," *Heliyon*, vol. 6, no. 11, p. e05543, 2020.

[96] J. D. German, A. K. S. Ong, A. A. N. P. Redi, K. P. E. Robas, "Predicting factors affecting the intention to use a 3PL during the COVID-19 pandemic: A machine learning ensemble approach," *Heliyon*, vol. 8, no. 11, p. e11382, 2022.

[97] K. Knop, "Evaluation of quality of services provided by transport & logistics operator from pharmaceutical industry for improvement purposes," *Transportation Research Procedia*, vol. 40, pp. 1080-1087, 2019.

[98] R. Tian, Y. Tang, "Multiobjective Planning for Logistics Distribution of Consumer Electronic Items Based on Improved Genetic Algorithm," *Wireless Communications & Mobile Computing (Online)*, vol. 2022, 2022.

[99] J. Lee, E. Suh, H. Park, S. Lee, "Determinants of users' intention to purchase probability-based items in mobile social network games: A case of South Korea," IEEE Access, vol. 6, pp. 12425-12437, 2018.

[100] M. Hu, F. Huang, H. Hou, Y. Chen, L. Bulysheva, "Customized logistics service and online shoppers' satisfaction: an empirical study," *Internet Research*, 2016.

[101] P. Christidis, C. Focas, "Factors affecting the uptake of hybrid and electric vehicles in the European Union," *Energies*, vol. 12, no. 18, p. 3414, 2019.

[102] E. C. A. Carreón, H. Nonaka, A. Hentona, H. Yamashiro, "Measuring the influence of mere exposure effect of TV commercial adverts on purchase behavior based on machine learning prediction models," *Information Processing & Management*, vol. 56, no. 4, pp. 1339-1355, 2019.

[103] S. Ghorbany, E. Noorzai, S. Yousefi, "BIM-Based Solution to Enhance the Performance of Public-Private Partnership Construction Projects using Copula Bayesian Network," *Expert Systems with Applications*, p. 119501, 2023.

[104] P. N. Ramkumar, J. M. Karnuta, H. S. Haeberle, K. A. Owusu-Akyaw, T. S. Warner, S. A. Rodeo,... R. J. Williams III, "Association between preoperative mental health and clinically meaningful outcomes after osteochondral allograft for cartilage defects of the knee: a machine learning analysis," *The American Journal of Sports Medicine*, vol. 49, no. 4, pp. 948-957, 2021.

[105] P. Wang, Z. Xu, "A novel consumer purchase behavior recognition method using ensemble learning algorithm," *Mathematical Problems in Engineering*, vol. 2020, pp. 1-10, 2020.

Cheng-Feng Wu

He received the Ph.D. degree from National Taiwan University, Taipei, Taiwan in 2015. He is currently an Assistant Professor in the Department of Business Administration at National Taipei University, Taipei, Taiwan. His research includes supply chain management, applied econometric time series analysis, data mining, and patient safety culture.

Kunkun Zhang

Student of the graduate degree in Accounting at Hubei University of Economics in China. Her research interest is the application of machine learning in financial and economic management problems.

Meng-Chen Lin

Dr. Lin is currently an Assistant Professor in the Department of Marketing and Distribution Management at Chien Hsin University of Science and Technology, Taoyuan, Taiwan. Her research includes marketing, consumer behavior, and organization management.

Chei-Chang Chiou

Chei-Chang Chiou completed his Master of Business Education at National Changhua University of Education, and Ph.D. in Accounting at National Chengchi University in Taiwan. Currently, he is a Professor of Department of Accounting at National Changhua University of Education in Taiwan. His research interest includes capital market research, ESG, artificial intelligence, educational technology, multimedia learning, concept map, and operating efficiency and productivity.

Constructing the Public Opinion Crisis Prediction Model Using CNN and LSTM Techniques Based on Social Network Mining

Yan Lou[1], Zhipeng Ren[2,3], Yong Zhang[4]*, Zhonghui Tao[1], Yiwu Zhao[1]

[1] National & Local Joint Engineering Research center of Space Opto-Electronic technology, Changchun University of Science and Technology, Changchun, 130022 (China)
[2] School of Computer Science and Technology, Changchun University of Science and Technology, Changchun, 130022 (China)
[3] State Key Laboratory of Applied Optics, Changchun Institute of Optics, Fine Mechanics and Physics, CAS, Changchun 130033 (China)
[4] School of Electrical Information, Changchun Guanghua University, Changchun, 130033 (China)

* Corresponding author: zhangy_19730626@163.com

Received 25 April 2023 | Accepted 4 April 2024 | Early Access 24 July 2024

Abstract

This research endeavors to address the persistent dissemination of public opinion within social networks, mitigate the propagation of inappropriate content on these platforms, and enhance the overall service quality of social networks. To achieve these objectives, Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) techniques are employed in this research to develop a predictive model for anticipating public opinion crises in social network mining. This model furnishes users with a valuable reference for subsequent decision-making processes. The initial phase of this research involves the collection of user behavior data from social networks using IoT technologies, serving as the basis for extensive big data analysis and neural network research. Subsequently, a social network text categorization model is constructed by amalgamating the Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) architecture, elucidating the training procedures of deep learning methodologies within CNN and LSTM networks. The effectiveness of this approach is subsequently validated through comparisons with other deep learning techniques. Based on the obtained results and findings, the CNN-LSTM model demonstrates a noteworthy accuracy rate of 92.19% and an exceptionally low loss value of 0.4075. Of particular significance is the classification accuracy of the CNN-LSTM algorithm within social network datasets, which surpasses that of alternative algorithms, including CNN (by 6.31%), LSTM (by 4.43%), RNN (by 3.51%), Transformer (by 40.29%), and Generative Adversarial Network (GAN) (by 4.49%). This underscores the effectiveness of the CNN-LSTM algorithm in the realm of social network text classification.

Keywords

Convolutional Neural Network, Deep Learning, Inappropriate Remarks, Internet of Things, Long Short-Term Memory, Social Network.

DOI: 10.9781/ijimai.2024.07.005

I. Introduction

The market for intelligent hardware in China has witnessed a rapid expansion. The infusion of intelligence into hardware components has enabled seamless connectivity, fostering the realization of Internet service integration. This has given rise to a distinctive architectural paradigm known as 'cloud + terminal', which effectively harnesses the potential of big data [1] ,[2]. Foundational elements of this platform, including software and hardware components, are characterized by emerging technologies such as intelligent sensor interconnectivity, human-computer interaction, innovative display technologies, and advanced capabilities for processing large-scale data. In their pursuit of cutting-edge designs, materials, and hardware processes, social networks are increasingly focusing on developing new intelligent terminal products and service network applications. This is accomplished by seamlessly integrating intelligent hardware via applications [3]. The field of social network analysis, utilizing deep learning and the Internet of Things (IoT), has emerged as an increasingly significant area of research. The development of IoT has led to enhanced production efficiency due to the evolution of information networks. However, a substantial opportunity may be overlooked if one merely considers the value of social networks from a communication standpoint. Social networks exert a deeper influence on understanding and engagement than traditional social networks, significantly impacting various facets

Please cite this article as:
Y. Lou, Z. Ren, Y. Zhang, Z. Tao, Y. Zhao. Constructing the Public Opinion Crisis Prediction Model Using CNN and LSTM Techniques Based on Social Network Mining, International Journal of Interactive Multimedia and Artificial Intelligence, vol. 8, no. 7, pp. 86-96, 2024, http://dx.doi.org/10.9781/ijimai.2024.07.005

of human civilization to a greater extent than their conventional counterparts [4]. The IoT is well-positioned to effectively monitor interactive data within social networks, particularly given individuals' abundant sharing of interactive data. When intervention is warranted, IoT technology can support users in optimizing social networking applications by offering insights into usage patterns of application software [5]. Within the domain of deep learning, developers have the capacity to construct intricate neural network models capable of identifying and evaluating extensive information from social media. This is accomplished by utilizing expansive and reliable datasets [6]-[8], enhancing the potential for precise analysis and comprehension of social network data.

The primary objective of this research is to address the persistent challenges associated with the propagation of public opinions and inappropriate content within social networks. Additionally, it aims to enhance the overall service quality of social networks. To achieve these goals, this paper leverages deep learning and IoT technologies to intelligently identify and mitigate inappropriate comments within social networks. Initially, IoT technology is employed to amass user behavior data from social networks, laying the foundation for subsequent big data analysis and neural network research. Subsequently, the training procedures of deep learning techniques, specifically Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM), are elucidated. A social network text classification model that integrates CNN and LSTM methodologies is established. Finally, a comparative analysis of the performance of various deep learning techniques is conducted to validate the effectiveness of the proposed algorithm.

The paper is structured into five sections. Section 1 serves as the introduction, providing insights into the research background of social network analysis based on deep learning and IoT. Furthermore, it outlines the research methodology and paper's structure. Section 2 comprises a literature review, presenting theoretical research pertaining to IoT applications and exploring the application of deep learning within social networks. It also expounds on the process of integrating IoT with social networks. Section 3, the model section, introduces the CNN and LSTM algorithms and details the establishment of a social network text classification model that amalgamates both techniques. Section 4 encompasses the results and discussion, wherein the paper elucidates the dataset used for experimentation and evaluates the performance of the social network text classification model that integrates CNN and LSTM. Finally, Section 5 concludes the paper, providing a comprehensive summary of the research conducted herein. It further discusses the achieved results and encountered limitations while outlining prospects for future research endeavors.

II. Literature Review

Al-Garadi et al. conducted a research study employing the social network analysis methodology to investigate variations in communicative language within the context of theoretical research concerning the IoT in the realm of social networks. They established a correlation between the intensity of communicative language and the distribution of languages, positing that the emergence of social networks has transformed the Internet into a social force that exerts a significant influence on interpersonal relationships [9]. Khalil et al. contended that online communication media share a common attribute in reinforcing existing social models while expanding the reach of social networks. These media enable users to engage in active communication for predefined periods without supplanting other communication tools [10]. Additionally, Andronie et al. argued that emerging technologies had heightened the prevalence of networked social behaviors [11]. With the aid of these novel technologies, individuals are redefining the mode of network interaction, giving rise to a new form of network society. Deep learning, as an influential approach, has proven effective in addressing a multitude of intricate challenges and has demonstrated remarkable efficacy in diverse domains, including object detection, speech recognition, and language translation. Rahman et al. applied machine learning techniques to analyze the emotions conveyed in social network texts. They employed a noise reduction autoencoder for text feature extraction and emotion classification, with experimental results highlighting the proficiency of machine learning in emotional data analysis [12].

The advent of novel technologies has afforded individuals the opportunity to reshape the social structure and interaction modalities within networks. Deep learning has emerged as a potent tool for tackling complex challenges, enabling target detection, speech recognition, and language translation. In the context of social network alignment, Vinayakumar et al. delved into the alignment processes utilized by dynamic social network users and proposed a dynamic social network model based on the depth sequence model. Their extensive experiments, conducted with real data, revealed a substantial 10% enhancement in the alignment effect by utilizing the dynamic social network model [13]. Humayun et al. aimed to generate user-matching scores through deep learning, employing CNN to map features derived from grid data logged by users in social networks. Their findings affirmed the heightened matching accuracy achieved through this approach [14]. Lv et al. conducted research focused on the extraction of public opinion information from social networks using deep learning-based techniques [15]. Khan and colleagues investigated the application of deep learning in social network surveillance and developed social computing to enhance the safety of social network environments [16]. Siboni et al. integrated deep self-coding and network representation learning while collecting implicit semantic data from the network to construct a social network information recommendation model [17]. Experimental results indicated that implicit social network information more accurately identifies user connections compared to explicit semantic information. Sarker explored the capacity of CNNs and LSTM networks to categorize public opinions and emotions within social networks, leveraging deep learning techniques for public opinion classification within these networks [18]. In another study, Yu et al. employed a web crawler to collect login information from Facebook and Twitter users, pre-processed the experimental data, and applied a deep learning algorithm for identity recognition. Research findings demonstrated that deep learning surpasses conventional identification algorithms, highlighting its potential benefits within the realm of social networks [19]. Furthermore, text mining technology can analyze textual content, identify user behaviors and characteristics, investigate social media connections, and create user profiles, emphasizing its relevance and utility within the domain of social network analysis.

Social media data has been harnessed for various urban planning and tourism analysis applications. Muñoz et al. (2022) employed social media data to extract insights into tourist characteristics, laying the groundwork for urban planning and tourism analysis [20]. Leveraging the vast reservoir of user-generated content, social media data provides valuable glimpses into tourists' preferences, interests, and behaviors. This resource facilitates an enhanced comprehension of tourist behavior patterns and demands, thereby enabling the development of more precisely targeted strategies for urban planning and tourism decision-making. In the domain of hotel reviews, the utilization of attention-based emotion prediction models has exhibited promise. Arroni et al. (2023) implemented an attention-based model to forecast sentiment in tweets pertaining to Las Vegas hotels. The study reported a similarity score of 0.64121 with actual hotel rankings, underscoring the model's efficacy in sentiment analysis [21]. This approach can aid hotel managers in gaining insights into customer sentiment feedback, swiftly identifying, and addressing issues, and ultimately elevating service quality.

In summary, notwithstanding the substantial theoretical and practical strides made by researchers in China and other nations, shallow learning methodologies continue to predominate in natural language processing research and applications. The integration of CNN or deep learning techniques for text training and classification remains relatively limited. Embracing the realm of deep learning and IoT-based social network analysis, however, offers the potential for a more comprehensive and reliable information source, serving as a valuable asset to inform subsequent decision-making processes for users. While extant research has predominantly focused on the exploration of changes in communication language, particularly in the context of language variation in communication, it has not delved deeply into social media discourse or discourse characteristics linked to public opinion crises. These research perspectives may fall short in providing adequate insights and solutions to tackle the challenges posed by public opinion crises. These identified limitations and gaps serve as the impetus and focal point of the present study. The objective of this paper is to develop a predictive model capable of addressing public opinion crises within social networks. The paper harnesses CNN and LSTM technologies to enhance the accuracy and efficacy of the public opinion crisis prediction model. CNN excels at capturing local patterns and features within textual data, while LSTM adeptly models and analyzes temporal dependencies within textual data. By synergistically deploying these two techniques, the paper aims to construct a model proficient in extracting pertinent features from social network user behavior data and accurately classifying and forecasting public opinion crises.

A. Integration of IoT Technology and Social Network

The IoT represents an extensive network, expanding its reach by amalgamating diverse information-sensing devices into the network infrastructure. This convergence empowers the IoT to comprehend the perpetual interconnectedness of people, machinery, and objects across all temporal and spatial dimensions [22]. The amalgamation of IoT and Artificial Intelligence (AI) into the concept of Artificial Intelligence of Things (AIoT) is poised to usher in a multitude of innovative applications and guide future trends. On one hand, the amalgamation of IoT devices with AI technologies enables the realization of intelligent production processes and automation control. For instance, within the manufacturing sector, AIoT has the capability to oversee device operational statuses, gather production data, and employ machine learning algorithms for comprehensive analysis and optimization. This, in turn, facilitates heightened efficiency in production scheduling and enhanced quality control. On the other hand, AIoT holds the potential to exert a transformative influence on various domains, including urban management, transportation, and healthcare. For instance, through the integration of IoT sensing devices with AI algorithms, AIoT can pave the way for smart transportation systems, ultimately streamlining traffic patterns and ameliorating congestion. In the healthcare domain, AIoT can be harnessed for health monitoring and remote medical services, affording individuals personalized health management and timely access to medical assistance. In summary, the amalgamation of IoT and AI into AIoT stands poised to significantly augment efficiency and the degree of intelligence across various sectors. This amalgamation heralds fresh prospects for the establishment of more intelligent and interconnected systems, ultimately fostering improved outcomes and enhancing quality of life.

The IoT can be envisaged as an extension and amplification of the conventional Internet, constituting its core and foundational underpinning. It encompasses a broad spectrum of devices and commodities designed for communication and the exchange of information. Within this context, the IoT denotes a network that establishes connections between diverse objects and the Internet, thereby enabling seamless information exchange and communication through well-established protocols. This exchange is facilitated by information sensing devices such as radio frequency identification (RFID), infrared sensors, global positioning systems, and laser scanners, which empower intelligent object identification, positioning, tracking, monitoring, and management. The overarching architecture of the IoT is visually depicted in Fig. 1.

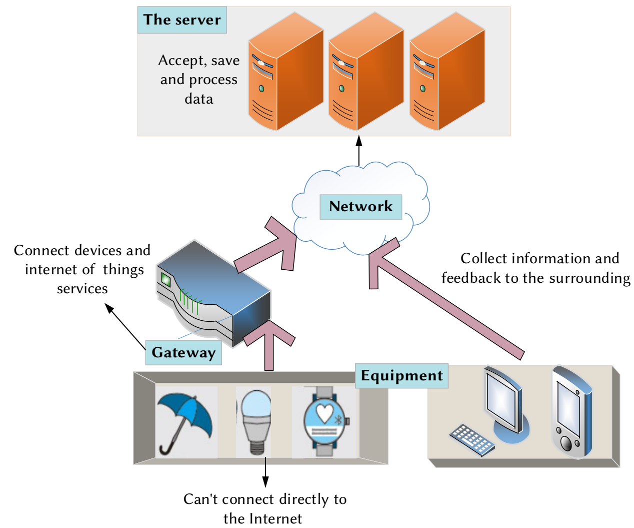


Fig. 1. Overall architecture of IoT. (Source: self-drawn by the author, icon source: Visio and https://www.iconfont.cn/).

The IoT, as illustrated in Fig. 1, operates within a meticulously structured framework comprising hardware, networking infrastructure, and cloud computing components. The seamless interconnection between the IoT and the broader Internet ecosystem is made possible by integrating pivotal IoT technologies. At the heart of the IoT's intricately layered architecture resides the sensor network, which harnesses cutting-edge two-dimensional code and RFID technology to discern and engage with a diverse array of interconnected entities. The application network, functioning as the ingress and egress control terminal, affords accessibility through an array of devices, including but not limited to mobile phones and other compatible interfaces. Conversely, the transmission network serves as the backbone for data dissemination and computation, capitalizing on the established infrastructure of the Internet, as well as conventional mediums such as radio and television networks, communication networks, and emerging next-generation networks. It is very important to underscore that each facet of social networking applications hinges on advanced hardware replete with specialized coding, differentiating between an extensive spectrum of devices, spanning from stereos to mobile phones. Through the judicious deployment of IoT, this paper attains its predefined objectives, concurrently harnessing data to elevate product quality and gain deeper insights into prevailing usage patterns within the realm of social media. A comprehensive elucidation of the operational workflow of the IoT is thoughtfully presented in Fig. 2.

As delineated in Fig. 2, an IoT device is thoughtfully equipped with both network and sensor modules, thus endowing it with the capability to seamlessly upload data to and retrieve data from the cloud server. Significantly, the cloud server serves as the central hub for orchestrating the management of intelligent hardware components. The network infrastructure emerges as the linchpin, facilitating the transmission and computation of data, thereby enabling effective control and interaction with the application network. The application network's input/output control terminal, which functions as the gateway to access its functionalities, exhibits compatibility with a multitude of established devices, including but not limited to mobile phones, personal computers (PCs), and other compatible terminals. In this context, it is paramount to underscore the pivotal role played by the IoT in the aggregation of user activity data extracted from social networks. This comprehensive data corpus serves as the foundation for conducting meticulous text data mining and analysis, empowering the various components of social networks to make well-informed judgments. For a deeper comprehension of IoT and social networks convergence, this paper presents the research findings concisely summarized in Table I.

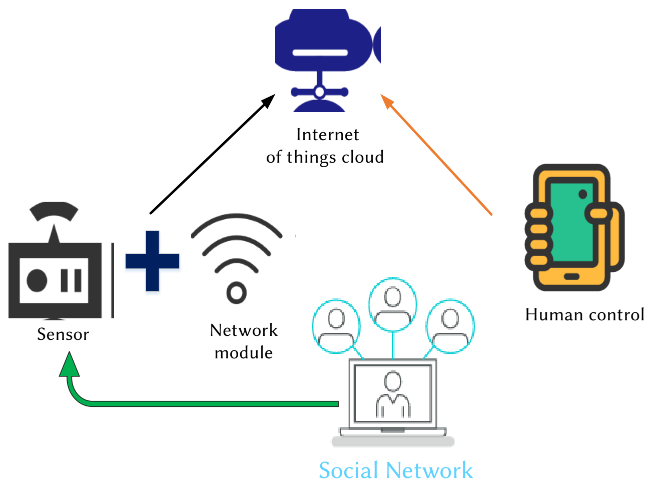


Fig. 2. The operation process of IoT. (Source: self-drawn by the author, icon source: Visio and https://www.iconfont.cn/).

TABLE I. Discussions on the Integration Technology of IoT Technology and Social Network

Research scholar	Specification
Wang, Wang, Li, Leung & Taleb (2020) [23]	Based on the social network, the sharing of things in the IoT is realized.
Lombardi, Pascale & Santaniello (2021) [24]	Social IoT realizes resource sharing and service between objects.
Gupta & Quamara (2020) [25]	The Architecture of Social IoT
Javaid & Khan (2021) [26]	A research method and model with social network characteristics is formed to realize the network application of the IoT, adopting the social network model, using the social relations among people, people and things, things, and things.

III. Research Model

A. CNN Algorithm

CNN, short for Convolutional Neural Network, represents a multi-layer perceptron explicitly designed for the extraction of localized features from images. It accomplishes this task through the implementation of convolution and pooling techniques. In stark contrast to fully connected neural networks, CNN is distinguished by three key attributes: localized perception, weight sharing, and the deployment of multiple convolution kernels. The process of localized perception hinges on the coordinated operation of upper neurons and convolution kernels. Their collaborative effort is directed toward the extraction of localized features from the input data originating from lower neurons, leading to the creation of a novel feature map. Weight sharing is a fundamental strategy employed to ensure that features are systematically extracted from the same feature graph. This technique serves to curtail the proliferation of network parameters by relying on a shared set of convolution kernels within the same feature graph. Furthermore, to encompass a broader spectrum of image information, a multi-convolution kernel conducts convolution operations on a feature map, employing multiple convolution kernels for the purpose. The structural organization of CNN is thoughtfully depicted in Fig. 3.

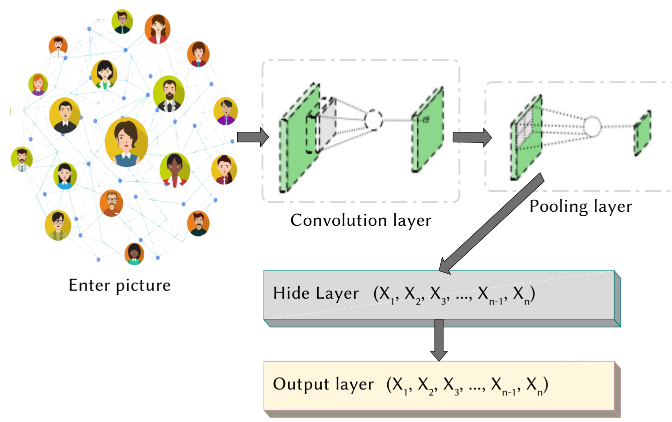


Fig. 3. CNN structure. (Source: self-drawn by the author, icon source: Visio and https://www.iconfont.cn/).

The foundational architectural framework of the CNN, as depicted in Fig. 3, comprises five principal layers: the input layer, convolution layer, pooling layer, fully connected layer, and output layer. Commencing with the input layer, it serves as the recipient of data, arriving in the configuration of an image pixel array. The convolution layer shoulders the responsibility of isolating local features from the input image. Subsequently, the pooling layer undertakes the task of transforming the extracted data features into a linear format, employing a down-sampling technique. This transformation yields data points that are subsequently transmitted to the output layer, which ultimately generates the desired results [27], [28]. Mathematically, the operation of the convolution layer is succinctly expressed as shown in (1):

$$y_{mn} = f\left(\sum_{j=0}^{Q-1}\sum_{j=0}^{P-1} x_{m+i,n+j} w_{ij} + b\right), 0 \le m < M, 0 \le n < N \quad (1)$$

In Equation (1), $x_{m+i,n+j}$ is the pixel size of the input image, and y_{mn} is the output data of convolution operation. b is the offset, $P \times Q$ is the size of the convolution kernel, and w_{ij} is the value of the convolution kernel in (i, j). M and N are the sizes of the input image on $P \times Q$. Equation (2) and (3) show the excitation function:

$$Softmax(x) = \frac{e^{x^i}}{\Sigma_i e^{x^i}} \quad (2)$$

$$ReLU(x) = \max(0, x) = \begin{cases} 0, x < 0 \\ x \;\; x > 0 \end{cases} \quad (3)$$

The image size after convolution is shown in Equation (4):

$$N = \frac{W - F + 2P}{S} + 1 \quad (4)$$

In Equation (4), W is the extracted data size, F is the size of the convolution kernel, P is the number of padding data columns at the edge of the original data, and S is the moving step size. Equation (5) shows the pooling operation:

$$a^l = pool(a^{l-1}) \quad (5)$$

In Equation (5), *pool* represents the process of reducing the size of input data by pooling area k and pooling standard, and a^{l-1} is the input tensor obtained by edge filling of the convolution input data matrix.

Equation (6) shows the output of the full connection layer:

$$a^l = \sigma(z^l) = \sigma(W^l a^{l-1} + b^l) \tag{6}$$

In Equation (6), b^l is the threshold of the fully connected layer l, and σ is the activation function. In the training process of CNN, the forward propagation is shown in (7)(8)(9):

$$z^{(l)} = W^{(L)} \cdot a^{(l-1)} + b^{(l)} \tag{7}$$

$$a^{(l)} = f_l(z^{(l)}) \tag{8}$$

$$x = a^{(0)} \to z^{(1)} \to a^{(1)} \to z^{(2)} \to a^{(2)} \to \cdots \to z^{(l)} \to a^{(l)} \tag{9}$$

In Equations (7)(8)(9), l is the number of layers of the CNN, L is the number of neurons, $W^{(L)}$ is the weight, $b^{(l)}$ is the bias, $a^{(l)}$ represents the network prediction output of the forward operation, $z^{(l)}$ is the input of the l-layer neuron, and $f_l(\cdot)$ is the activation function. Equation (10) shows the difference between the test value and the average value of the backward propagation output:

$$\delta^l = \frac{\partial J(W,b,x,y)}{\partial a^l} \odot \sigma'(z^l) \tag{10}$$

In Equation (10), $J(W, b, x, y)$ is the variance, z^l is the input of l-layer neurons, a^l is the output of l-layer neurons, and $\sigma'(\cdot)$ directs the number operation. The process of updating the parameters (W, b) by gradient descent of the weights and offset vectors of the l-layer network is shown in (11)(12)(13):

$$\delta^l = (W^{(l+1)})^T \cdot \delta^{(l-1)} \odot \sigma'(z^l) \tag{11}$$

$$\frac{\partial J(W,b,x,y)}{\partial W^l} = \delta^l (a^{l-1})^T \tag{12}$$

$$\frac{\partial J(W,b,x,y)}{\partial b^l} = \delta^l \tag{13}$$

In Equation (11), the gradient calculation of the weight and the offset vector can be obtained, as shown in Equations (12) and (13). Then, all parameters (W, b) are updated using random gradient descent to complete the feedback operation [29].

B. Neural Network Algorithm for LSTM

The LSTM network stands apart from other deep learning architectures due to its distinctive unit state mechanism. It exercises selective information transmission through the utilization of a gate structure, enabling the incorporation or removal of data from the cell state. This operation generally entails the application of a sigmoid neural network layer in conjunction with point-wise multiplication procedures [30]. The structural arrangement of the LSTM neural networks is delineated in Fig. 4:

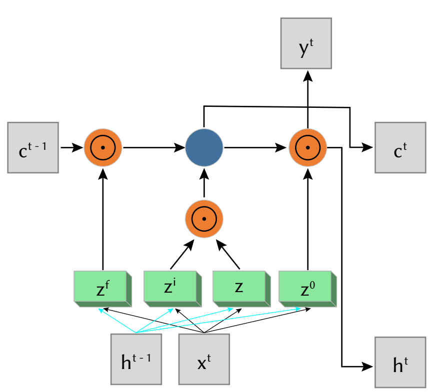


Fig. 4. Neural network structure of LSTM (Source: self-drawn by the author).

The LSTM neural network structure illustrated in Fig. 4 comprises two LSTM units, each featuring three gate structures: the input gate, the forgetting gate, and the output gate. The primary function of these 'gates' is to regulate and update the cell state within the network. The forgetting gate employs the *sigmoid* activation function to evaluate whether specific cell state information should be retained or forgotten. The input gate, in conjunction with the *tanh* layer, is responsible for storing and preserving cell information, while the output gate facilitates the output of pertinent information. The sigmoid activation function plays a pivotal role in determining particular aspects of the output unit state. Subsequently, the *tanh* layer is employed to manipulate the cell state, and the final output value is derived as the product of the *sigmoid* layer and the *tanh* layer. The equations governing the LSTM network are delineated in (14)(15)(16)(17):

$$z = \tanh\,(w \cdot [x^t, h^{t-1}] + b) \tag{14}$$

$$z^i = \sigma(w^i \cdot [x^t, h^{t-1}] + b^i) \tag{15}$$

$$z^f = \sigma(w^f \cdot [x^t, h^{t-1}] + b^f) \tag{16}$$

$$z^o = \sigma(w^o \cdot [x^t, h^{t-1}] + b^o) \tag{17}$$

z^i, z^f and z^0 are values converted into [0,1] by *sigmoid* layer, z is values converted into $[-1,1]$ by *tanh* layer. b^i, b^f and b^0 are corresponding thresholds. The forgetting gate reads x^t and h^{t-1}, and decides which historical information to forget according to the current input information, and finally gets z. When the information is stored in the cell, the sigmoid layer determines the updated value, the tanh layer creates a new candidate value vector, and z is added to the state to update the cell state information of the sigmoid layer and tanh layer. The output h^{t-1} at the last moment, and the current data input x^t get z through the input gate. Then update the old cell state, and c^{t-1} is updated to c^t. The old state is multiplied by z^f and the information that is determined to be discarded is discarded. Then $z^i *$ z is added to generate new candidate values. According to the degree of the decision to update each cell state, the process of obtaining the temporary state c^t at the current moment through the unit state is changed. Finally, the output value determines which part of the cell state is output by running the *sigmoid* layer. Then, the cell state is processed by *tanh* layer and multiplied by the output of *sigmoid* gate, and the result is converted into numerical value by *tanh* activation function as the input signal. The final output determines the output port, the output h^{t-1} at the last moment, and the current data input x^t, and z^0 is obtained through the output gate, and the final output h^t is obtained by combining the cell states c^t and z^0 of the current cell. Through the *tanh* layer, the c^t obtained in the previous stage is scaled down. After obtaining the hidden layer state h^t, the output y^t and W' are often obtained by changing h^t as the output weight matrix [31], [32], the internal training process of LSTM is shown in (18)(19)(20).

$$c^t = z^f * c^{t-1} + z^i * z \tag{18}$$

$$h^t = z^o * \tanh\,(c^t) \tag{19}$$

$$y^t = \sigma(W' h^t) \tag{20}$$

C. Construction of a Social Network Text Classification Model Based on CNN and LSTM

Deep learning and IoT technology are used to identify and intelligently manage inappropriate comments on social networks to avoid public opinion crises. A social network text categorization model using CNN-LSTM and CNN is shown in Fig. 5.

In Fig. 5, initially, IoT technology is employed to gather social media user behavior data, encompassing users' activities such as posts, comments, likes, and shares on social networks. Subsequently, the assembled social media text data undergoes a pre-processing phase, which encompasses tasks such as text cleansing, tokenization, and the elimination of stop-words. These operations are undertaken to

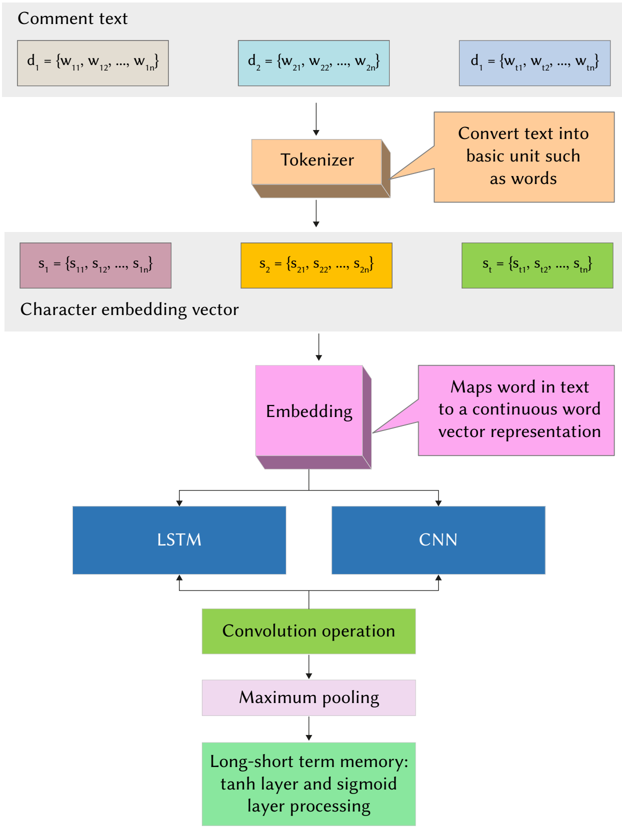


Fig. 5. Social network text classification model of CNN-LSTM network. (Source: self-drawn by the author).

transform the raw textual content into a format amenable for utilization as input within the model. The text cleaning phase involves the removal of noise, special characters, and unnecessary formatting from the text. This helps ensure that the text data input to the model is clean and consistently formatted. Tokenization is the process of segmenting the original text into words or tokens. This means converting the text into basic units, which are words or tokens, that the model can understand. This process aids the model in understanding the structure of the text. Stop words refer to words that frequently appear in the text but typically do not carry important information (e.g., "the," "is," "in"). In this step, stop words are removed to reduce the dimensionality and noise of the text data. Secondly, Word2Vec, a word embedding model, is used to map each word to a continuous word vector representation. This enables the capture of semantic relationships between words. The embedding layer is a crucial component used to map words in the text to continuous word vector representations. This helps the model understand the semantic relationships and contextual information between words. Word2Vec is a commonly used embedding model that transforms each word into a high-dimensional vector to capture the semantic similarity between words. This enables the model to better understand the meaning and associations of words in the text. In the initial layer of the model, a convolutional layer is employed to perform feature extraction. This layer conducts local feature extraction on the input sequence using sliding windows. Multiple convolutional kernels of various sizes can be utilized to capture features at different scales. In the subsequent layer of the model, an LSTM layer is utilized to capture long-term dependencies within the sequence effectively. The LSTM layer is well-suited for handling sequential data and incorporates a memory mechanism to retain crucial information. The output features from both the CNN and LSTM layers are then combined through a weighted sum. Following this, the fused features are fed into a fully connected layer for the purpose of classification, resulting in the prediction of the text's category. Subsequently, the model is trained using annotated training data, and the model parameters are updated through the backpropagation algorithm. To expedite convergence during this process, the Adam optimization algorithm can be employed. Once the model has been trained and evaluated, it can be utilized to predict and classify new social media text. As users post new content or comments on social networks, the constructed model can be employed to determine whether it falls within the category of a public opinion crisis. Based on the model's predictions, users can make informed decisions, which may include removing or revising inappropriate remarks or responding promptly to public sentiments.

The algorithm pseudocode for the social network text classification model based on the CNN-LSTM network is presented in Table 2.

TABLE II. Algorithm, Pseudocode of Social Network Text Classification Model, Based on CNN-LSTM Network

Step	**Algorithm Pseudocode**
Step 1: Data pre-processing; assume that the steps of data pre-processing and division into the training set and test set have been completed	import numpy as np from TensorFlow. Keras. Models import Sequential from tensorflow. Keras.layers import Embedding, Conv1D, MaxPooling1D, LSTM, Dense
Step2: Define the model	model = Sequential() model.add(Embedding(input_dim=vocab_size, output_dim=embedding_dim, input_length=max_length)) model.add(Conv1D(filters=num_filters, kernel_size=filter_size, activation='relu')) model.add(MaxPooling1D(pool_size=pool_size)) model.add(LSTM(units=lstm_units)) model.add(Dense(units=num_classes, activation='softmax'))
Step3: Compile the model	model.compile(loss='categorical_crossentropy', optimizer='adam', metrics=['accuracy'])
Step4: Model training	model.fit(train_data, train_labels, epochs=num_epochs, batch_size=batch_size)
Step5: Model evaluation	loss, accuracy = model.evaluate(test_data, test_labels)
Step6: Model application	predictions = model.predict(new_data)
Step7: Result analysis	—
Step8: Analyze and compare the results as needed	—
Step9: Complete the model building and evaluation process	—

IV. Results and Discussion

A. Experimental System Construction and Model Performance Evaluation

The experimental hardware environment comprises an Intel® Core™ i5-7200U CPU @ 2.50GHz. The software environment is Windows 10. The deep learning framework utilized is Keras, with the underlying deep learning framework being TensorFlow. The implementation is performed using the Python programming language with 20 iterations.

To evaluate the performance of the social network text classification model of the CNN-LSTM network, the study adopted a random split to divide the dataset into two parts, resulting in 70% of the data being used as the training set and 30% as the testing set. The purpose of splitting the dataset is to use the training set for parameter estimation when building the model and to evaluate the model's performance on the testing set. The accuracy, precision, recall rate, F1 value, and loss value are used for evaluation [33], as shown in (21)(22)(23)(24)(25):

$$Accuracy = \frac{TP+TN}{TP+FN+FP+TN} \tag{21}$$

$$Precision = \frac{TP}{TP+FP} \tag{22}$$

$$Recall = \frac{TP}{TP+FN} \tag{23}$$

$$F_1 = \frac{2*Precision*Recall}{Precision+Recal} \tag{24}$$

$$Loss = -\frac{1}{n}\sum_{i=1}^{n}\sum y^{(i)} log\hat{y}^{(i)} \tag{25}$$

In the context of the experimental evaluation, True Positive (TP) denotes the number of positive samples accurately identified, while True Negative (TN) represents the number of negative samples correctly classified. False Positive (FP) indicates the count of negative samples incorrectly identified as positive, and False Negative (FN) refers to the number of positive samples mistakenly classified as negative. The variable n represents the total number of training samples, $y^{(i)}$ corresponds to the actual (real) value, and $\hat{y}^{(i)}$ stands for the predicted value.

B. Experimental Data

The data extraction process primarily centered on the identification of specific keywords, including "network civility," "universities," "economy," "internal spiritual conflict," "culture," "New Year's Day," "Spring Festival," "Lantern Festival," "consumer rights," and various others. These keywords served as the focal points in assembling the research dataset for this paper. It is noteworthy that most of the Weibo comments comprising this dataset are characterized by brevity and conciseness. Specifically, there were 8,833 texts containing up to 30 Chinese characters, 9,799 texts with up to 40 Chinese characters, and 10,687 texts containing up to 100 Chinese characters. Furthermore, it is imperative to underscore that these comments exhibit colloquial expressions, and they may feature diverse emoji representations. Given these distinct attributes of the comment data, the experiment necessitated meticulous consideration of feature extraction integrity when selecting the appropriate text vectorization algorithm. Additionally, a denoising procedure was applied to eliminate extraneous comment elements, such as superfluous words, phrases, and punctuation marks. As a result of this data refinement process, a total of 7,368 comments were retained for inclusion in this paper. This corpus comprises 3,425 comments categorized as offensive remarks and 3,943 comments falling under the category of other comments.

C. Hyperparameter Settings

In the course of constructing the model, the learning rate assumes a critical role in governing the step size for updating model parameters, with a predetermined value of 0.001 being employed. Concurrently, the batch size, which dictates the quantity of samples utilized for parameter updates in each iterative step, was set at 128. In configuring the model's architecture, the word embedding dimension was established at 150, while the maximum permissible length for text sequences was defined as 380 characters. The convolutional layer featured 256 filters employing a kernel size of 5, while the pooling layer incorporated a window size of 2. The LSTM layer was equipped with 64 units, and the fully connected layer had an output category count of 16. Model training spanned 25 epochs to achieve the desired convergence.

D. Evaluation Performance of Social Network Text Classification Model Combining CNN-LSTM

Within the framework of the CNN-LSTM network algorithm, predicated on the theoretical underpinnings and the model's training regimen, the aforementioned text datasets were categorized into two distinct groups, denoted as Data1 and Data2. The specific outcomes of the model's classification process are graphically illustrated in Fig. 6:

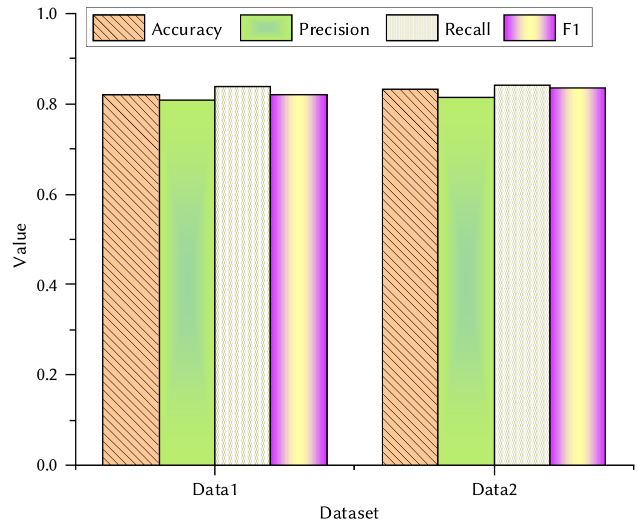


Fig. 6. Classification results based on Data1 and Data2 in this paper. (Source: self-drawn by the author).

Through a meticulous comparative examination of the Data1 and Data2 datasets, Fig. 6 effectively delineates that the model introduced in this scholarly work exhibits a nominal degree of divergence concerning its classification and evaluation metrics when applied to the two distinct datasets. Specifically, the performance metric for the social network text classification model, predicated on the CNN-LSTM network architecture, attains a value of 0.8217 when applied to Data1 and registers at 0.8315 for Data2. It is noteworthy that the Data2 model manifests an enhanced classification performance, denoting a 1.19% amelioration relative to its Data1 counterpart. Thus, it becomes apparent that the CNN's influence on the outcome becomes discernible, particularly in connection with the convolutional block encompassing the pooling operation, as the model interacts with the sequence data within the dataset. The experimental findings elucidating the performance of the CNN-LSTM network-based social network text categorization model is graphically depicted in Fig. 7.

In Fig. 7, the CNN-LSTM model demonstrates a commendable accuracy rate of 92.19% while concurrently achieving a relatively low loss value of 0.4075. Notably, when the CNN-LSTM model processes a solitary convolution block preceding each pooling operation, its performance registers a notable enhancement of 2.62%, concomitant with a substantial reduction of 16.11% in its loss value. Conversely, upon configuring the CNN-LSTM model to process three convolution blocks preceding each pooling operation, an even more substantial performance improvement of 4.43% is observed, accompanied by a notable loss reduction of 24.29%. These outcomes illuminate that the CNN-LSTM model's capacity is not profoundly influenced by the number of network layers it encompasses. An excessive increase in network layers may potentially lead to network degradation, ultimately detrimentally impacting the model's overall performance.

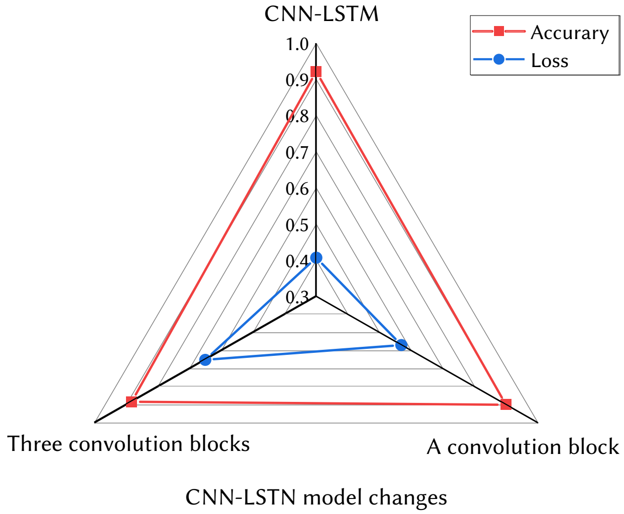


Fig. 7. Experimental effect of social network text classification model of CNN-LSTM network. (Source: self-drawn by the author).

E. Effectiveness Analysis of CNN-LSTM Algorithm in Social Network Text Classification

To substantiate the efficacy of the CNN-LSTM algorithm, this research undertakes a comparative assessment against independent CNN, LSTM, RNN, Transformer, and Generative Adversarial Network (GAN) algorithms employed for the task of social network text classification. The classification outcomes of these diverse algorithms on social network datasets are thoughtfully presented in Fig. 8.

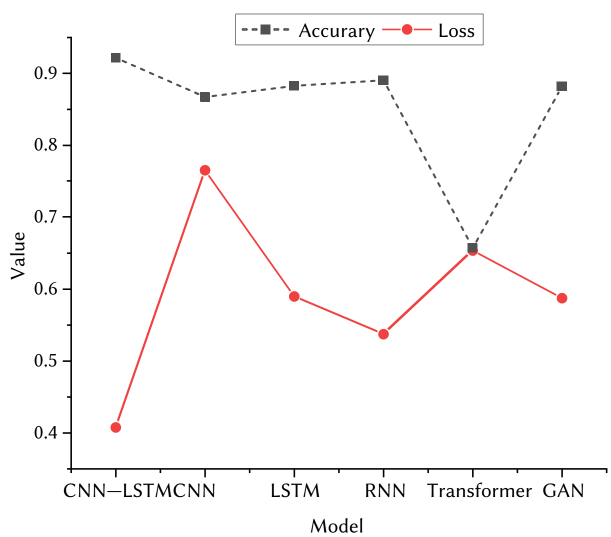


Fig. 8. Classification results of different algorithms in social network data sets. (Source: self-drawn by the author).

Fig. 8 provides a comprehensive comparative analysis of various deep learning algorithms, with CNN-LSTM emerging as the frontrunner, exhibiting the highest accuracy rate and the lowest loss rate within the context of this paper. Specifically, when applied to social network datasets, the CNN algorithm attains a classification accuracy of 86.72% alongside a loss value of 0.7653. The LSTM methodology achieves a commendable classification accuracy of 88.28%, accompanied by a loss value of 0.5898 on social network datasets. In contrast, the RNN algorithm impressively registers a classification accuracy of 89.06% while maintaining a relatively low loss value of 0.5373 on social network datasets. Notably, while demonstrating a reasonable classification accuracy of 65.71%, the Transformer algorithm corresponds to a loss value of 0.6535 on social network datasets. Furthermore, the GAN algorithm notably achieves a classification accuracy of 88.23% alongside a loss value of 0.5874 when operating on social network datasets. Conversely, the CNN-LSTM algorithm consistently outperforms the CNN algorithm, LSTM methodology, RNN algorithm, Transformer algorithm, and GAN algorithm in terms of classification accuracy when confronted with the intricacies of social network datasets. These compelling results underscore the remarkable efficacy and aptitude of the CNN-LSTM algorithm for the task of social network text categorization. The variation in classification performance across different datasets is thoughtfully presented in Fig. 9.

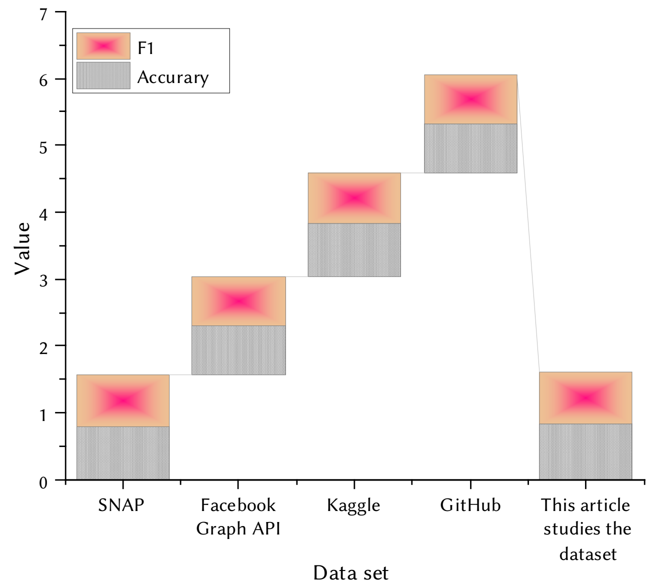


Fig. 9. Classification Effect of the Model Under Different Data Sets. (Source: self-drawn by the author).

Fig. 9 meticulously outlines the model's classification performance across distinct datasets. Notably, the dataset employed within the scope of this paper attains the highest level of accuracy, impressively reaching 84.29%. The Stanford Network Analysis Project (SNAP) dataset exhibits a commendable accuracy rate of 80.09%. Meanwhile, the accuracy for the Facebook Graph API dataset and the Kaggle dataset stands at 74.57% and 79.14%, respectively. The GitHub dataset yields an accuracy of 73.86%. It is noteworthy that the dataset utilized in this paper consistently demonstrates superior performance, not only in terms of accuracy but also in the F1 score metric. Equally noteworthy is the SNAP dataset, which also showcases commendable results in accuracy and F1 score. The remaining datasets, while still demonstrating noteworthy performance, exhibit slightly lower accuracy rates and F1 scores in comparison.

F. Discussion

This paper utilized LSTM and CNN, two deep learning techniques, to construct a predictive model for public opinion crises. Compared to standalone models such as CNN, LSTM, RNN, GAN, Transformer, etc., the CNN-LSTM model proposed in this paper demonstrates the highest accuracy and the lowest loss rate, achieving a significant advantage. The accuracy and reliability of the model have also been thoroughly validated in the experiments. The results indicate its outstanding performance in crisis prediction tasks. Furthermore, the consistent performance of the model on different datasets suggests its stability and robustness. These findings align with the research conducted by Kang et al. (2020), where it was observed that LSTM-based methods exhibited the capability to effectively address imbalanced datasets, subsequently leading to enhanced classification performance [34]. Consequently, given the inclusion of LSTM as a pivotal component in this paper, it is reasonable to anticipate favorable outcomes when dealing with opinion data within the context of social networks. Furthermore, the attention mechanism introduced by He et al. (2022) for fine-grained text classification introduces the prospect of information exchange between sentences and keywords during the text classification process. This innovative mechanism has the potential to bolster the model's comprehension and classification proficiency when handling social media text [35]. Additionally, Dang et al. (2020) conducted research that employed deep learning techniques for the analysis of question texts, achieving an impressive F-score of 80.20% within the specific text length range of 50 to 100 words [36]. Given the typically concise nature of text in social networks, these findings bear relevance to the present study. As the model presented in this paper is tailored for social network mining, it may encounter shorter text segments or sentences. Therefore, it is conceivable that the model in this paper may demonstrate similar effectiveness in the realm of social media text analysis. In addition, Saxena et al.'s (2023) research involves link prediction, which is the task of determining whether a connection will exist between two entities in a network, primarily applied in social network analysis, biological networks, and other network science domains. The study used a combination of network centrality and graph convolutional networks, enhancing model accuracy by selecting nodes with high betweenness centrality for model training [37]. This research provides a robust approach to link prediction and thoroughly discusses its performance in both experimental and theoretical aspects. Arroni et al. (2023) proposed a simpler attention-based model that utilizes a transformer architecture to predict sentiment expressed in tweets about Las Vegas hotels, comparing its performance with traditional sentiment analysis methods. Experimental results demonstrated the outstanding predictive performance of their model [38]. Although these two studies focus on different areas than the present paper, they also leverage deep learning techniques and have achieved significant results in different application domains, which is relevant to the discussion of the effectiveness and potential of social network text classification tasks in this paper. In summary, the CNN-LSTM model in this paper exhibits outstanding performance in handling social network text classification tasks, providing powerful tools and methods for predicting and managing public sentiment crises.

V. Conclusion

This paper harnessed the capabilities of the IoT and a web crawler to amass a corpus of 11,238 comments originating from the trending entertainment topics on the Weibo social networking platform. The primary objective was to discern and intelligently manage inappropriate comments within the realm of social networks, thereby proactively averting potential public opinion crises that could arise from the dissemination of such objectionable content. The performance of our social network text classification model, which amalgamates the CNN-LSTM network, underwent comprehensive evaluation, encompassing various metrics such as accuracy rate, precision, recall rate, F1 score, and loss value. The research outcomes revealed minimal disparities in classification across distinct data sets and assessment criteria when comparing the CNN-LSTM network with alternative models for social network text classification. Notably, the CNN-LSTM model exhibited an impressive accuracy rate of 92.19% along with a commendably low loss value of 0.4075. Intriguingly, our model's performance exhibited a diminishing trend as the number of network layers increased, thereby indicating that the network's efficacy is not significantly influenced by layer count. Additionally, the CNN-LSTM algorithm outperformed several other deep learning algorithms, underscoring its preeminence in the domain of social network text classification. When juxtaposed with the CNN algorithm, our CNN-LSTM algorithm showcased a noteworthy 6.31% enhancement in accuracy. Similarly, when compared to the LSTM algorithm, the CNN-LSTM algorithm demonstrated a commendable 4.43% increase in classification accuracy. Furthermore, it surpassed the RNN algorithm by 3.51%, the Transformer algorithm by a substantial 40.29%, and the GAN algorithm by 4.49% when applied to social network data sets. These compelling findings unequivocally validate the efficacy of the CNN-LSTM algorithm in the domain of social network text classification. The principal research findings presented in this paper carry substantial practical implications for the application of deep learning techniques in the context of managing social network crises and processing high-dimensional data. These findings furnish valuable tools and methodologies for social network governance and data analytics, which, in turn, can significantly contribute to the enhancement of social network services and the overall quality of decision-making processes.

Nonetheless, it is imperative to acknowledge certain caveats associated with this research. The ever-expanding landscape of social networks may have introduced unaccounted-for nuances and unrepresented social network datasets, potentially affecting the generalizability and robustness of our experiment's findings. Furthermore, our acquisition of the hot search comment dataset from microblog public opinion sources relied on a combination of web crawling and manual screening. Consequently, the dataset this paper compiled may be afflicted by incomplete coverage, which could potentially introduce biases into our overall results. The experiments and tests carried out in this paper were conducted within a simulated computer environment. Therefore, disparities may exist between our experimental outcomes and real-world results. The evolution of information technology and ongoing updates to filtering techniques may compel malicious actors to employ diverse countermeasures aimed at circumventing information filtering while disseminating their content. Thus, future research should prioritize developing and applying novel information filtering methods integrated with deep learning technology to tackle the multifaceted challenges posed by social networks. By establishing a comprehensive social network dataset for deep learning application research, this paper can pioneer advancements in filtering methodologies for mining public opinion information, thereby fortifying network security and enhancing information filtering within the dynamic landscape of social media.

Acknowledgements

This research was supported by the Natural Science Fund of Jilin Province. (Grant No.20220402014GH).

References

[1] F. Wei, S. Zeadally, P. Vijayakumar, N. Kumar, and D. He, "An intelligent terminal based privacy-preserving multi-modal implicit authentication protocol for internet of connected vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 3939-3951, 2020.

[2] W. Wang, N. Kumar, J. Chen, Z. Gong, X. Kong, W. Wei, and H. Gao, "Realizing the potential of the internet of things for smart tourism with 5G and AI," *IEEE Network*, vol. 34, no. 6, pp. 295-301, 2020.

[3] A. Shmatko, S. Barykin, S. Sergeev, and A. Thirakulwanich, "Modeling a logistics hub using the digital footprint method—the implication for open innovation engineering," *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 7, no. 1, pp. 59, 2021.

[4] Z. S. Ageed, S. R. Zeebaree, M. A. Sadeeq, M. B. Abdulrazzaq, B. W. Salim, A. A. Salih, and A. M. Ahmed, "A state of art survey for intelligent energy monitoring systems," *Asian Journal of Research in Computer Science*, vol. 8, no. 1, pp. 46-61, 2021.

[5] L. Nie, Y. Wu, X. Wang, L. Guo, G. Wang, X. Gao, and S. Li, "Intrusion detection for secure social internet of things based on collaborative edge computing: A generative adversarial network-based approach," *IEEE Transactions on Computational Social Systems*, vol. 9, no. 1, pp. 134-145, 2021.

[6] Z. Guo, K. Yu, Y. Li, G. Srivastava, and J. C. W. Lin, "Deep learning-embedded social internet of things for ambiguity-aware social recommendations," *IEEE Transactions on Network Science and Engineering*, vol. 9, no. 3, pp. 1067-1081, 2021.

[7] H. Zhang, W. Qu, H. Long, M. Chen, "The Intelligent Advertising Image Generation using Generative Adversarial Networks and Vision Transformer: A Novel Approach in Digital Marketing," *Journal of Organizational and End User Computing*, vol. 36, no. 1,pp.1-26, 2024.

[8] Z. Guo and H. Wang, "A deep graph neural network-based mechanism for social recommendations," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2776-2783, 2020.

[9] M. A. Al-Garadi, A. Mohamed, A. K. Al-Ali, X. Du, I. Ali, and M. Guizani, "A survey of machine and deep learning methods for internet of things (IoT) security," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1646-1685, 2020.

[10] R. A. Khalil, N. Saeed, M. Masood, Y. M. Fard, M. S. Alouini, and T. Y. Al-Naffouri, "Deep learning in the industrial internet of things: Potentials, challenges, and emerging applications," *IEEE Internet of Things Journal*, vol. 8, no. 14, pp. 11016-11040, 2021.

[11] M. Andronie, G. Lăzăroiu, M. Iatagan, C. Uță, R. Ștefănescu, and M. Cocoșatu, "Artificial intelligence-based decision-making algorithms, internet of things sensing networks, and deep learning-assisted smart process management in cyber-physical production systems," *Electronics*, vol. 10, no. 20, p. 2497, 2021.

[12] M. S. Rahman, N. C. Peeri, N. Shrestha, R. Zaki, U. Haque, and S. H. Ab Hamid, "Defending against the Novel Coronavirus (COVID-19) outbreak: How can the Internet of Things (IoT) help to save the world?," *Health Policy and Technology*, vol. 9, no. 2, pp. 136-138, 2020.

[13] R. Vinayakumar, M. Alazab, S. Srinivasan, Q. V. Pham, S. K. Padannayil, and K. Simran, "A visualized botnet detection system based deep learning for the internet of things networks of smart cities," *IEEE Transactions on Industry Applications*, vol. 56, no. 4, pp. 4436-4456, 2020.

[14] M. Humayun, N. Z. Jhanjhi, A. Alsayat, and V. Ponnusamy, "Internet of things and ransomware: Evolution, mitigation and prevention," *Egyptian Informatics Journal*, vol. 22, no. 1, pp. 105-117, 2021.

[15] Z. Lv, L. Qiao, J. Li, and H. Song, "Deep-learning-enabled security issues in the internet of things," *IEEE Internet of Things Journal*, vol. 8, no. 12, pp. 9531-9538, 2020.

[16] L. U. Khan, W. Saad, Z. Han, E. Hossain, and C. S. Hong, "Federated learning for internet of things: Recent advances, taxonomy, and open challenges," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1759-1799, 2021.

[17] S. Siboni, V. Sachidananda, Y. Meidan, M. Bohadana, Y. Mathov, S. Bhairav, and Y. Elovici, "Security testbed for Internet-of-Things devices," *IEEE Transactions on Reliability*, vol. 68, no. 1, pp. 23-44, 2019.

[18] I. H. Sarker, "Machine learning: Algorithms, real-world applications and research directions," *SN Computer Science*, vol. 2, no. 3, p. 160, 2021.

[19] K. Yu, L. Tan, S. Mumtaz, S. Al-Rubaye, A. Al-Dulaimi, A. K. Bashir, and F. A. Khan, "Securing critical infrastructures: deep-learning-based threat detection in IIoT," *IEEE Communications Magazine*, vol. 59, no. 10, pp. 76-82, 2021.

[20] P. Muñoz, E. Doñaque, A. Larrañaga, and J. Martínez Torres, "Tourism-Related Placeness Feature Extraction from Social Media Data Using Machine Learning Models," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 8, no. 4, pp. 176-181, 2023.

[21] S. Arroni, Y. Galán, X. Guzmán-Guzmán, E. R. Núñez-Valdez, and A. Gómez, "Sentiment Analysis and Classification of Hotel Opinions in Twitter With the Transformer Architecture," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 8, no. 1, pp. 53-63, 2023.

[22] X. Zhou, W. Liang, I. Kevin, K. Wang, H. Wang, L. T. Yang, and Q. Jin, "Deep-learning-enhanced human activity recognition for Internet of healthcare things," *IEEE Internet of Things Journal*, vol. 7, no. 7, pp. 6429-6438, 2020.

[23] X. Wang, C. Wang, X. Li, V. C. Leung, and T. Taleb, "Federated deep reinforcement learning for internet of things with decentralized cooperative edge caching," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 9441-9455, 2020.

[24] M. Lombardi, F. Pascale, and D. Santaniello, "Internet of things: A general overview between architectures, protocols and applications," *Information*, vol. 12, no. 2, pp. 87, 2021.

[25] B. B. Gupta and M. Quamara, "An overview of Internet of Things (IoT): Architectural aspects, challenges, and protocols," *Concurrency and Computation: Practice and Experience*, vol. 32, no. 21, e4946, 2020.

[26] Javaid, M., & Khan, I. H. (2021). Internet of Things (IoT) enabled healthcare helps to take the challenges of COVID-19 Pandemic. Journal of Oral Biology and Craniofacial Research, 11(2), 209-214.

[27] G. Lou and H. Shi, "Face image recognition based on convolutional neural network," *China Communications*, vol. 17, no. 2, pp. 117-124, 2020.

[28] N. Lu, G. Wu, Z. Zhang, Y. Zheng, Y. Ren, and K. K. R. Choo, "Cyberbullying detection in social media text based on character-level convolutional neural network with shortcuts," *Concurrency and Computation: Practice and Experience*, vol. 32, no. 23, e5627, 2020.

[29] H. Hu, Z. Liu, and J. An, "Mining mobile intelligence for wireless systems: a deep neural network approach," *IEEE Computational Intelligence Magazine*, vol. 15, no. 1, pp. 24-31, 2020.

[30] K. Kumari and J. P. Singh, "Identification of cyberbullying on multi-modal social media posts using genetic algorithm," *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 2, e3907, 2021.

[31] M. Umer, I. Ashraf, A. Mehmood, S. Kumari, S. Ullah, and G. Sang Choi, "Sentiment analysis of tweets using a unified convolutional neural network-long short-term memory network model," *Computational Intelligence*, vol. 37, no. 1, pp. 409-434, 2021.

[32] I. Priyadarshini and C. Cotton, "A novel LSTM–CNN–grid search-based deep neural network for sentiment analysis," *The Journal of Supercomputing*, vol. 77, no. 12, pp. 13911-13932, 2021.

[33] Y. Bao, Z. Huang, L. Li, Y. Wang, and Y. Liu, "A BiLSTM-CNN model for predicting users' next locations based on geotagged social media," *International Journal of Geographical Information Science*, vol. 35, no. 4, pp. 639-660, 2021.

[34] H. Kang, H. Wu, and X. Zhang, "Generative text steganography based on LSTM network and attention mechanism with keywords," *Electronic Imaging*, vol. 2020, no. 4, pp. 291-1, 2020.

[35] X. He, P. Tai, H. Lu, X. Huang, and Y. Ren, "A biomedical event extraction method based on fine-grained and attention mechanism," *BMC Bioinformatics*, vol. 23, no. 1, pp. 1-17, 2022.

[36] N. C. Dang, M. N. Moreno-García, and F. De la Prieta, "Sentiment analysis based on deep learning: A comparative study," *Electronics*, vol. 9, no. 3, pp. 483, 2020.

[37] R. Saxena, S. P. Pati, A. Kumar, M. Jadeja, P. Vyas, V. Bhateja, and J. C. W. Lin, "An Efficient Bet-GCN Approach for Link Prediction," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 8, no. 1, pp. 38-52, 2023.

[38] S. Arroni, Y. Galán, X. Guzmán-Guzmán, E. R. Núñez-Valdez, and A. Gómez, "Sentiment Analysis and Classification of Hotel Opinions in Twitter With the Transformer Architecture," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 8, no. 1, pp. 53-63, 2023.

Yan Lou

Yan Lou was born in Jilin. P.R. China, in 1981. She received the Doctor's degree from Changchun University of Science and Technology, P.R. China. Now, she is the associate professor of Changchun University of Science and Technology. Her research interest include computational intelligence, free space optical communication. E-mail: Lyan@cust.edu.cn

Zhipeng Ren

Zhipeng Ren was born in Jilin. P.R. China, in 1980. She received the Master's degree from Changchun University of Science and Technology, P.R. China. Now, He is the teacher of Changchun University of Science and Technology. He also is a PhD candidate working on computer science and technologyPhD candidateHis research interest include computational intelligence. E-mail: rzp@cust.edu.cn

Yong Zhang

Yong Zhang was born in Changchun, Jilin, P.R. China, in 1973. She received the Doctor‘s degree from Jilin University, P.R. China. Now, she is the professor of Changchun Guanghua University. Her research interest include machine vision & artificial intelligence.
E-mail: zhangy_19730626@163.com

Zonghui Tao

Zonghui Tao was born in Liaoning. P. R. China, in 1986. She received the Master‘s degree from Changchun University of Science and Technology, P.R. China. Now, she is the experimental division of Changchun University of Science and Technology. Her research interest include information security and big data analysis.
E-mail: taozonghui@cust.edu.cn

Yiwu Zhao

Yiwu Zhao was born in Changchun.Jilin P.R.China,in 1972. He received the Doctor’s degree from Jilin University, P.R. China. Now, He is the Professor of Changchun University of Science and Technology. His research interest include Computational mathematics, neural networks .
E-mail: zyw@cust.edu.cn

Design of Traffic Electronic Information Signal Acquisition System Based on Internet of Things Technology and Artificial Intelligence

Hongling Wang

Harbin University Key Laboratory of Heilongjiang Underground Engineering Technology, Harbin University, Harbin, 150086 (China)

* Corresponding author: wanghongling@hrbu.edu.cn

Received 14 April 2023 | Accepted 3 May 2024 | Early Access 20 August 2024

Abstract

This study aims to devise a traffic electronic information signal acquisition system employing Internet of Things and artificial intelligence technologies, offering a novel approach to address prevailing challenges related to traffic congestion and safety. Initially, the hardware circuit for the high-speed signal acquisition control core is developed, leveraging Field-Programmable Gate Array technology. This facilitates wireless monitoring of signal acquisition. Subsequently, a comprehensive time signal acquisition system is formulated, encompassing modules for communication, acquisition, storage, adaptive measurement, and signal analysis. The geomagnetic acquisition module within this system is utilized for collecting geomagnetic signals, which are then translated into switch signals indicating the presence or absence of vehicles. These signals are subsequently transmitted to the geomagnetic signal processor. Experimental results pertaining to the signal acquisition system reveal a notable peak storage speed of 200KB/s, considering the utilization of one million sampling points. Across a series of tests, the maximum relative error of the obtained results ranges from 2.2% to 2.7%, underscoring the consistency and reliability of the measurements. In comparison to existing testing devices, the system exhibits heightened accuracy in test results, rendering it more apt for traffic signal acquisition applications. In conclusion, this study accomplishes the collection and dissemination of diverse traffic information, furnishing robust support for traffic control and ensuring safe operations.

Keywords

Artificial Intelligence, Internet of Things, Signal Acquisition System, Traffic Electronic Information.

DOI: 10.9781/ijimai.2024.08.002

I. Introduction

PRESENTLY, urban transportation has faced significant and urgent challenges, including inadequate transportation infrastructure development, escalating traffic congestion, frequent traffic accidents, and a fundamentally flawed travel structure. These challenges profoundly impact the quality of life for urban inhabitants and hinder the sustainable advancement of cities. Traditional approaches, however, have limitations in effectively addressing these complexities. For example, conventional methods of acquiring traffic information rely predominantly on stationary sensor apparatus, restricting both the scope and immediacy of data collection. Additionally, established vehicle positioning technologies, such as the global positioning system (GPS), provide positional precision but are hindered by elevated energy consumption and cost, limiting their extensive integration within intelligent transportation systems (ITSs). However, vehicle positioning remains a crucial supporting technology for intelligent traffic information acquisition, holding vital practical significance [1].

Traffic information collection is a pivotal aspect of urban ITSs and the broader traffic domain. Effective traffic management and control heavily depend on acquiring precise and up-to-the-minute traffic information [2]-[4]. By leveraging contemporary digital technologies, such as the Internet of Things (IoT) and wireless sensor networks (WSNs), enables the automated acquisition, amalgamation, and transmission of critical data like vehicle positioning, traffic flow, and road occupancy. This technological integration facilitates the provision of more accurate and timely data, empowering urban traffic management and control endeavors. Furthermore, this technological approach plays a pivotal role in addressing persistent challenges related to traffic congestion and vehicular accidents. A WSN is a novel form of an intelligent application network capable of autonomously collecting, fusing, and transmitting data. It assumes a significant role in urban ITS, effectively addressing urban traffic problems [5]-[7]. The communication tree tracking method solves the target tracking challenge in the sensor network, ensuring efficient target tracking and minimal node communication costs.

Please cite this article as:
H. Wang. Design of Traffic Electronic Information Signal Acquisition System Based on Internet of Things Technology and Artificial Intelligence, International Journal of Interactive Multimedia and Artificial Intelligence, vol. 8, no. 7, pp. 97-105, 2024, http://dx.doi.org/10.9781/ijimai.2024.08.002

WSNs play a crucial role in acquiring traffic information, encompassing data on vehicle speed, traffic flow, road occupancy rates, and intersection traffic conditions [8]. Overcoming the limitations of traditional monitoring sensors, which hinder system scalability and network efficiency, WSNs offer enhanced information acquisition accuracy. The integration of multi-source traffic information further augments the efficiency of monitoring and management tasks, including parking management, electronic toll collection, energy conservation, and emission reduction. In certain scenarios, sensor nodes deployed randomly within an area can obtain location information through positioning technology [9], [10]. While the Global Positioning System (GPS) is the prevailing positioning technology, GPS receivers are unsuitable for sensor networks due to their high energy consumption and cost. Locating vehicles via WSN presents several advantages. Firstly, it ensures high tracking accuracy, as WSN nodes are widely distributed, providing a clear understanding of changes in the target's geographic location. Secondly, the tracking is reliable, allowing system configurations to be fine-tuned for detailed, accurate, and reliable motion information during target tracking. Thirdly, it offers efficient tracking, as existing WSNs can simultaneously monitor and track various sensors within a specified range. Lastly, it is easy to implement, as sensor nodes are cost-effective, compact, and easily portable, facilitating concealment [11]-[13]. Therefore, WSN target tracking offers significant concealment and feasibility.

This study attempts to deploy WSNs across various stages of information collection, transmission, processing, and feedback within ITS to modernize and elevate the current traffic electronic information signal collection. Initially, an analysis is conducted on the characteristics and requisites of intelligent transportation in the IoT environment. Building upon this analysis, a high-speed signal acquisition control core, based on a Field Programmable Gate Array (FPGA), is designed to enable wireless monitoring of signal acquisition. Subsequently, a real-time signal acquisition system is formulated to effectively gather and process signals, comprising communication, acquisition, storage, adaptive measurement, and signal analysis modules. Additionally, the geomagnetic acquisition module's key technology is thoroughly examined and designed. Anticipated outcomes include enhancing traffic electronic information signal collection effectiveness and elevating traffic information service quality within ITSs.

This study aims to design and implement a traffic electronic information signal acquisition system utilizing IoT and AI technologies. The primary aim is to address existing challenges related to urban traffic congestion and traffic safety. Specifically, the study seeks to achieve wireless monitoring and acquisition of traffic signals through the innovative application of geomagnetic signal acquisition and processing techniques. This initiative targets the enhancement of signal acquisition speed, precision, and efficiency to provide robust information support for traffic control, management, and safe operations. The overarching goal is to optimize traffic flow, alleviate congestion, elevate traffic safety, and offer pioneering resolutions for urban traffic issues. The devised traffic electronic information signal acquisition system is constructed based on IoT technology and artificial intelligence. The distinctiveness of this approach emerges from the fusion of IoT and AI to facilitate intelligent data collection and processing for accurate traffic signal monitoring, transcending the limitations of conventional methodologies. Furthermore, the system employs geomagnetic signal acquisition and conversion to acquire precise vehicle presence or absence information, thereby enabling high-precision vehicle monitoring and furnishing dependable data for traffic control and management. The system caters to a wide spectrum of traffic information, thereby furnishing robust information backing for traffic control and safety operations. Consequently, it contributes to optimizing traffic flow, mitigating traffic congestion, and enhancing overall traffic safety. The integration and application of IoT and AI technologies empower the proposed traffic electronic information signal acquisition system to accommodate diverse scenarios and environments, showcasing commendable applicability and adaptability.

The structure of this study unfolds in the following sections: Section I, designated as the introduction, provides an overview of prevalent issues and challenges in urban transportation. It emphasizes the limitations of conventional approaches, underscores the significance of employing IoT and AI technologies to address transportation predicaments, and outlines the research objectives. Section II, comprising the literature review, conducts an examination of pertinent research domains. It delves into the application of IoT technology within intelligent transportation and scrutinizes the constraints inherent in existing methodologies. Section III, encompassing the research methodology, expounds upon the pivotal methodologies and technologies leveraged to tackle urban transportation issues. This section covers topics such as IoT-supported intelligent transportation, FPGA-based high-speed signal acquisition, the underpinning principles of geomagnetic signal acquisition, and the design of an adaptive, real-time signal acquisition system for traffic. It culminates in the formulation and design of a real-time signal acquisition system. Section IV, titled "Experimental Design and Performance Evaluation," delineates the orchestrated experimental procedures, including data acquisition and performance assessment. It further elucidates the test outcomes and the system's performance prowess. Section V, denominated as "Conclusion," provides a synthesis of the research content and methodologies encapsulated in the thesis. This segment culminates in overarching conclusions, and avenues for future research are envisaged.

II. Literature Review

A. Research Progress and Applications of AI

With the rapid evolution of AI technology, its integration into intelligent transportation has grown steadily. Akhtar and Moridpour (2021) provided a comprehensive synthesis of existing research on traffic congestion anticipation, incorporating various AI approaches with a prominent focus on diverse machine learning models. The authors meticulously categorized these models, offering a succinct overview of their merits and drawbacks [14]. Abduljabbar et al. (2019) elucidated the swift progress of AI within the transportation domain, highlighting its versatile applications in overcoming transportation challenges. The study specifically underscored AI's potential in data analysis, predictive modeling, and optimized decision-making, thereby enhancing transportation system efficiency, reliability, and sustainability [15]. Wu et al. (2022) explored AI's role in the context of smart city construction, particularly its current standing in intelligent transportation infrastructure. The research delved into scrutinizing diverse dimensions such as spatial typology, functional classifications, and facility utilization. The findings illuminated the substantial advantages of AI technology in classifying and administrating transportation infrastructure [16]. In summary, numerous AI methodologies have found widespread application in fields encompassing traffic flow projection, signal optimization, and accident forewarning. These methodologies possess the capability to refine and elevate transportation systems. Their potency lies in their adeptness at scrutinizing extensive traffic datasets, distilling pertinent features, and executing astute decisions.

B. Research Progress and Applications of IoT

Recently, the rise of IoT technology has opened new avenues for advancing ITSs. Wang and Ma (2022) conducted a focused investigation on the recognition and classification of stationary vehicles and seat belts within intelligent IoT-based traffic management systems. An innovative identification algorithm was introduced for the surveillance of stationary vehicles, leading to a substantial enhancement in detection accuracy compared to conventional background differential algorithms. Moreover, a proficient driver localization algorithm was formulated for driver seatbelt detection, uniting a target detection algorithm with a streamlined network structure, effectively elevating localization precision [17]. Ushakov et al. (2022) gathered insights from multiple European transportation agencies concerning public transportation. Through a comprehensive case study, they probed the far-reaching effects of IoT on the global transportation system. The study illuminated IoT's expansive potential within transportation, foreseen to amplify both system efficiency and safety [18]. Muthuramalingam et al. (2019) underscored the pivotal role of IoT solutions within the worldwide ITS, particularly in the domain of intelligent transportation marked by vehicle-to-vehicle communication. The authors delineated how IoT-based ITS can automate transport across railways, roadways, airways, and oceans, augmenting the logistics of cargo transportation, monitoring, and delivery, consequently enhancing customer experiences [19]. By interconnecting sensors, devices, and networks, real-time aggregation, transmission, and analysis of transportation data can be effectively achieved. IoT technology seamlessly amalgamates various facets of the transportation system, fostering comprehensive data utilization and augmenting traffic management intelligence.

C. Research Review

The aforementioned studies exemplify the utilization of AI and IoT technologies in intelligent transportation, approaching the subject from diverse vantage points. However, a dearth of research specifically focuses on the traffic signal acquisition system in isolation. Drawing from the principles of fuzzy control, this study introduces a fresh methodology to refine the traffic information acquisition system, aiming to enhance both the efficiency and the intelligence of traffic management.

III. Research Methodology

A. Intelligent Transportation Supported by IoT Technology

Entities need to communicate with each other, giving rise to the necessity for machine-to-machine (M2M) communication. Utilizing wireless short-range communication technologies is a viable approach, such as Wi-Fi, Bluetooth, and ZigBee, or large-scale mobile communication technologies, including World Interoperability for Microwave Access, Long Range, Sigfox, CAT M1, NB-IoT, Global System for Mobile Communications, General Packet Radio Service, the 3rd Generation Telecommunication, the 4th Generation Telecommunication, Long Term Evolution, and the 5th Generation Telecommunication [20]. Maintaining the affordability of IoT devices is paramount, especially considering their extensive usage across various daily life applications. Furthermore, IoT devices must possess the capability to fulfill basic tasks such as data collection, M2M communication, and even pre-processing data according to application requirements.

Intelligent traffic research represents a promising avenue for addressing urban traffic challenges. Advanced urban rail transit systems play a crucial role in providing residents with access to both dynamic traffic data and static information [21], [22]. Public transportation offers significant advantages, including substantial passenger capacity, enhanced transportation efficiency, minimal energy consumption, and low transportation costs. Consequently, urban informatization becomes a vital focus of research. IoT serves as a robust platform for intelligent transportation, facilitating the exchange of vehicle information through the network without human intervention. This enables intelligent transmission and sharing among vehicles. Navigation and route optimization stand out as pivotal aspects of intelligent transportation. Applications can leverage data from a user's mobile device or a roadside unit at a designated location to estimate traffic congestion and propose optimal route options. This approach minimizes travel time, thereby mitigating vehicle emissions and reducing energy consumption. Furthermore, a proposal is made for the introduction of smart streetlights equipped to detect traffic conditions and adjust illumination accordingly, aiming to contribute to energy conservation [23]-[25].

Fig. 1 shows the architecture of the IoT-based ITS.

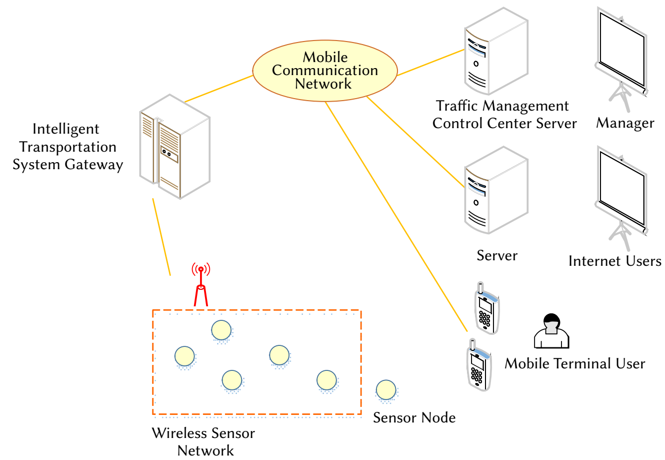


Fig. 1. IoT-based ITS architecture.

Intelligent transportation WSN is a purpose-driven wireless self-organizing network system, typically composed of multiple data convergence points and an array of sensor nodes dispersed throughout the ambient surveillance region. These nodes incorporate radio transponders, sensors, embedded processors, and more, enabling them to acquire, process, and transmit traffic data [26]. Network simulation software is employed to make these models functional, allowing a vehicular ad hoc network to utilize vehicle motion models [27]. Different scenarios are generated prior to the simulation, and the emulation program analyzes these scenarios based on a predefined path layout. Special applications in vehicular communication impose essential interaction requirements, facilitating communication between the two domains. Fig. 2 illustrates an isolated method of interaction between the transportation emulation program and the network simulation software.

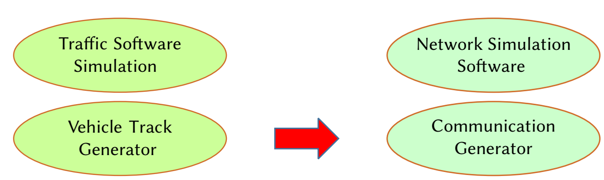


Fig. 2. An isolated method of interaction between traffic simulation software and network simulation software.

The two simulators are seamlessly integrated into a unified system to facilitate comprehensive interaction between the network and traffic simulation software. A straightforward collaboration between the web and mobile domains compensates for the absence of a protocol. The embedded approach provides the advantage of a streamlined and efficient interaction between the network and mobility models. This method utilizes validated vehicle motion models and strictly adheres to standard protocols.

B. FPGA-Based High-Speed Signal Acquisition

While high-resolution imagery is not obligatory for traffic sign detection, it does enhance the detection range in the Advanced Driver Assistance System (ADAS). Contemporary high-end processors boast sufficient computing power for executing these tasks, albeit at the cost of significant energy consumption [28]-[30]. Nevertheless, low power consumption and reliability are paramount for embedded systems like ADAS. In this context, FPGAs emerge as a potential solution to this challenge, as they can dynamically adjust their hardware to meet the current requirements of the application.

Information collection and manipulation systems have broad applications in metering and regulating systems. The information collection process involves measuring diverse electrical phenomena, including sound, pressure, temperature, current, or voltage. Various types of sensors are employed to measure heterogeneous parameters, such as velocity, viscosity, pressure, temperature, friction level, and vibration [31]. Chip-integrated information collection systems efficiently consolidate extensive functionalities onto a single compact chip, resulting in cost reduction, size diminution, and enhanced performance. Utilizing an FPGA network to govern the design module enables multi-channel data processing, which minimizes hardware requirements and enhances reconfigurability.

The NI LabVIEW FPGA Module extends the capabilities of the LabVIEW graphical development platform, making it particularly suitable for FPGA programming due to its explicit representation of parallelism and data streams. The modules responsible for multi-channel data collection and processing are executed within a dedicated NI cRIO device featuring an embedded FPGA. The design and implementation of the FPGA are carried out using the Project Explorer window. It involves creating FPGA target.vi and Host.vi, as well as configuring the relevant hardware before initiating the project implementation. The Target.vi serves as the FPGA target, accessing the desired number of inputs and selecting the input ports of the block. It determines the data type, memory size, and procedures for data reading and writing. Once compiled, it generates the bit file, which can be loaded into Host.vi. Fig. 3 illustrates the six-step FPGA design process, culminating in the transference of generated files to the FPGA for test verification in the final stage of the design.

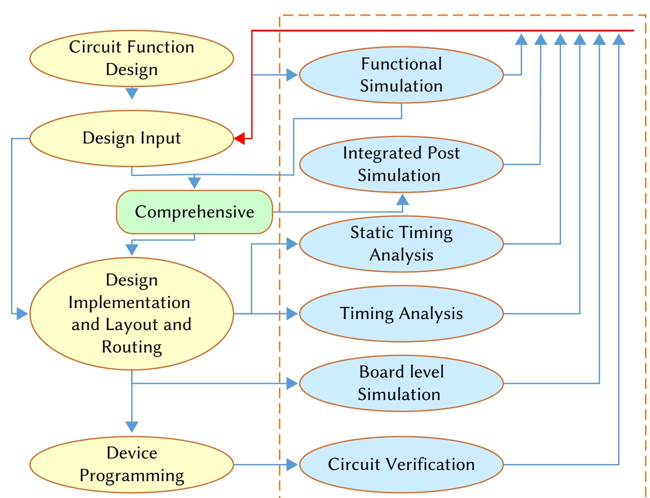


Fig. 3. FPGA design process.

C. Realization Principle of Geomagnetic Signal Acquisition

Geomagnetic detection represents an innovative vehicle detection technology involving the sensing of the magnetic field within the geomagnetic field using an anisotropic magneto-resistive sensor to ascertain the vehicle's condition. Currently, geomagnetic vehicle detectors predominantly rely on wireless transmission, offering advantages such as high detection accuracy, stability, reliability, and ease of installation and maintenance, making them highly sought after in the market. Integrating WSN and geomagnetic sensors, the traffic flow acquisition system collects vehicle induction data through these sensors. Users can access real-time road traffic flow information in the background through centralized management. Geomagnetic sensor detection technology is widely recognized as one of the most effective traffic data collection methods.

The presence of ferromagnetic substances within the vehicle influences the geomagnetic signal in the surrounding area, causing a distortion in the earth's magnetic field lines [32], [33]. When a vehicle passes near the sensor of the vehicle detector, the sensor sensitively perceives the signal change and extracts relevant information about the detected target through signal analysis. The WSN-based geomagnetic signal acquisition system offers ease in construction and maintenance, enabling real-time monitoring of traffic flow conditions [34]-[36]. Within geomagnetic signal acquisition, the anisotropic magnetoresistance effect demonstrates directionality. Equation (1) describes the relationship between the magnitude of the resistance value of a metal with anomalous reluctance effects between the bias current and the direction of the magnetic field.

$$R(\theta) = R1\, sin\, 2\,\theta + R2\, sin\, 2\,\theta \tag{1}$$

In Equation (1), θ stands for the angle between the magnetic field and the current direction; $R1$ and $R2$ represent the resistance values of the metal when the magnetic field direction and the current direction are parallel and perpendicular to each other, respectively. Fig. 4 illustrates the principle of the magnetoresistance effect.

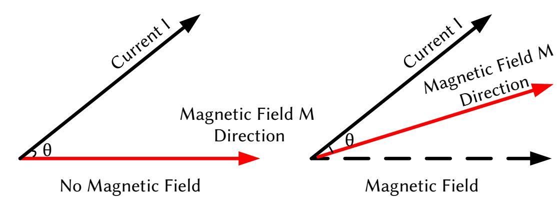


Fig 4. Principle of the magnetoresistance effect.

The geomagnetic signal collector is employed to capture geomagnetic signals and transform them into switch signals, indicating the presence or absence of a vehicle. These signals are then transmitted to the geomagnetic signal processor. Vehicle induction data is conveyed through the multi-hop transmission mechanism of the WSN, and the geomagnetic sensor is utilized for collecting the vehicle induction data. Real-time road traffic flow information is accessible to users through centralized background management. Compared with traditional traffic flow collection systems, the WSN-based traffic flow monitoring system is advantageous for its ease of construction and maintenance, facilitating real-time monitoring of traffic flow conditions. Fig. 5 illustrates the network topology of the geomagnetic data acquisition system.

The geomagnetic data acquisition system is built around an optimized single-chip microcomputer (SCM), serving as the hardware core. Discrete components are replaced with integrated circuit chips to enhance the operational reliability of the system. The core is the SCM P89C668, and its peripheral devices are configured to form a comprehensive hardware structure. This structure includes a liquid crystal display, keyboard, communication interface, non-volatile data memory, 16-bit A/D conversion, clock chip, GPS receiving module, and wireless modem. The communication interface employs the MAX202 chip for TTL and RS232 level conversion. On the transmitter side, the interface connects to the data acquisition unit, collects data, processes it, and transmits it to a modem via a microcontroller. At the receiving end, the port receives remote data through the modem. Subsequently, the SCM transmits the processed data to the computer. While the actual local geomagnetic field may vary with the environment, positional differences and surrounding structures have minimal impact on the magnetic field length in the short range and can be disregarded. Assessing the geomagnetic field distribution is necessary for determining variations in magnetic field length over long distances. Disparities between the calibration environment and operating environments have the most significant effect on the true local geomagnetic field. Over the long term, the geomagnetic vehicle detector industry is expected to continue improving, with signaling and parking management serving as key drivers for future development.

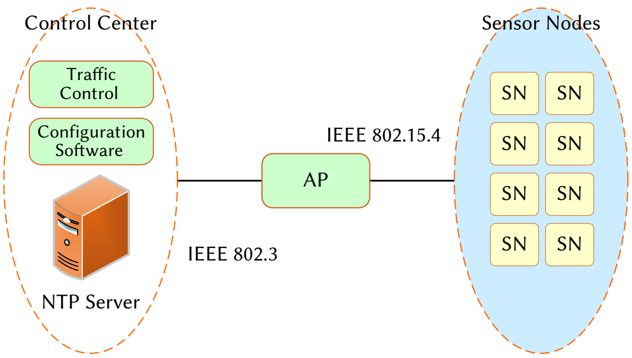


Fig 5. Network topology of the geomagnetic data acquisition system.

D. Adaptive Measurement of the Traffic Real-Time Signal Acquisition System

Transportation control plays a pivotal role as a technological tool in alleviating traffic congestion, managing traffic volumes, and reducing emissions. Its advancement is closely intertwined with progress in system science, information technology, and computer technology. The self-adaptive control system stands out by adjusting signal timing parameters in real-time, aligning with the manager's control objectives and the characteristics of intersection transportation flow. In comparison to timing and driving control methods, this system more effectively utilizes the entire road network's throughput, enhancing transportation efficiency. Current transportation management systems, employing inductive loop detectors and other sensing devices, are constrained in the extent of transportation information they collect. However, with the ongoing development of wireless communication technologies and vehicle-to-vehicle/vehicle-to-infrastructure systems (referred to as vehicle-to-everything), the optimization of urban transportation networks through the collaboration of traffic signal control and driving behavior regulation has become feasible [37]. This study introduces an optimization method that adjusts the vertical position of the oscilloscope over time using classic fuzzy control theory based on the measured traffic signal amplitude. This approach facilitates adaptive signal measurement.

The fundamental concept behind fuzzy control is to employ a computer to replicate human control experiences, often conveyed through language using fuzzy control rules [38], [39]. The primary factor contributing to the substantial success of the fuzzy controller (FC) is its rule-based nature. It directly applies language-based control rules and does not necessitate the development of a precise mathematical model of the controlled object during the design phase. Consequently, its control mechanism and strategy are easily comprehensible and accessible.

Firstly, the fuzzy set is defined. Given a domain of discourse, the mapping from U to the unit interval [0, 1] can be referred to as a fuzzy set on U, which can be expressed as in Equation (2).

$$\mu_A: U \to [0,1] \tag{2}$$

The membership function of each fuzzy set A on U is $\mu_A(u_i)$. A can be expressed as Equation (3).

$$\sum_{i=1}^{n} \mu_A(u_i)/u_i \tag{3}$$

Equation (4) indicates the arbitrary fuzzy set of U.

$$A = \int_U \mu_A(u)/\,u \tag{4}$$

Several membership functions commonly used in FCs are described as the followings.

(1) Trapezoidal membership function in Equation (5) and (6):

$$f(x,a,b,c,d) = \begin{cases} \frac{x\cdot a}{b\cdot a} & a \le x \le b \\ 1 & b \le x \le c \\ \frac{d\cdot x}{d\cdot c} & c \le x \le d \\ 0 & x < a, x > d \end{cases} \tag{5}$$

$$a \le b \le c \le d \tag{6}$$

(2) Triangular membership function in Equation (7) and (8):

$$f(x,a,b,c) = \begin{cases} \frac{x\cdot a}{b\cdot a} & a \le x \le b \\ \frac{c\cdot x}{c\cdot b} & b \le x \le c \\ 0 & x < a, x > c \end{cases} \tag{7}$$

$$a \le b \le c \tag{8}$$

(3) Gaussian membership function in Equation (9):

$$f(x,\sigma,c) = e^{-\frac{(x-c)^2}{2a^2}} \tag{9}$$

where c refers to the midpoint position of the Gaussian function, and the width of the Gaussian function depends on the value of σ.

This study leverages the concepts and techniques of fuzzy control to optimize the vertical positioning of the oscilloscope, enabling adaptive measurement of traffic signals to suit varying signal amplitudes. Initially, a set of fuzzy sets is defined, encompassing categories like "small," "medium," and "large" to signify distinct signal amplitude ranges. Subsequently, a corresponding collection of fuzzy rules is devised for each fuzzy set, delineating how adjustments to the oscilloscope's vertical position should be enacted in response to signal amplitudes within specific ranges. These rules can be formulated based on domain experts' insights, accrued knowledge, and signal acquisition requisites. The activation level of each fuzzy set is then ascertained through a fuzzy inference process, which hinges on real signal amplitude values. This inference entails employing fuzzy rules that map actual signal amplitudes to the activation degree of the corresponding fuzzy set. Lastly, the fuzzy output derived from the fuzzy reasoning process is transformed into a definitive control operation, specifically the manipulation of the oscilloscope's vertical position. This step may involve employing defuzzification techniques inherent to fuzzy control, such as using the average or maximum value of the fuzzy output as the ultimate control operation.

Based on the above analysis, this study formulates an effective digital model. The input to the FC consists of the first and second significant digits of the signal amplitude. The effective digit denoted as $f(x)$ of the signal amplitude can be expressed as Equation (10).

$$f(x) = \begin{cases} 10f_1(x) + f_2(x) & f_1(x) = 1 \\ f_1(x) + 0.1f_2(x) & other \end{cases} \tag{10}$$

Table I lists the effective figures obtained through function calculation.

TABLE I. Relationship Between Input Effective Figures and Signal Amplitude

Serial number	Signal amplitude (v)	Effective number
1	35.4	3.5
2	1.64	16
3	0.893	9.0
4	0.430	4.3
5	0.133	1 3
6	0.068	6.8

Fig. 6 reveals the relationship between the digital gear and the significant digits of the signal amplitude.

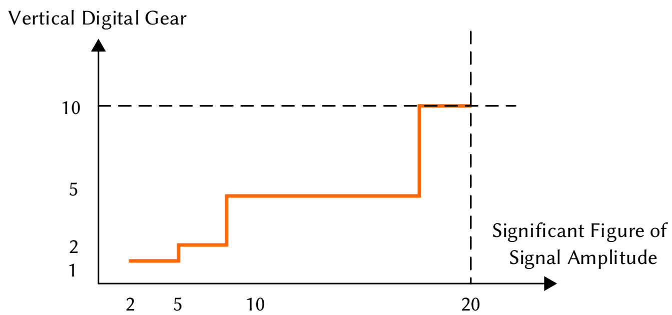


Fig 6. Relationship between the digital gear and the effective figure of the signal amplitude.

This study employs a set of representative signal parameter test data to assess and validate the program's output results. The computed results are displayed in Table 2, clearly demonstrating that the program has successfully achieved its intended functionality. This outcome confirms the feasibility and effectiveness of the fuzzy control algorithm.

TABLE II. Relationship Between Signal Amplitude and Vertical Scale

Signal amplitude (V)	Digital gear	Range gear (V)	Vertical gear (V/div)	Range (V)
0.010	5	0.001	0.005	0.02
0.016	5	0.001	0.005	0.02
0.4	1	0.1	0.1	0.4
1.63	5	0.1	0.5	2
1.92	10	0.1	1	4
2.5	1	1	1	4
7.22	2	1	2	8
22.5	1	10	10	40

IV. Experimental Design and Performance Evaluation

A. Experimental Materials

The principal aim of this experiment is to empirically validate the performance and efficacy of the proposed system. A comprehensive set of experiments has been carefully designed to scrutinize the system's performance across diverse scenarios, with a particular focus on temporal aspects related to acquisition tasks and storage velocity. Throughout the experimental process, the time required for signal transmission and code execution has been emulated to reflect real-world conditions. The system's storage speed has been rigorously assessed using a finely calibrated oscilloscope set to a resolution of 1ms/div to minimize potential testing inaccuracies. A meticulous approach has been adopted, synthesizing the final performance evaluation results by averaging outcomes from multiple experimental trials.

The experimental data used here consists of authentic signal data, including information about acquisition task duration, storage velocity, and the system's relative error. This empirical data serves a dual purpose: conducting a comprehensive assessment of the system's performance across varied scenarios and substantiating the system's dependability and precision.

B. Experimental Environment

The experiment is conducted within the premises of a high-speed railroad technical test station, replicating a real-world testing environment. In this setting, the test station is tasked with capturing equipment signals while potentially facing instances of high-amplitude transient pulse interference. These challenges serve as rigorous tests for evaluating the system's performance robustness and stability.

C. Parameters Setting

During the experimental phase, the oscilloscope's sampling rate is set at 1GSa/s to ensure an ample number of sampling points for proficient signal acquisition. In alignment with distinct testing scenarios, the system's adaptive parameters are systematically tuned through diffusion control, stepping control, and the dichotomous search algorithm, thus achieving the objective of adaptive measurement. Simultaneously, an assessment of the system's maximum relative errors is conducted across varying vertical resolutions using multiple sets of experimental data, facilitating an evaluation of the system's measurement accuracy and stability.

D. Performance Evaluation

1. System Verification and Index Analysis

The time required for the test system to complete a data collection task includes the time consumed during signal transmission to the PC and the time spent executing specific code instructions (adaptive adjustment). The system's storage speed is assessed at the oscilloscope's 1ms/div setting to prevent test inaccuracies, and the average of multiple test results is considered the final outcome. The test outcomes are depicted in Fig. 7. It is observable that the system's individual acquisition time increases as the oscilloscope's sampling points increase. The system's peak storage speed can reach 200KB/s when the number of sampling points reaches 1M. In the research conducted by Jin and Ma (2019) [40], a Constrained Markov Decision Process model is utilized to depict agent decisions, thereby optimizing multiple strategic objectives. The outcomes of their study closely align with the results presented in this study.

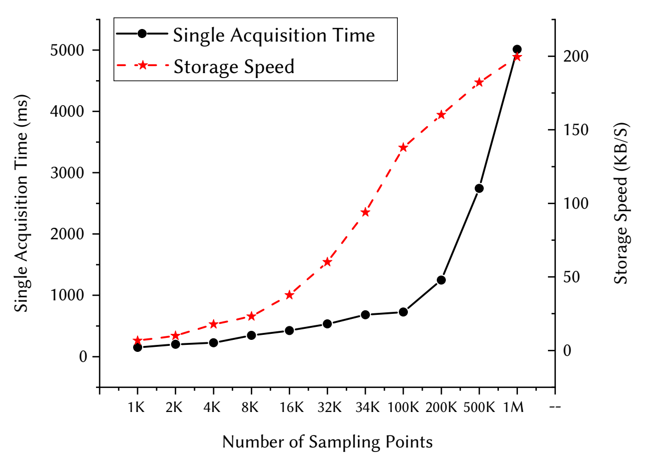


Fig 7. Relationship between collection points and single collection time.

2. Analysis of System Errors

Currently, high-speed railway technology test stations require digital recorders to capture equipment signals. Typically, the test subject installs the equipment at a designated test location, and personnel departs the area after configuring various parameters. The test equipment must achieve long-term continuous acquisition and automated signal storage. Given the presence of high-amplitude transient pulse interferences on railways, measurement parameters must be dynamically adjusted based on emergency conditions to ensure measurement accuracy and adaptive measurement of burst signals. It is recognized that digital oscilloscopes can achieve sampling rates of up to 1GSa/s. In this context, sufficient sampling points can be captured; hence, the test system's sampling rate is set to 1GSa/s. For signals being tested within the oscilloscope's current testing apparatus, adaptive parameters of the testing device are adjusted through diffusion control. Conversely, when the signal to be tested is not within the current testing device, the adaptive parameters are established using step control and a binary search algorithm. The maximum relative error of the system across multiple test configurations is computed through numerous sets of experiments. Table 3 presents a summary of the specific outcomes. After several tests, it is apparent that the maximum relative error of the test outcomes remains below ± 2.7%. This system attains more precise test outcomes in comparison to existing testing equipment, rendering it better suited for application in the collection of traffic signals. Chen and Zhang (2022) [41] developed a traffic flow prediction model utilizing the Deep Belief Network (DBN) algorithm. By collecting and pre-processing historical traffic flow data and incorporating multiple Restricted Boltzmann Machines in the DBN, they established a generative model for training. This procedure adds further validation to the reliability of the results presented in this study.

TABLE III. Maximum Relative Error at Different Vertical Gears

Vertical gear (V/div)	Minimum voltage increment (mV)	Current range (V)	Measuring range (V)	Maximum relative error (±)
0.002	0.08	0.008	0.0035~0.0075	2.2%
0.005	0.2	0.02	0.0075~0.018	2.6%
0.01	0.4	0.04	0.018~0.036	2.2%
0.02	0.8	0.08	0.036~0.075	2.7%
0.05	2	0.2	0.075~0.176	2.3%
0.1	4	0.4	0.176~0.362	2.2%
0.2	8	0.8	0.362~0.745	2.6%
0.5	20	2	0.745~1.73	2.3%
1	40	4	1.73~3.60	2.2%
2	80	8	3.60~7.45	2.2%
5	200	20	7.45~17.3	2.7%

V. Discussion

This study successfully designed and implemented a traffic electronic information signal collection system based on IoT technology and artificial intelligence. By leveraging FPGA technology, the hardware circuitry for the high-speed signal acquisition control core was developed, enabling wireless monitoring of signal collection. This innovative design achieves wireless monitoring of signal collection and imparts efficient data processing capabilities to the system, ensuring stable operation even in complex traffic environments. As detailed in this study, the time signal acquisition system encompasses multiple modules: communication, acquisition, storage, adaptive measurement, and signal analysis. The magnetic field acquisition module stands out for its effective collection of magnetic field signals and their conversion into switch signals, indicating the presence or absence of vehicles. This design enhances not only the practicality of the system but also its adaptability to dynamically changing traffic conditions.

The experimental results demonstrate the excellent performance of the system designed in this study in data storage and processing, achieving a significant peak storage speed of 200KB/s. Considering the substantial volume of data the system needs to handle, this achievement undoubtedly showcases the system's outstanding capabilities. In a series of tests, the maximum relative error of the obtained results ranged from 2.2% to 2.7%, further emphasizing the consistency and reliability of the measurements. Compared to existing testing devices, the system designed in this study exhibits higher accuracy in test results, rendering it more suitable for collecting traffic signals. It is noteworthy that the designed system can collect and process traffic information in real-time and be able to self-adjust and optimize based on changes in traffic conditions. This feature grants the designed system strong adaptability and flexibility, allowing it to maximize utility in various traffic environments.

In conclusion, the incorporation of AI technology, specifically fuzzy control, into traffic signal acquisition systems has proven to be a valuable strategy for enhancing system precision. A comprehensive literature review conducted by Ranyal et al. (2022) on road condition monitoring, spanning from 2017 to 2022, explored various approaches, innovative contributions, and limitations in the field. The authors underscored the importance of smart sensors and data acquisition platforms while addressing challenges in AI technology development. Their analysis provided valuable insights outlined directions and perspectives for future research in the realm of road condition monitoring [42]. In summary, a growing body of evidence suggests that the integration of AI into intelligent transportation and smart cities holds the potential to significantly optimize road conditions, thereby advancing the overall transportation system.

The algorithm proposed in this study facilitates the development of the hardware circuit for the high-speed signal acquisition control core, employing FPGA technology. This innovative design enables the system to achieve wireless monitoring of signal acquisition and demonstrates efficient data processing capabilities. Distinguished from other state-of-the-art algorithms, the proposed algorithm greatly emphasizes hardware-level optimization and innovation, thereby enhancing the system's overall performance and stability. The proposed algorithm incorporates a geomagnetic collection module that effectively gathers geomagnetic signals, transforming them into switch signals indicating the presence or absence of vehicles. This design allows the system to dynamically adapt to changing traffic environments, thereby increasing its practicality. In contrast to other advanced algorithms that typically rely on traditional sensors or cameras for data collection, the algorithm presented here is characterized by its innovation and adaptability. The study successfully achieves the collection and sharing of various traffic information, providing robust information support for traffic control and safe operations. However, the system still possesses certain limitations, such as potential performance bottlenecks when dealing with large-scale complex data. Future research directions will focus on optimizing system performance and enhancing data processing capabilities to achieve more efficient and accurate traffic information collection and processing.

VI. Conclusion

1. Research Contribution

Deploying various intelligent technologies and equipment to advance the digitization, interconnection, and intelligence of transportation defines intelligent transportation. Network connectivity emerges as a critical application integral to the evolution of intelligent transportation, with the IoT playing a pivotal role in seamlessly connecting all components of transportation. This technology has the potential to revolutionize the traffic industry by optimizing the efficient utilization and management of traffic information, reinforcing traffic oversight, and elevating traffic services. It encompasses information collection, policy control, output execution, data transmission, and communication between subsystems. Experimental results demonstrate that, with a sampling point count of 1M, the system achieves a maximum storage speed of up to 200KB/s. Throughout numerous tests, the peak relative error in test outcomes ranges from 2.7% to a mere 2.2%. Notably, the test results from this system showcase enhanced accuracy compared to existing testing equipment, making it more suitable for traffic signal acquisition applications. This study affirms that the real-time signal acquisition system within the IoT environment can promptly gather, analyze, and process collected signals. An intelligent traffic signal optimization control system is established through the integration of the intelligent collection system and the comprehensive analysis of big data.

2. Future Works and Research Limitations

This study introduces the design and experimentation of a traffic electronic information signal acquisition system based on IoT and AI technologies. However, several limitations should be acknowledged. The experiments primarily focus on validating the performance of the signal acquisition system, particularly regarding storage speed and relative error. Nevertheless, the experimental scenarios and datasets are relatively limited, potentially not capturing the entirety of real-world traffic situations and intricacies. The study predominantly emphasizes the design and performance evaluation of signal acquisition systems without comprehensive integration within the broader context of ITSs. Actual ITS scenarios often involve additional factors such as traffic flow management and incident prediction. Future research endeavors could explore the integration of this traffic electronic information signal acquisition system with other ITS components, such as traffic flow management and vehicle behavior prediction. This integration could lead to more comprehensive traffic management and optimization, addressing a broader spectrum of challenges in the field. Although this study successfully attains a noteworthy peak storage speed of 200KB/s, the system can face performance bottlenecks when confronted with extensive and intricate datasets. The design and optimization strategies are primarily tailored to specific traffic scenarios, necessitating further validation of the system's adaptability to diverse contexts. Subsequent research endeavors should prioritize augmenting the algorithm's generalization capabilities, thereby enabling it to adeptly accommodate a broader spectrum of traffic scenarios.

References

[1] B. Jan, H. Farman, M. Khan, M. Talha, & I. U. Din, "Designing a smart transportation system: An internet of things and big data approach," *IEEE Wireless Communications*, vol. 26, no. 4, pp. 73-79, 2019.

[2] F. Al-Turjman, J. P. Lemayian., "Intelligence, security, and vehicular sensor networks in internet of things (IoT)-enabled smart-cities: An overview," *Computers & Electrical Engineering*, vol. 87, p. 106776, 2020.

[3] M. Babar, F. Arif., "Real-time data processing scheme using big data analytics in internet of things based smart transportation environment," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 10, pp. 4167-4177, 2019.

[4] Y. Zhang, H. Huang, L. X. Yang, Y. Xiang, & M. Li, "Serious challenges and potential solutions for the industrial Internet of Things with edge intelligence," *IEEE Network*, vol. 33, no. 5, pp. 41-45, 2019.

[5] D. Jiang, "The construction of smart city information system based on the Internet of Things and cloud computing," *Computer Communications*, vol. 150, pp. 158-166, 2020.

[6] A. Ahmad, S. Din, A. Paul, G. Jeon, M. Aloqaily, & M. Ahmad, "Real-time route planning and data dissemination for urban scenarios using the Internet of Things," *IEEE Wireless Communications*, vol. 26, no. 6, pp. 50-55, 2019.

[7] Y. Ding, M. Jin, S. Li, & D. Feng, "Smart logistics based on the internet of things technology: an overview," *International Journal of Logistics Research and Applications*, vol. 24, no. 4, pp. 323-345, 2021.

[8] J. Yang, J. Zhang, H. Wang, "Urban traffic control in software defined internet of things via a multi-agent deep reinforcement learning approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 6, pp. 3742-3754, 2020.

[9] R. W. Liu, J. Nie, S. Garg, Z. Xiong, Y. Zhang, & M. S. Hossain, "Data-driven trajectory quality improvement for promoting intelligent vessel traffic services in 6G-enabled maritime IoT systems," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5374-5385, 2020.

[10] H. Tran-Dang, N. Krommenacker, P. Charpentier, & D. S. Kim, "The Internet of Things for logistics: Perspectives, application review, and challenges," *IETE Technical Review*, vol. 39, no. 1, pp. 93-121, 2022.

[11] D. Tabernik, D. Skočaj, "Deep learning for large-scale traffic-sign detection and recognition," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 4, pp. 1427-1440, 2019.

[12] D. Temel, M. H. Chen, G. AlRegib, "Traffic sign detection under challenging conditions: A deeper look into performance variations and spectral characteristics," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 9, pp. 3663-3673, 2019.

[13] F. Zhu, Y. Lv, Y. Chen, X. Wang, G. Xiong, & F. Y. Wang, "Parallel transportation systems: Toward IoT-enabled smart urban traffic control and management," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 10, pp. 4063-4071, 2019.

[14] M. Akhtar, S. Moridpour, "A review of traffic congestion prediction using artificial intelligence," *Journal of Advanced Transportation*, vol. 2021, pp. 1-18, 2021.

[15] R. Abduljabbar, H. Dia, S. Liyanage, & S. A. Bagloee, "Applications of artificial intelligence in transport: An overview," *Sustainability*, vol. 11, no. 1, p. 189, 2019.

[16] J. Wu, X. Wang, Y. Dang, & Z. Lv, "Digital twins and artificial intelligence in transportation infrastructure: Classification, application, and future research directions," *Computers and Electrical Engineering*, vol. 101, p. 107983, 2022.

[17] Z. Wang, Y. Ma, "Detection and recognition of stationary vehicles and seat belts in intelligent Internet of Things traffic management system," *Neural Computing and Applications*, pp. 1-10, 2022.

[18] D. Ushakov, E. Dudukalov, E. Kozlova, & K. Shatila, "The Internet of Things impact on smart public transportation," *Transportation Research Procedia*, vol. 63, pp. 2392-2400, 2022.

[19] S. Muthuramalingam, A. Bharathi, S. Rakesh Kumar, N. Gayathri, R. Sathiyaraj, & B. Balamurugan, "IoT based intelligent transportation system (IoT-ITS) for global perspective: A case study," *Internet of Things and Big Data Analytics for Smart Generation*, 2019, pp. 279-300.

[20] A. Avramović, D. Sluga, D. Tabernik, D. Skočaj, V. Stojnić, & N. Ilc, "Neural-network-based traffic sign detection and recognition in high-definition images using region focusing and parallelization," *IEEE Access*, vol. 8, pp. 189855-189868, 2020..

[21] M. Koresh, J. Deva. "Computer vision based traffic sign sensing for smart transport," *Journal of Innovative Image Processing (JIIP)*, vol. 26, no. 4, pp. 11-19, 2019.

[22] D. De Silva, S. Sierla, D. Alahakoon, E. Osipov, X. Yu, & V. Vyatkin, "Toward intelligent industrial informatics: A review of current developments and future directions of artificial intelligence in industrial applications," *IEEE Industrial Electronics Magazine*, vol. 14, no. 2, pp. 57-72, 2020.

[23] S. Xu, Y. Qian, R. Q. Hu, "Data-driven edge intelligence for robust network anomaly detection," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 3, pp. 1481-1492, 2019.

[24] S. D. Mou, Z. X. Xiong, Y. Y. Tian, "Intelligent traffic control system based on cloud computing and big data mining," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 12, pp. 6583-6592, 2019.

[25] H. Guan, W. Yan, Y. Yu, L. Zhong, & D. Li, "Robust traffic-sign detection and classification using mobile LiDAR data with digital images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 5, pp. 1715-1724, 2018.

[26] Y. Tian, J. Gelernter, X. Wang, J. Li, & Y. Yu, "Traffic sign detection using a multi-scale recurrent attention network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 12, pp. 4466-4475, 2019.

[27] W. Li, Z. Su, R. Li, K. Zhang, & Y. Wang, "Blockchain-based data security for artificial intelligence applications in 6G networks," *IEEE Network*, vol. 34, no. 6, pp. 31-37, 2020.

[28] Z. Liang, J. Shao, D. Zhang, & L. Gao, "Traffic sign detection and recognition based on pyramidal convolutional networks," *Neural Computing and Applications*, vol. 32, no. 11, pp. 6533-6543, 2020.

[29] A. Boukerche, Y. Tao, P. Sun., "Artificial intelligence-based vehicular traffic flow prediction methods for supporting intelligent transportation systems," *Computer Networks*, vol. 182, p. 107484, 2020.

[30] X. Tao, Y. Peng, F. Zhao, P. Zhao, & Y. Wang, "A parallel algorithm for network traffic anomaly detection based on Isolation Forest," *International Journal of Distributed Sensor Networks*, vol. 14, no. 11, p. 1550147718814471, 2018.

[31] Y. Ma, Z. Wang, H. Yang, & L. Yang, "Artificial intelligence applications in the development of autonomous vehicles: a survey," *IEEE/CAA Journal of Automatica Sinica*, vol. 7, no. 2, pp. 315-329, 2020.

[32] R. A. Gonzalez, R. E. Ferro, D. Liberona, "Government and governance in intelligent cities, smart transportation study case in Bogotá Colombia," *Ain Shams Engineering Journal*, vol. 11, no. 1, pp. 25-34, 2020.

[33] W. Shi, M. Zhang, R. Zhang, S. Chen, & Z. Zhan, "Change detection based on artificial intelligence: State-of-the-art and challenges," *Remote Sensing*, vol. 12, no. 10, p. 1688, 2020.

[34] J. Yang, C. Wang, B. Jiang, & Q. Meng, "Visual perception enabled industry intelligence: state of the art, challenges and prospects," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 2204-2219, 2020.

[35] Z. Tong, F. Ye, M. Yan, H. Liu, & S. Basodi, "A survey on algorithms for intelligent computing and smart city applications," *Big Data Mining and Analytics*, vol. 4, no. 3, pp. 155-172, 2021.

[36] D. Impedovo, G. Pirlo. "Artificial intelligence applications to smart city and smart enterprise," *Applied Sciences*, vol. 10, no. 8, p. 2944, 2020.

[37] Y. Hou, Q. Li, C. Zhang, G. Lu, Z. Ye, Y. Chen, L. B. Wang, & D. Cao, "The state-of-the-art review on applications of intrusive sensing, image processing techniques, and machine learning methods in pavement monitoring and analysis," *Engineering*, vol. 7, no. 6, pp. 845-856, 2021.

[38] S. Latif, Z. e Huma, S. S. Jamal, F. Ahmed, J. Ahmad, A. Zahid, K. Dashtipour, M. U. Aftab, M. Ahmad, & Q. H. Abbasi, "Intrusion detection framework for the internet of things using a dense random neural network," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 9, pp. 6435-6444, 2021.

[39] S. Mishra, D. Bhattacharya, A. Gupta, "Congestion adaptive traffic light control and notification architecture using google maps APIs," *Data*, vol. 3, no. 4, p. 67, 2018.

[40] J. Jin, X. Ma, "A multi-objective agent-based control approach with application in intelligent traffic signal system," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 10, pp. 3900-3912, 2019.

[41] G. Chen, J. Zhang, "Applying Artificial Intelligence and Deep Belief Network to predict traffic congestion evacuation performance in smart cities," *Applied Soft Computing*, vol. 121, p. 108692, 2022.

[42] E. Ranyal, A. Sadhu, K. Jain, "Road condition monitoring using smart sensing and artificial intelligence: A review," *Sensors*, vol. 22, no. 8, p. 3044, 2022.

Hongling Wang

Hongling Wang was born in Yantai, Shan Dong P.R. China, in 1976. She received the Master degree from Harbin Engineering University, P.R. China. Now, she works in College of Civil Engineering, Harbin University. Her research direction is communication and information processing. E-mail: wanghongling@hrbu.edu.cn

LA UNIVERSIDAD
EN INTERNET

Rectorado
Avenida de la Paz, 137
26006 Logroño (La Rioja)
t (+34) 941 21 02 11

www.unir.net