

# Journal of Advances in Engineering and Technology

Editor-in-Chief:

**Prof. Dr. José Pereira**

*Universidade de Lisboa, Portugal*

Copyright © 2025. ASIA PACIFIC SCIENCE PUBLICATIONS  
COMPANY LIMITED. Complimentary Copy.

## Journal of Advances in Engineering and Technology

*Journal of Advances in Engineering and Technology (JAET)* is an international, peer-reviewed and open access journal which publishes original articles, reviews, short communications, case studies and letters in the field of electronic research and application. It covers mainly but not limits to the following areas:

- Civil Engineering
- Mechanical Engineering
- Electrical Engineering
- Chemical Engineering
- Aerospace Engineering
- Computer Science and Engineering
- Environmental Engineering
- Materials Science and Engineering
- Biomedical Engineering
- Energy Engineering
- Robotics and Automation

### About Publisher:

Asia Pacific Science Press (APSP) is a swiftly expanding publisher of peer-reviewed and open-access journals, strategically located in Hong Kong. As a reliable and esteemed corporation, APSP is dedicated to promoting and serving a wide array of subject areas, ultimately contributing to the betterment of humanity. By disseminating knowledge to a global community of scholars, practitioners, researchers, and students, we strive to establish ourselves as the world's leading independent academic and professional publisher.

Submission instructions: You can submit your manuscript through the official website ([www.apspublisher.com](http://www.apspublisher.com)) or email ([editor.jaet@apspublisher.com](mailto:editor.jaet@apspublisher.com)), All manuscripts will go through a rapid peer review and production, making the process of publishing simpler and more efficient.

### Publisher Headquarter

Room 03, 7th Floor, Block B, Tuen Mun Industrial Centre, 2 New Ping Street, Tuen Mun, Hong Kong, China  
Website : [www.apspublisher.com](http://www.apspublisher.com)  
Email : [info@apspublisher.com](mailto:info@apspublisher.com)

### Fujian Province Office, China

603-1, 6th Floor, Building B20, Chengyi North Street, Software Park, Jimei District, Xiamen City, Fujian Province, China  
Website : <https://ojs.apspublisher.com/index.php/amit>  
Email : [amit@apspublisher.com](mailto:amit@apspublisher.com)

## Table of Contents

- 1    MADENet: Explainable AI-Driven Bike-Sharing Demand Forecasting for Sustainable Urban Mobility**  
*Jianbo Huang, Yanbin Zheng, Lihua Lan, Jia Chen*
  
- 18   Low-Cost CCTV Repurposing for Sustainable Parking Management: A Non-AI Computer Vision Case Study**  
*Namya Kamboj, Kongwen Zhang*
  
- 30   A Review of Research on the Development and Application of Modular Steel Structures**  
*Chunyang Hu, Siqi Li, Yibing Wang, Zijing Song, Chunhao Wang*
  
- 43   Public Data Access and AI Adoption for Sustainable Digital Transformation: Evidence from China**  
*Wei Zhao*
  
- 52   Indoor Pathfinding with the A Algorithm: A Cross-Platform Mobile Implementation Case**  
*Kongwen Zhang, Boris Massesa, Jingwei Gao*
  
- 59   Research on Optimization Pathways for Reverse Logistics Networks Based on Ecological Civilization**  
*Jinzhao Song, Xiaofeng Zhang*
  
- 67   Construction of SnS<sub>2</sub>/TiO<sub>2</sub> Heterojunction and Study of Its Photocatalytic Performance**  
*Xiaofeng Zhang, Jinzhao Song*
  
- 75   Vocabulary Semantic Similarity Calculation in Natural Language Processing**  
*Huixiang Xiao, Kaige Zheng, Xiangyu Li*
  
- 86   Image Classification in Coal Production Using Deep Neural Networks: A Comprehensive Benchmarking Study**  
*Wenmi Chai, Zhiyao Yang, Rui Zhao, Qian Xiang, Xinxin Niu, Ling Liang*
  
- 98   Coupled Topology–Material Design Enables Manufacturable Lattice Metamaterials with Tailored Mechanics**  
*Xinru Li, Yifei Zhou, Danyang Qiu, Ningping Zhan, Jianxiang Qiao, Yaoxin Huang*



# MADENet: Explainable AI-Driven Bike-Sharing Demand Forecasting for Sustainable Urban Mobility

Jianbo Huang\*, Yanbin Zheng, Lihua Lan, Jia Chen

School of Computer Application, Guilin University of Technology, Guilin, 541000, China

\*Corresponding author: Jianbo Huang, [lucassivan@163.com](mailto:lucassivan@163.com)

**Copyright:** 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

**Abstract:** With rapid urbanization and increasing motorization, bike-sharing systems have emerged as sustainable solutions for urban “last-mile” connectivity. However, existing demand forecasting approaches face a critical trade-off between predictive accuracy and operational interpretability, limiting their practical deployment in municipal decision-making contexts where both reliable predictions and transparent insights are essential. The study proposes MADENet, a novel neural architecture that systematically addresses this accuracy-interpretability challenge. The framework integrates three key innovations: multi-head attention mechanisms to dynamically capture cross-regional demand dependencies and temporal periodicity patterns; adaptive dropout with early-stopping regularization to mitigate overfitting in high-dimensional spatio-temporal scenarios; and multilayer perceptron components to model complex nonlinear interactions between heterogeneous external factors and urban mobility patterns. Experimental evaluation demonstrates MADENet’s superior performance, achieving 95.1% prediction accuracy ( $R^2=0.9515$ ,  $MAE=0.2320$ ) and outperforming 15 baseline algorithms with MAE improvements ranging from 7.7% to 70% across different algorithmic paradigms. Embedded SHAP and LIME explainable AI frameworks systematically identify hour-of-day, temperature, and humidity as dominant spatio-temporal drivers while quantifying their nonlinear interactions with demand patterns. These innovations provide transparent operational protocols for station layout optimization, dynamic fleet rebalancing, and evidence-based policy formulation, ultimately advancing data-driven governance of sustainable urban mobility systems through actionable insights that bridge algorithmic predictions with practical urban planning requirements.

**Keywords:** Bike-Sharing; Explainable AI; Smart City; Sustainable Transportation

**Published:** Oct 15, 2025

**DOI:** <https://doi.org/10.62177/jaet.v2i4.662>

## 1.Introduction

The rapid urbanization observed in recent decades has intensified challenges such as traffic congestion, environmental degradation, and strained public transportation systems<sup>[1]</sup>. A critical contributor to these issues is the overreliance on private vehicles, which account for approximately 11% of global carbon dioxide emissions<sup>[2]</sup> and impose severe economic burdens—for instance, traffic congestion alone costs China an estimated USD 35 billion annually<sup>[3]</sup>. In response to these dual crises of sustainability and urban efficiency, bike-sharing systems have emerged as a transformative mobility solution. Since their inception, these programs have experienced explosive global growth: active bike-sharing fleets surged from 700,000 to 2 million in the United States between 2013 and 2016, with over 2,000 systems now operating nearly 10 million bicycles

worldwide<sup>[4]</sup>. By providing affordable, eco-friendly “last-mile” connectivity, bike-sharing networks reduce reliance on private cars while alleviating congestion and emissions.

Accurate demand forecasting and bicycle allocation remain critical bottlenecks for optimizing bike-sharing system efficiency. While traditional time-series models such as ARIMA and linear regression<sup>[5-10]</sup> often fail to capture intricate spatio-temporal dependencies and heterogeneous external factors, recent deep learning approaches have improved predictive performance through temporal dependency modeling<sup>[11-12]</sup>. However, these solutions face a fundamental accuracy-interpretability trade-off: high-dimensional meteorological and urban infrastructure data exacerbate overfitting risks in sparse datasets, while the inherent opacity of neural networks undermines stakeholders’ trust in critical urban planning decisions where interpretability is essential—such as justifying infrastructure investments or optimizing fleet rebalancing strategies<sup>[13-15]</sup>.

Current literature reveals significant methodological gaps in bike-sharing demand forecasting, where existing research typically addresses predictive accuracy and model interpretability as separate processes, resulting in suboptimal frameworks for urban mobility management. Contemporary deep learning approaches predominantly focus on maximizing prediction performance through increasingly complex architectures, while traditional interpretable models sacrifice accuracy for transparency, creating a fundamental dichotomy that fails to meet the dual requirements of municipal decision-making contexts. The unique challenges in bike-sharing demand prediction—dynamic spatio-temporal dependencies, complex meteorological-mobility interactions, and stringent requirements for both operational accuracy and transparent decision-making support—demand specialized methodological considerations that contemporary approaches inadequately address.

To address these limitations, this investigation introduces MADENet, a novel neural architecture synergistically combining multi-head attention mechanisms with adaptive dropout regularization and early-stopping protocols for enhanced bike-sharing demand forecasting. The framework systematically addresses the accuracy-interpretability trade-off by embedding explainable AI techniques directly into the prediction pipeline, enabling efficient capture of cross-regional demand correlations while mitigating overfitting without compromising computational efficiency.

The principal contributions encompass:

- **Novel Architecture Design:** Introduction of an innovative neural framework integrating multi-head attention mechanisms that dynamically capture cross-regional demand dependencies and temporal periodicity patterns, coupled with multilayer perceptron components and adaptive regularization strategies specifically tailored for spatio-temporal bike-sharing data;
- **Superior Predictive Performance:** Achieved exceptional forecasting accuracy (95.1% accuracy,  $R^2=0.9515$ ,  $MAE=0.2320$ ) outperforming 15 baseline algorithms with MAE improvements ranging from 7.7% to 70% across different algorithmic paradigms including deep learning, ensemble methods, and traditional statistical approaches;
- **Comprehensive Explainability Integration:** Embedded SHAP and LIME interpretability frameworks ensuring operational transparency for municipal decision-making applications, systematically identifying key spatio-temporal drivers including hour-of-day, temperature, and humidity while quantifying their nonlinear interactions with urban mobility patterns.

The remainder of this paper is organized as follows: Section 2 reviews existing methodologies in bike-sharing demand prediction and highlights unresolved challenges. Section 3 presents preliminary work, including data preprocessing and exploratory visualization analysis of temporal and environmental factors influencing bike-sharing demand. Section 4 details the MADENet architecture, covering its attention mechanisms and regularization strategies. Section 5 presents experimental results and comparative analyses, exploring interpretability outcomes via SHAP and LIME to connect algorithmic behavior with practical urban mobility strategies. Finally, Section 6 concludes by discussing implications for sustainable smart city governance and outlines future research directions.

## 2.Related Works

The rapid growth of bike-sharing systems as sustainable urban mobility solutions has intensified the need for accurate demand forecasting. Data-driven approaches, particularly machine learning, have propelled this field forward, yet existing frameworks still struggle to capture dynamic spatio-temporal dependencies while maintaining operational interpretability for practical deployment<sup>[5,13,30]</sup>.

Early research predominantly relied on statistical methods such as time series analysis and linear regression<sup>[6-10]</sup>. Although

these approaches provided foundational in-sights, they often failed to capture nonlinear interactions arising from weather fluctuations, special events, and complex urban dynamics. Studies that integrated seasonal and weather factors<sup>[5-6,16]</sup> established valuable baselines but exhibited limited adaptability to real-world volatility. Sathishkumar et al.<sup>[7]</sup> incorporated weather and usage data into a Gradient Boosting Machine to predict hourly bike-sharing demand, yet its robustness to rare events remained unclear.

Machine learning models marked a paradigm shift by addressing nonlinear effects and temporal interactions. Random Forest and Gradient Boosting Machines (GBMs) produced promising results<sup>[28-29]</sup>, while Schnieder<sup>[18]</sup> demonstrated that temperature, distance, wind, and elevation accounted for 21–27% of potential e-bike usage. Hu et al.<sup>[19]</sup> employed a grid search-optimized XGBoost model for Washington rental data, but it did not adequately handle abrupt real-time demand shifts. Lee et al.<sup>[14]</sup> illustrated that incorporating air pollution, traffic, and socio-economic variables can enhance predictive performance, reinforcing the potential of diverse data sources.

Deep learning approaches have attempted to bridge gaps by jointly modeling spatial and temporal complexities. LSTMs outperformed GBMs for weekend demand<sup>[10]</sup>, and CNN-based models introduced spatial awareness through demand heatmaps<sup>[11]</sup>. However, rigid grid structures often clashed with organic urban layouts. Li et al.<sup>[17]</sup> introduced STG2Vec, an attention-based graph embedding model, to learn heterogeneous spatio-temporal patterns for improved demand prediction. Li et al.<sup>[20]</sup> developed a Spatial-Temporal Memory Network (STMN) to capture short-term spatio-temporal patterns more effectively.

Recent hybrid approaches have shown promising results in addressing specific operational challenges. Yu et al.<sup>[30]</sup> integrated SARIMA with LSTM to predict bicycle flows around metro stations, while Wang et al.<sup>[38]</sup> proposed a model-data dual-driven approach combining SARIMA and extended Long Short-Term Memory (xLSTM) networks, achieving high R-squared values (0.9928-0.9535) with 8% improvement over conventional LSTM. However, dual-component fusion introduces computational complexity and potential calibration challenges when adapting to different metropolitan contexts.

Graph Neural Networks (GNNs) emerged as a state-of-the-art paradigm by representing stations as graph nodes<sup>[12]</sup>. However, their static graph architectures struggled to adapt to rapid demand fluctuations due to weather changes. Liang et al.<sup>[22]</sup> proposed a Domain-Adversarial Multi-Relational Graph Neural Network (DA-MRGNN) that leverages multimodal transport data, mitigating negative transfer between modes.

Recent advances have focused on improving GNN architectures for bike-sharing applications. Behroozi et al.<sup>[36]</sup> proposed a gate graph convolutional neural network integrating trajectory, weather, and access data, though the framework may struggle with computational scalability due to dynamic graph topology changes. Qian et al.<sup>[37]</sup> developed CGA-STNet for dockless bike-sharing demand prediction, integrating multi-dimensional spatial features and time periodicity through Fourier transforms, achieving 16.3% MSE reduction over benchmark models but with limitations in handling long-term seasonal variations. Xiang et al.<sup>[41]</sup> combined dynamic time warping with spatio-temporal graph attention networks, using data-driven adjacency matrices and multi-scale temporal features, though computational complexity may limit scalability.

Recent research has explored sophisticated frameworks addressing both prediction accuracy and operational challenges. Guo et al.<sup>[39]</sup> developed an XGBoost-based three-stage prediction approach that addresses the gap between observed bike pickup/drop-off records and true user demand by incorporating unsatisfied demand from empty or full stations. While demonstrating superior performance using Citi Bike data from New York, the framework's computational complexity and dependency on historical patterns may limit real-time implementation and transferability across different systems.

Beyond predictive modeling, recent research has explored optimization-driven approaches. Shi et al.<sup>[40]</sup> introduced a generative-model-informed reinforcement learning approach for long-term inventory management in hybrid bike-sharing systems, utilizing a recurrent-attentive neural process (RANP) for demand prediction and a cooperative two-agent MDP framework for bike-e-bike allocation optimization. While demonstrating superior performance, the framework's complexity may pose challenges for real-time deployment. Giner Fabregat et al.<sup>[42]</sup> developed an intelligent optimization framework for Barcelona's Bicing system, combining clustering analysis and machine learning-based demand prediction with optimization algorithms for efficient rebalancing strategies.

Despite noteworthy gains in accuracy, hybrid and deep learning models often lack explainability, limiting their usefulness in municipal decision-making contexts <sup>[25–26]</sup>. The “black-box” nature of these complex models raised concerns regarding transparency and stakeholder trust. Explainable AI (XAI) methodologies have thus emerged to balance predictive power with interpretability <sup>[27]</sup>, offering mechanisms to unveil the decision logic of otherwise opaque models <sup>[25,26]</sup>. However, existing XAI applications in bike-sharing prediction rarely undergo thorough validation against empirical urban mobility patterns, underscoring the need for interpretable frameworks that align with real-world operational constraints <sup>[35]</sup>.

Despite significant advances in bike-sharing demand forecasting, critical gaps remain in achieving the optimal balance between predictive accuracy and operational interpretability. Current deep learning approaches are predominantly adapted from general time-series prediction without adequate consideration of urban mobility-specific requirements, such as real-time deployment constraints, cross-regional transferability, and municipal decision-making transparency. The relationship between model architectural complexity and explanation reliability represents a fundamental but understudied aspect, potentially leading to urban planning decisions based on explanations derived from models that may not adequately capture the nuanced spatio-temporal dynamics of bike-sharing systems. Most studies focus on algorithmic performance metrics without sufficient consideration of practical integration into municipal planning workflows, where computational efficiency, regulatory transparency, and stakeholder interpretability significantly influence real-world applicability.

Our work addresses these gaps by introducing MADENet, a novel neural architecture that integrates multi-head attention mechanisms with adaptive dropout regularization, specifically designed for bike-sharing demand forecasting. By embedding explainable AI frameworks (SHAP and LIME) directly into the predictive pipeline, we provide both theoretical foundations and empirical validation of the approach’s effectiveness in delivering transparent, actionable insights for sustainable urban mobility management.

### 3. Preliminary

#### 3.1 Data Overview

As shown in Table 1, this study selects data provided by Capital Bicycle, which includes information such as date, season, and weather conditions.

*Table 1: Data Field Description.*

Field Name	Description
year	Year of observation
month	Month of observation
day	Day of observation
hour	Hour of observation
weekday	Day of the week
season	Season of observation
holiday	Whether it is a holiday
workingday	Whether it is a working day
weather	Weather condition
temp	Temperature in Celsius
atemp	Body temperature in Celsius
humidity	Humidity of the environment
windspeed	Wind speed in m/s
casual	Number of casual users
registered	Number of registered users
count	Total number of bike rentals
day_type	Type of the day (weekday/weekend)

### 3.2 Data processing

Data preprocessing was a critical step to ensure the dataset was clean, well-structured, and suitable for accurate prediction. The initial dataset included 10,886 records for training and 6,493 records for testing, with 12 and 9 features, respectively. First, missing data were examined and no missing values were identified in either the training or test datasets.

Outliers were then identified through the use of statistical methods. Variables like humidity, wind speed, and the target variable count exhibited irregularities. Specifically, wind speed contained numerous zero values, which were treated as potential missing data. To handle such cases, outlier removal was performed based on the criterion of values deviating more than 3 standard deviations from the mean.

For outlier correction in the atemp (apparent temperature), a linear regression model was employed to predict the correct values based on temperature, as defined by the formula:

$$atemp_{predicted} = \beta_0 + \beta_1 \cdot temperature \quad (1)$$

where  $\beta_0$  and  $\beta_1$  are the coefficients derived from a linear regression model.

In the case of the target variable count, extreme outliers were eliminated using the following criteria:

$$|count - \mu_{count}| > 3\sigma_{count} \quad (2)$$

where  $\mu_{count}$  is the mean of the count variable and  $\sigma_{count}$  is its standard deviation. Data points falling outside of this range were removed from the dataset.

A logarithmic transformation was then applied to the count variable to stabilize variance, as this transformation is effective in reducing skewness and making the data more normally distributed:

$$count_{log} = \log(count + 1) \quad (3)$$

Temporal features were also extracted from the timestamp column, including the day of the week, month, and hour, which were treated as categorical variables to capture time-dependent patterns.

These preprocessing operations ensured that the data were ready for model training, mitigating issues like outliers and skewed distributions while enriching the dataset with additional time-related features for improved predictive performance.

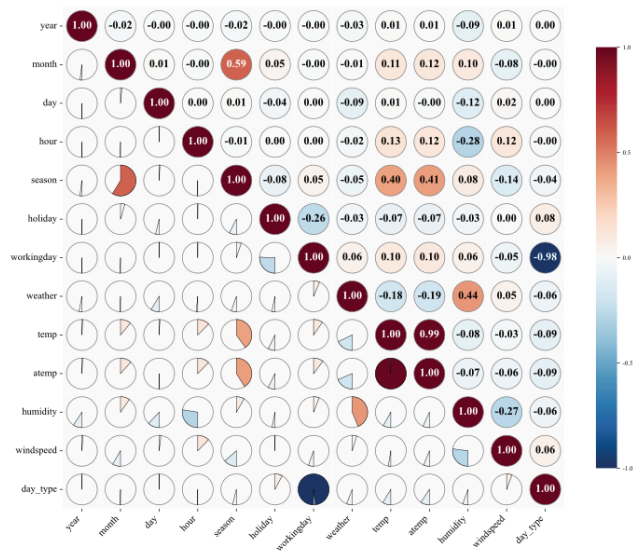
### 3.3 Correlation analysis

The data analysis aimed to identify key factors influencing shared bicycle demand, leveraging correlation analysis and trend evaluations. The correlation matrix (Figure 1) was calculated to explore the relationships between different features and the target variable, count. The correlation coefficient ( $r$ ) between two variables,  $x$  and  $y$ , was computed using the formula:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (4)$$

Where:  $x_i$  and  $y_i$  are the individual data points for variables  $x$  and  $y$ ,  $\bar{x}$  and  $\bar{y}$  are the mean values of  $x$  and  $y$ , respectively, the summation is taken over all data points in the dataset.

Figure 1. Correlation Matrix of Bike-sharing Dataset Features.



As shown in Figure 1, this correlation matrix illustrates the strength and direction of linear relationships between bike-sharing system variables. The analysis reveals several notable patterns among temporal, environmental, and operational factors. Temperature (temp) and apparent temperature (atemp) demonstrate a strong positive correlation (0.99), indicating their close relationship in weather conditions. Environmental variables show moderate correlations, with temperature exhibiting positive associations with seasonal patterns (0.40-0.44). Temporal variables display expected relationships, such as the moderate correlation between working days and day type (-0.98), reflecting the distinction between weekdays and weekends. Holiday patterns show negative correlations with working days (-0.26), confirming the inverse relationship between holidays and regular work schedules. Wind speed and humidity demonstrate relatively weak correlations with most other variables, suggesting their more independent influence on system usage patterns.

### 3.4 Influence of Time Conditions on Rental Demand

Figure 2. Monthly Bike Rental Trends.

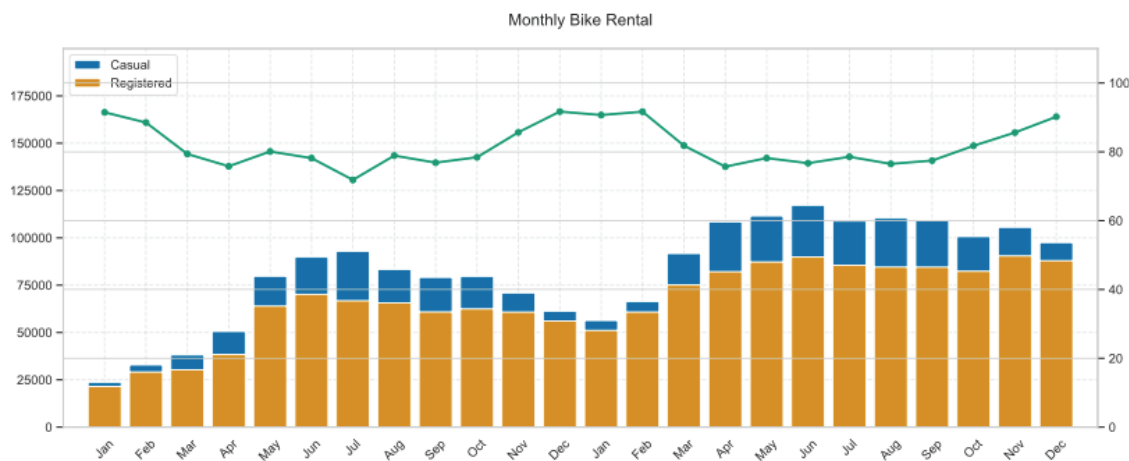


Figure 2 demonstrated a general upward trajectory in shared bicycle usage from 2 years, with notable seasonal fluctuations. Demand peaked during the summer and autumn months, with a significant decline observed in spring and winter. Registered users consistently represented over 75% of the total rentals each month, although their proportion slightly decreased in the warmer months, indicating a shift in user behavior. Non-registered users exhibited a preference for rentals during hotter months, further supporting the notion that shared bicycles are often used for short-term, seasonal needs.

Figure 3. Daily Bike Rental Distribution by Weekday and Day Type.

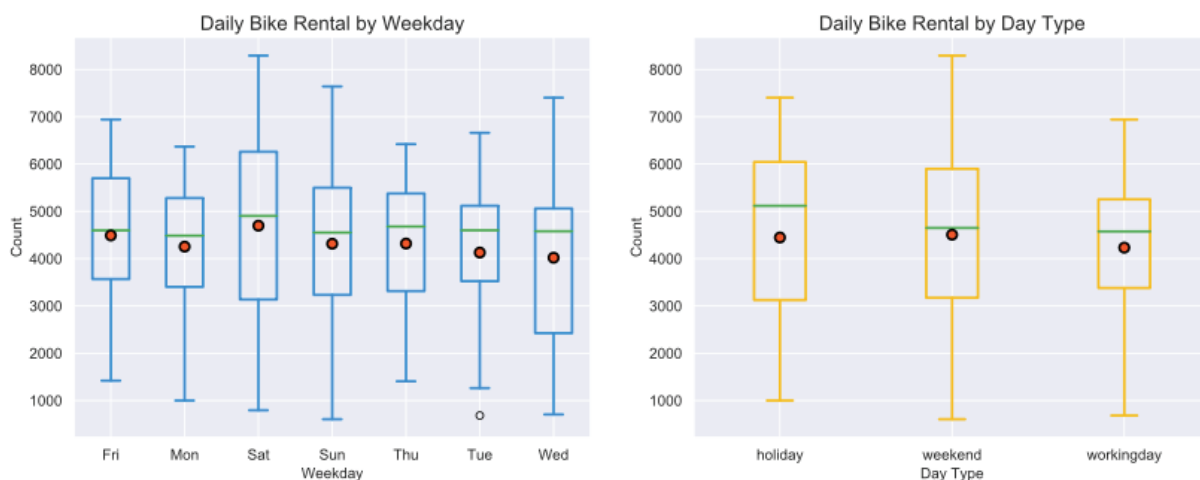


Figure 3 revealed that Saturdays experienced the highest rental volumes, likely driven by leisure activities, while Sundays saw slightly reduced demand, possibly due to lower mobility. Rentals on weekdays were more consistent, with Fridays exhibiting the lowest demand, albeit still higher than other weekdays. The impact of holidays on demand was more variable, with seasonal effects influencing demand peaks, particularly lower usage in the spring and winter months.

Figure 4. Hourly Bike Rental Patterns.

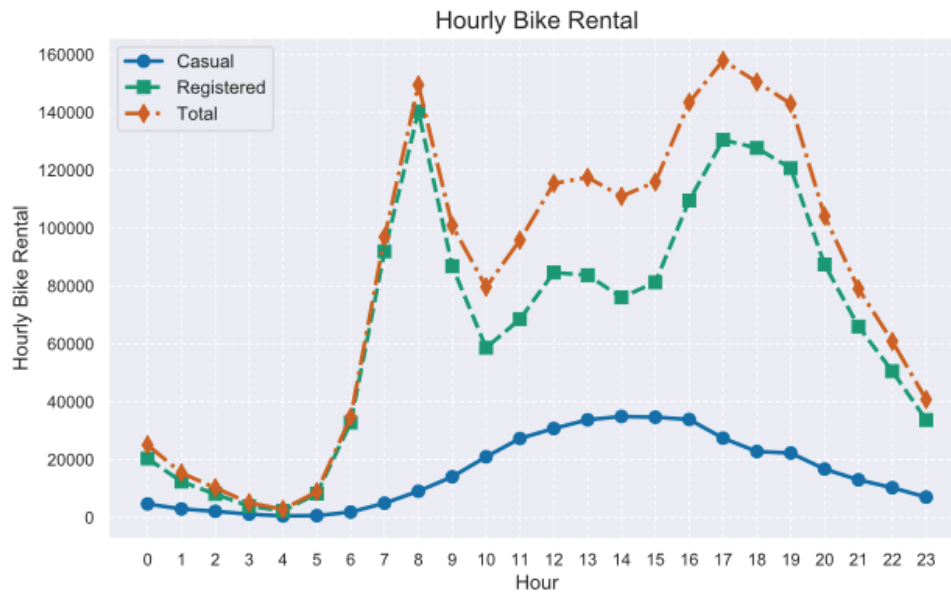
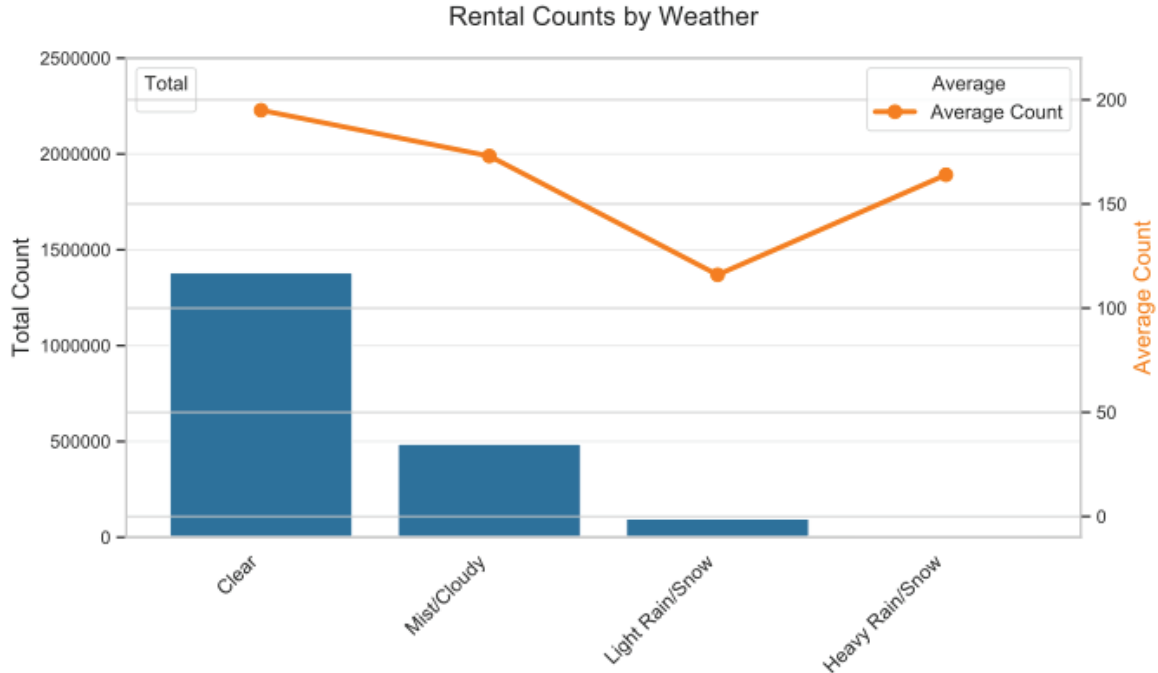


Figure 4 showed that registered users exhibited relatively stable demand throughout the day, while casual users demonstrated peak demand during morning and evening rush hours (7:00–8:00 and 17:00–18:00), reflecting their use of bicycles for commuting. These findings highlight the importance of considering time-related factors such as time of day, workdays, weekends, and holidays when planning for bicycle distribution and management.

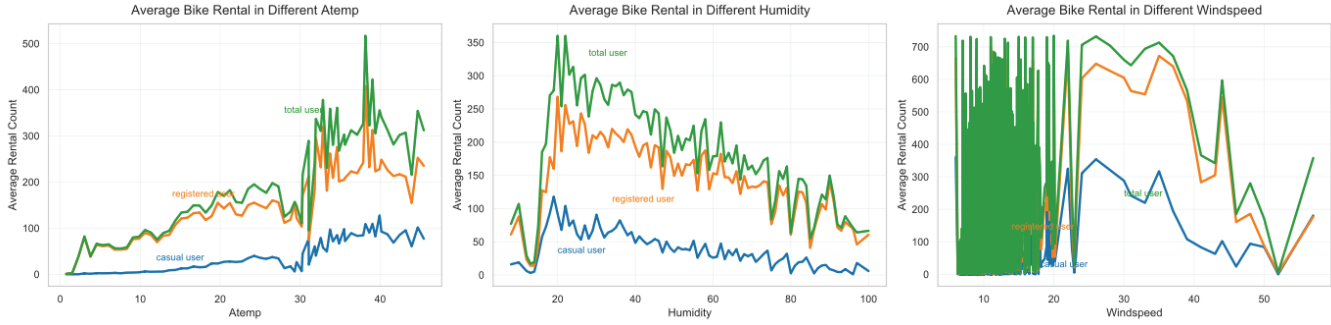
### 3.5 Influence of Weather on Rental Demand

Figure 5. Impact of Weather Conditions on Bike Rental Demand.



Weather conditions emerged as a significant factor influencing rental demand. The analysis of weather conditions (Figure 5) demonstrated a clear correlation between weather type and rental volumes. Interestingly, despite initial expectations, demand remained relatively high even during extreme weather events, such as heavy rain and snow. However, the distribution of weather-related data indicated that severe weather conditions, such as storms, led to a dramatic decline in rentals, with total demand during such events being just 1/9000th of the demand during clear weather, confirming a strong negative correlation between weather severity and demand.

Figure 6. Environmental Factor Influence on Average Bike Rental.



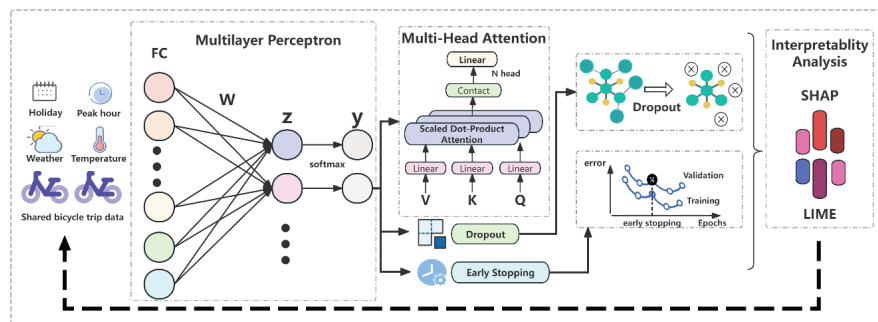
Environmental factors, including temperature, humidity, and windspeed (Figure 6), significantly influenced rental demand. Demand was lowest around 4°C, increasing with temperature up to 36°C, beyond which it declined, indicating a preference for moderate weather. Humidity showed a negative correlation, with demand highest at 20% humidity, which is more favorable for outdoor activities. Windy speed also impacted demand. Rentals were stable at wind speeds between 10 and 40 km/h, but sharply declined above 40 km/h. Interestingly, a brief rebound in demand was observed during high wind speeds around 17:00, likely due to commuter patterns during the evening rush hour. This suggests that while high wind speeds generally reduce rentals, commuter demand during peak hours can still drive usage.

These findings demonstrate the significant impact of both temporal and environmental conditions on shared bicycle rentals, with time-related factors revealing distinct hourly, weekly, and seasonal patterns, while environmental variables exhibit complex nonlinear relationships including temperature optima (4°C-36°C), humidity thresholds, and wind speed effects that interact dynamically across different temporal contexts. These empirical insights directly inform our MADENet architecture design, where the observed temporal periodicity patterns necessitate attention mechanisms for dynamic feature weighting, the nonlinear environmental relationships require multilayer perceptron components to capture complex meteorological-demand interactions, and the identified threshold effects guide our regularization strategy to ensure systematic integration of these multifaceted spatio-temporal and environmental factors without overfitting. The following methodology section details how these empirical findings are systematically integrated into our proposed framework.

## 4. Methodology

Chapter 4 introduces the MADENet model for bicycle-sharing demand forecasting, designed to address random user behavior and dynamic external factor that leads to supply–demand imbalance. MADENet integrates a multi-head attention mechanism with adaptive dropout regularization and early-stopping within a multilayer perceptron framework to strengthen key spatio-temporal feature representations while preventing overfitting. The architecture consists of five sequential components—Input, Attention, Dropout, MLP and Output layers—where the attention module assigns probabilistic weights to critical regional and temporal signals, the adaptive dropout adapts its rate during training, and the MLP captures complex nonlinear dependencies. Based on the described architecture and functional components, the proposed MADENet framework is structurally illustrated in Figure 7, which visually elucidates the synergistic integration of its multi-head attention mechanisms, adaptive dropout layers, and hierarchical feature processing pathways.

Figure 7. MADENet Framework Architecture.



## 4.1 Multilayer Perceptron

The Multilayer Perceptron is a fundamental component of MADENet. It consists of an input layer, one or more hidden layers, and an output layer. Each neuron in a layer is connected to all neurons in the next layer:

The output of a neuron  $y_j$  in the  $l$ -th layer is calculated as:

$$y_j^{(l)} = f \left( \sum_{i=0}^{n_{l-1}} E[w_{ji}^{(l)} x_i^{(l-1)} + b_j^{(l)}] \right) \quad (5)$$

Where:

$n_{l-1}$  is the number of neurons in the  $(l-1)$ -th layer,

$w_{ji}^{(l)}$  is the weight between the  $i$ -th neuron in the  $(l-1)$ -th layer and the  $j$ -th neuron in the  $l$ -th layer,

$x_i^{(l-1)}$  is the output of the  $i$ -th neuron in the  $(l-1)$ -th layer,

$b_j^{(l)}$  is the bias of the  $j$ -th neuron in the  $l$ -th layer,

and  $f$  is the activation function, such as the sigmoid function  $\sigma(x) = \frac{1}{1+e^{-x}}$  or the ReLU function  $f(x) = \max(0, x)$ ,  $E[\cdot]$  signifies the expectation operator applied to the weighted input.

## 4.2 Multi-Head Attention Mechanism

The multi-head attention mechanism in MADENet is used to dynamically capture cross-regional demand dependencies and temporal periodicity patterns.

The scaled dot-product attention is calculated as:

$$Attention(Q, K, V) = \text{softmax} \left( \frac{QK^T}{\sqrt{d_k}} \right) V \quad (6)$$

Where:  $Q$  is the query matrix,  $K$  is the key matrix,  $V$  is the value matrix, and  $d_k$  is the dimension of the keys.

The multi-head attention is composed of multiple parallel attention heads. The output of the multi-head attention is:

$$MultiHead(Q, K, V) = \text{Concat}(head_1, \dots, head_h) W^O \quad (7)$$

Where:  $head_i = Attention(QW_i^Q, KW_i^K, VW_i^V)$ ,  $W_i^Q, W_i^K, W_i^V$  are the weight matrices for the  $i$ -th head, and  $W^O$  is the output weight matrix.

## 4.3 Dropout

Dropout is a regularization technique used in MADENet to prevent overfitting. During training, neurons in a layer are randomly “dropped out” with a probability  $p$ .

Let  $x$  be the input to a layer. The output  $y$  after applying dropout is:

$$y = \frac{r}{1-p} \odot x \quad (8)$$

Where:  $r$  is a binary mask vector of the same length as  $x$ , and each element  $r_i$  is randomly set to 0 with probability  $p$  and 1 with probability  $1-p$ .

## 4.4 Early Stopping

Early stopping is another technique to prevent overfitting. It monitors the performance of the model on a validation set during training.

Let  $E_{val}(t)$  be the error on the validation set at the  $t$ -th training epoch. The training stops when the following condition is met:

$$E_{val}(t) > E_{val}(t-k) + \epsilon \quad (9)$$

Where:  $k$  is a predefined patience parameter, and  $\epsilon$  is a small positive constant. This ensures that the model does not over-train on the training data and generalizes well to new data.

# 5. Experiment

## 5.1 Experimental Configuration and Setup

The experimental evaluation was conducted using the Capital Bicycle dataset comprising over 17,000 hourly demand records with comprehensive temporal and environmental features including weather conditions, temperature, humidity, and seasonal

variables. The dataset was partitioned using an 8:2 stratified split for training and testing to ensure representative sampling of underlying demand patterns and variability.

MADENet was implemented using Python and TensorFlow framework, with training executed on a high-performance computing cluster using optimized parallel processing. Input data was formatted as 3D tensors (samples, 1, features) to accommodate the multi-head attention mechanism requirements. The model architecture employed 4 attention heads with key dimension of 32, multilayer perceptron structure of [512, 256, 128] neurons, and adaptive dropout rates (0.2-0.3). Training optimization utilized the Adam optimizer (learning rate 0.001) with mean squared error loss function, batch size of 64, and early stopping strategy (patience=20) monitoring validation loss to prevent overfitting. Hyperparameter optimization was conducted through systematic grid search across attention heads, hidden layer configurations, dropout rates, L2 regularization strengths (0.0, 0.001, 0.01), and batch normalization settings, with training limited to 100 epochs maximum.

## 5.2 Performance Metrics

Model performance was evaluated using the following metrics:

$R^2$  measures how well the model explains the variance in the target variable and indicates the proportion of the total variation that is captured by the model. It is defined as:

$$1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (10)$$

MAE measures the average absolute difference between predicted and actual values, providing a straightforward interpretation of prediction accuracy. It is defined as:

$$\frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (11)$$

MSE calculates the average of the squared differences between the predicted and actual values, offering a measure of how far predictions deviate from actual observations. It is defined as:

$$\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (12)$$

RMSE provides the square root of MSE, offering an interpretable estimate of the average magnitude of prediction error in the same units as the target variable. It is defined as:

$$\sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (13)$$

MAPE expresses the prediction error as a percentage of the actual values, providing an intuitive and scale-independent measure of model accuracy. It is defined as:

$$\frac{\sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|}{n} \quad (14)$$

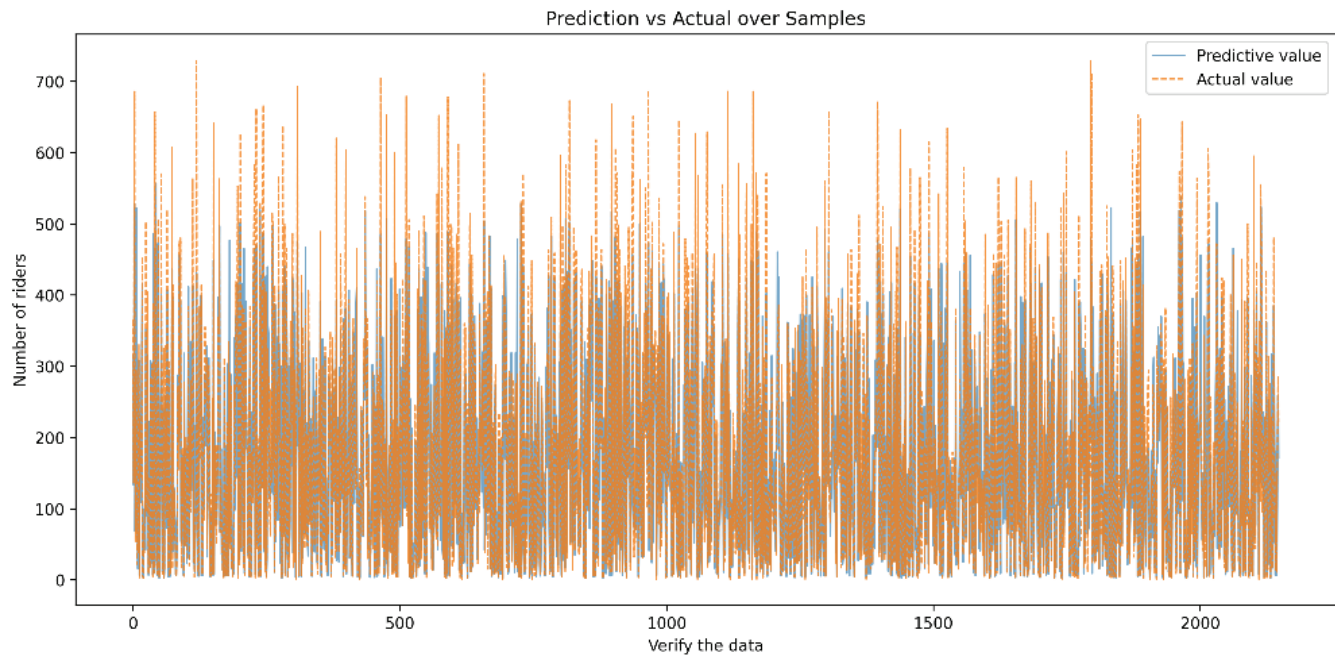
## 5.3 Comparative Analysis of Forecast Results

To comprehensively assess the performance of the MADENet model, Figure. 8 visualizes the predicted hourly bike-sharing demand alongside the actual recorded values across the test dataset. The continuous blue line represents the model's predicted values, while the orange dashed line denotes the ground truth. The graph clearly shows that MADENet captures both the short-term fluctuations and long-term periodic patterns in user demand with high fidelity.

Throughout more than 2,000 test samples, the predicted curves closely track the empirical values without significant lag or

deviation, even in regions with high variance and sharp peaks. This alignment indicates that the model not only learns overall trends but also adapts well to abrupt changes likely caused by external factors such as holidays, weather anomalies, or peak commuting hours.

Figure 8. Comparison of predicted and true values.



As shown in Table 2, to evaluate the predictive performance and generalization ability of the proposed MADENet model, a comprehensive comparative analysis was conducted against a variety of baseline models, including traditional machine learning methods and advanced deep learning architectures. These models included IrConv-LSTM [32], CNN-LSTM [33], Bi-LSTM [9], DeepAR [34], Decision Tree Regression [10], LSTM [12], RNN, AdaBoost Regression, K-Nearest Neighbors (KNN), Support Vector Regression (SVR), Elastic Net, Bayesian Ridge, Ridge Regression, Linear Regression, and Lasso Regression.

Table 2. Model Evaluation Results.

Model	R <sup>2</sup>	MAE	MSE	RMSE	MAPE
MADENet	0.9515	0.2320	0.1046	0.3234	7.3083
IrConv-LSTM	0.9363	0.2515	0.1263	0.3554	8.2922
CNN-LSTM	0.9205	0.2990	0.1600	0.4000	9.1285
Bi-LSTM	0.8281	0.4492	0.3437	0.5862	12.4637
DeepAR	0.9182	0.3028	0.1653	0.4066	9.4387
Decision Tree	0.9088	0.2933	0.1781	0.4220	8.8363
LSTM	0.8416	0.4416	0.3073	0.5544	12.7009
RNN	0.8046	0.4692	0.3924	0.6264	16.1000
AdaBoost	0.7829	0.5320	0.4301	0.6558	15.3739
KNN	0.7673	0.5050	0.4611	0.6790	16.4689
SVR	0.6367	0.6140	0.7198	0.8484	22.1008
Elastic Net	0.4742	0.7968	1.0267	1.0133	25.5929
Bayesian Ridge	0.4742	0.7959	1.0267	1.0133	25.5597
Ridge Regression	0.4740	0.7958	1.0271	1.0135	25.5543
Linear Regression	0.4739	0.7958	1.0273	1.0135	25.5526
Lasso Regression	0.4705	0.8014	1.0340	1.0169	25.7206

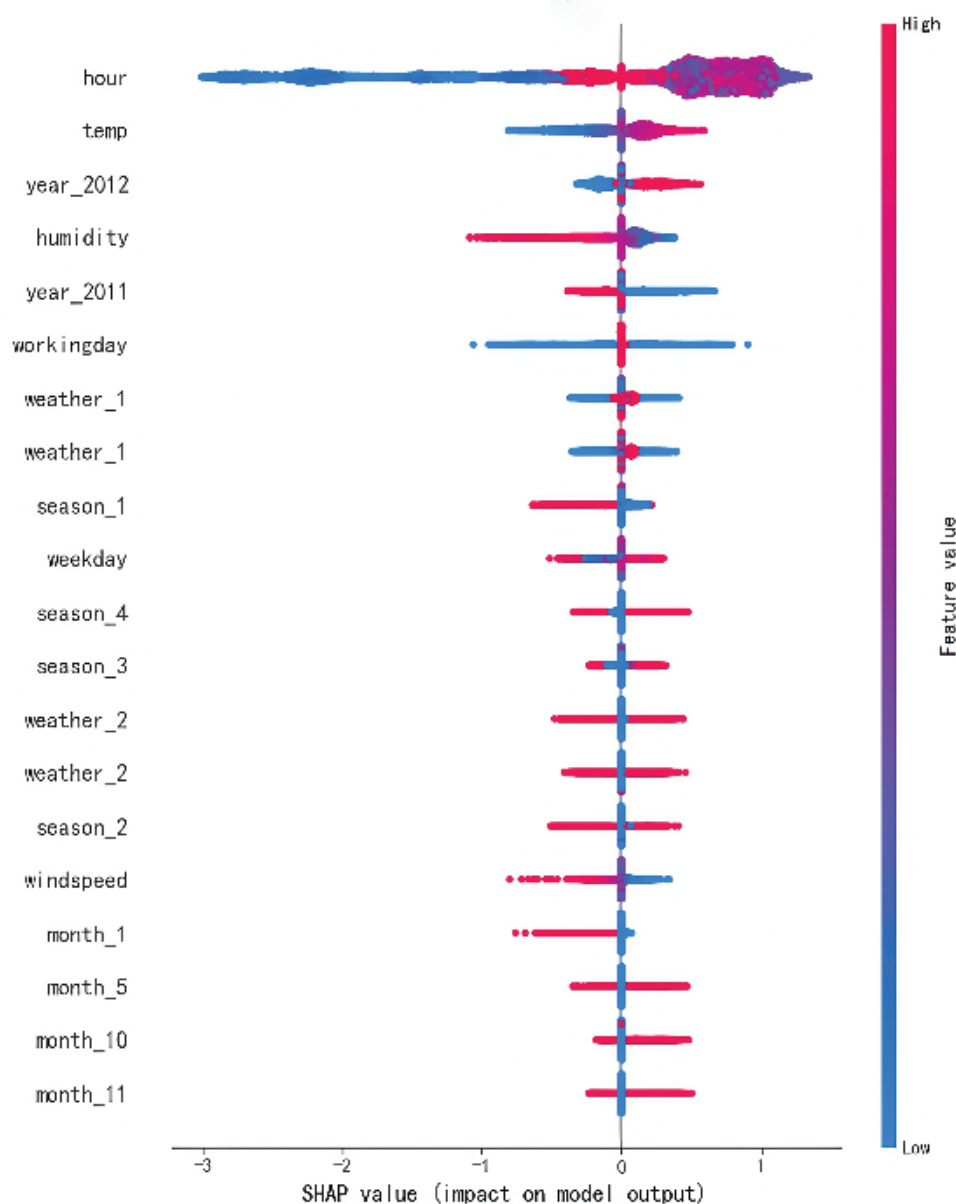
MADENet achieves superior performance across all evaluation metrics, demonstrating exceptional predictive accuracy with  $R^2=0.9515$ ,  $MAE=0.2320$ ,  $MSE=0.1046$ ,  $RMSE=0.3234$ , and  $MAPE=7.31\%$ . This comprehensive performance establishes MADENet as the leading method among all tested approaches.

Compared to advanced deep learning architectures, MADENet significantly outperforms CNN-LSTM with 22.4% MAE reduction and 34.6% MSE reduction, while achieving notable improvements of 7.7% in MAE over IrConv-LSTM. MADENet demonstrates substantial advantages over recurrent architectures, with 47.4% and 50.2% MAE improvements compared to Bi-LSTM and LSTM, respectively. Notably, MADENet surpasses the state-of-the-art probabilistic DeepAR model ( $R^2=0.9182$ ) by 3.6% in variance explanation while achieving 23.4% improvement in MAE, demonstrating superior effectiveness in capturing complex spatio-temporal demand patterns.

Against traditional machine learning approaches, MADENet exhibits remarkable performance gains. Compared to Decision Tree regression, MADENet achieves 4.7% higher  $R^2$  and 20.9% lower MAE. The superiority extends to ensemble methods, with MADENet outperforming AdaBoost by 56.4% in MAE reduction and achieving 21.8% improvement over KNN. When evaluated against linear regression methods, MADENet demonstrates exceptional advancement with over 100% improvement in  $R^2$  (from 0.47 to 0.9515) and approximately 70% reduction in prediction errors across all metrics.

#### 5.4 Factor Analysis Based on SHAP Values

Figure 9. SHAP Analysis of the importance of impact factors.

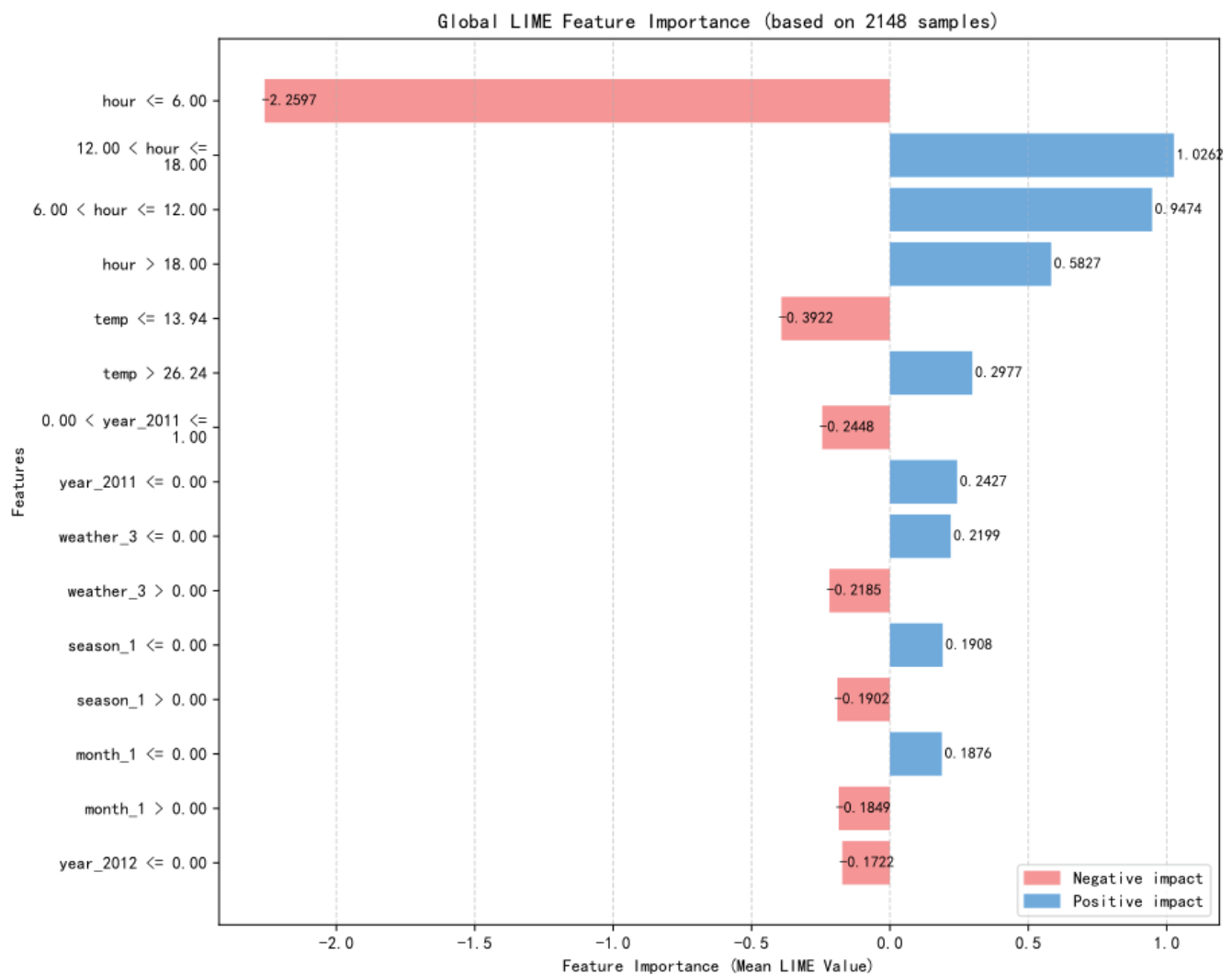


To enhance interpretability and understand MADENet's internal decision logic, SHAP values were employed to evaluate feature contributions to model outputs. Figure 9 presents the SHAP summary plot where each point represents a single SHAP value, with color gradients denoting feature values from low (blue) to high (red). The analysis reveals that temporal indicators, particularly hour-of-day variables, exert the greatest influence on model predictions, with the "hour" feature demonstrating both high density and wide spread, indicating its dominant role in shaping ridership forecasts and reflecting strong regularity in daily urban mobility patterns.

Climatic and seasonal features such as temperature and humidity exhibit substantial but nuanced impacts, with warmer temperatures generally corresponding to increased predicted demand while humidity demonstrates complex nonlinear effects. Calendar context variables including working days, weekdays, and seasons present moderate SHAP influence with values clustering near zero, suggesting their effects depend primarily on interactions with other features rather than exerting strong independent influence, while environmental indicators like wind speed show relatively limited marginal contributions in the current model configuration.

### 5.5 Factor Analysis Based on LIME Values

Figure 10. LIME Analysis of the importance of impact factors.



LIME values were utilized to further investigate feature importance through global feature impact quantification. Figure 10 shows the global LIME feature importance plot where positive values represent positive contributions and negative values indicate negative impacts. The analysis confirms that time-of-day variables dominate the predictive power, with early morning hours exhibiting the highest positive impact, strongly suggesting commuter-driven demand patterns, while midday periods also demonstrate substantial positive influence, reinforcing the critical importance of specific temporal windows in

demand forecasting.

Temperature emerges as another crucial factor with consistent positive LIME values, indicating that warmer conditions generally increase bike-sharing demand and aligning with observed weather-behavior relationships in urban cycling environments. Notably, negative impacts are observed for certain months and seasonal variables, suggesting weather-dependent demand variations, while some features display minimal contribution with LIME values near zero, indicating marginal influence and highlighting the model's ability to distinguish between critical and peripheral predictive factors.

## 6. Conclusions

This study successfully addresses the dual challenge of achieving high-precision and interpretable demand forecasting in urban bike-sharing systems through MADENet, a novel neural architecture combining multi-head attention, adaptive dropout, and early stopping mechanisms. The framework achieves 95.1% prediction accuracy while providing systematic transparency through integrated SHAP and LIME analysis, identifying hour-of-day, temperature, and humidity as dominant drivers. These interpretable insights empower urban planners with actionable guidance for station placement optimization, dynamic fleet rebalancing, and evidence-based policy formulation, ultimately supporting sustainable urban mobility through more effective resource allocation and environmental responsiveness.

Despite these advances, several limitations constrain the model's broader applicability. The evaluation on a single operator's dataset (Capital Bicycle) limits generalization across different operators, urban contexts, and regional characteristics with varying infrastructure, user behaviors, and operational constraints. Additionally, the reliance on static historical data rather than real-time dynamic data sources may reduce model effectiveness in rapidly evolving urban environments where demand patterns shift due to unexpected events, policy changes, or seasonal disruptions. These constraints highlight the need for more comprehensive validation frameworks that account for cross-regional variability and real-time operational dynamics.

Future research should prioritize expanding data sources through multi-operator collaborations and cross-city validation to enhance model generalizability and robustness across diverse urban environments. Critical development areas include federated learning architectures for privacy-preserving multi-operator training, online learning capabilities for real-time adaptability, and integration of unstructured data streams such as event schedules and social media indicators. Furthermore, extending MADENet's framework to other shared mobility modes (car-sharing, scooter-sharing) and developing scalable deployment strategies for different city sizes and infrastructure levels will enhance the model's practical utility for comprehensive multimodal transportation ecosystem management, fostering data-driven governance in smart city development.

## Funding

This research was funded by the National Natural Science Foundation of China (62462019, 62172350), Guangdong Basic and Applied Basic Research Foundation (2023A1515012846), Guangxi Science and Technology Major Program (AA24263010), The Key Research and Development Program of Guangxi (AB24010085, AB23026120), Key Laboratory of Equipment Data Security and Guarantee Technology, Ministry of Education (GDZB2024060500), Natural Science Foundation of Guangxi Province (2025GXNSFBA069410), and Basic Scientific Research Ability Improvement Project for Young and Middle-aged Teachers of Guangxi Higher Education Institutions (2024KY0233, 2025KY0243).

## Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Reference

- [1] Afsari, M., Dastmard, M., Bresciani Miristice, L. M., & Gentile, G. (2024). Utilizing P-median and Machine Learning for Bike-Sharing Demand Prediction. 2024 IEEE International Conference on Environment and Electrical Engineering and 2024 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe), 1–5. <https://doi.org/10.1109/EEEIC/ICPSEurope61470.2024.10751398>
- [2] Andor, M. A., Gerster, A., Gillingham, K. T., & Horvath, M. (2020). Running a car costs much more than people think—

- Stalling the uptake of green travel. *Nature*, 580(7804), 453–455. <https://doi.org/10.1038/d41586-020-01118-w>
- [3] Biswas, Md. E., Hossain, Md. D., Huh, E.-N., & Sultana, T. (2025). Spatiotemporal Data-Driven Hourly Bike-Sharing Demand Prediction Using ApexBoost Regression. *International Journal of Data Science and Analytics*. <https://doi.org/10.1007/s41060-025-00820-0>
- [4] Chen, P., Hsieh, H., Su, K., Sigalingging, X. K., Chen, Y., & Leu, J. (2020). Predicting station level demand in a bike-sharing system using recurrent neural networks. *IET Intelligent Transport Systems*, 14(6), 554–561. <https://doi.org/10.1049/iet-its.2019.0007>
- [5] Cheng, C., & Yue, W. (2024). Construction of a Green and Low-Carbon Travel Order Prediction Model Based on Shared Bicycle Big Data. *International Journal of Computational Intelligence Systems*, 17(1), 130. <https://doi.org/10.1007/s44196-024-00519-3>
- [6] Chengula, T. J., Mwakalonge, J., Comert, G., Sulle, M., Siuhi, S., & Osei, E. (2024). Enhancing advanced driver assistance systems through explainable artificial intelligence for driver anomaly detection. *Machine Learning with Applications*, 17, 100580. <https://doi.org/10.1016/j.mlwa.2024.100580>
- [7] Collini, E., Nesi, P., & Pantaleo, G. (2021). Deep Learning for Short-Term Prediction of Available Bikes on Bike-Sharing Stations. *IEEE Access*, 9, 124337–124347. <https://doi.org/10.1109/ACCESS.2021.3110794>
- [8] E, S. V., Park, J., & Cho, Y. (2020). Using data mining techniques for bike sharing demand prediction in metropolitan city. *Computer Communications*, 153, 353–366. <https://doi.org/10.1016/j.comcom.2020.02.007>
- [9] Feng, J., Liu, H., Zhou, J., & Zhou, Y. (2024). A Spatial-Temporal Aggregated Graph Neural Network for Docked Bike-sharing Demand Forecasting. *ACM Transactions on Knowledge Discovery from Data*, 18(9), 1–27. <https://doi.org/10.1145/3690388>
- [10] Galatoulas, N.-F., Genikomsakis, K. N., & Ioakimidis, C. S. (2020). Spatio-Temporal Trends of E-Bike Sharing System Deployment: A Review in Europe, North America and Asia. *Sustainability*, 12(11), 4611. <https://doi.org/10.3390/su12114611>
- [11] Gao, C., & Chen, Y. (2022). Using Machine Learning Methods to Predict Demand for Bike Sharing. In J. L. Stienmetz, B. Ferrer-Rosell, & D. Massimo (Eds.), *Information and Communication Technologies in Tourism 2022* (pp. 282–296). Springer International Publishing. [https://doi.org/10.1007/978-3-030-94751-4\\_25](https://doi.org/10.1007/978-3-030-94751-4_25)
- [12] Giner Fabregat, G., Fonseca I Casas, P., & Rivero Martínez, A. (2025). Intelligent Optimization of Bike-Sharing Systems: Predictive Models and Algorithms for Equitable Bicycle Distribution in Barcelona. *Sustainability*, 17(10), 4316. <https://doi.org/10.3390/su17104316>
- [13] Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G.-Z. (2019). XAI—Explainable artificial intelligence. *Science Robotics*, 4(37), eaay7120. <https://doi.org/10.1126/scirobotics.aay7120>
- [14] Guo, H., Li, K., & Rou, Y. (2024). Optimized Demand Forecasting for Bike-Sharing Stations Through Multi-Method Fusion and Gated Graph Convolutional Neural Networks. *IEEE Access*, 12, 174017–174027. <https://doi.org/10.1109/ACCESS.2024.3501572>
- [15] Guo, H., Zhao, S., Ren, Y., Li, J., & Xu, S. (2025). An XGBoost-Based Three-Stage Prediction Approach for True User Demand of Bike-Sharing Systems Based on Spatio-Temporal Analysis. *IEEE Transactions on Intelligent Transportation Systems*, 1–12. <https://doi.org/10.1109/TITS.2025.3586747>
- [16] Hu, Y., Sun, Z., Li, W., & Pei, L. (2022). Forecasting public bicycle rental demand using an optimized eXtreme Gradient Boosting model. *Journal of Intelligent & Fuzzy Systems*, 42(3), 1783–1801. <https://doi.org/10.3233/JIFS-211202>
- [17] Lee, Y., Son, H., Ahn, J., & Cho, S. (2024). Bike-sharing Demand Prediction based on Artificial Intelligence Algorithm Using Weather Data. *2024 IEEE International Conference on Consumer Electronics (ICCE)*, 1–6. <https://doi.org/10.1109/ICCE59016.2024.10444462>
- [18] Li, X., Xu, Y., Chen, Q., Wang, L., Zhang, X., & Shi, W. (2022). Short-Term Forecast of Bicycle Usage in Bike Sharing Systems: A Spatial-Temporal Memory Network. *IEEE Transactions on Intelligent Transportation Systems*, 23(8), 10923–10934. <https://doi.org/10.1109/TITS.2021.3097240>

- [19] Li, X., Xu, Y., Zhang, X., Shi, W., Yue, Y., & Li, Q. (2023). Improving short-term bike sharing demand forecast through an irregular convolutional neural network. *Transportation Research Part C: Emerging Technologies*, 147, 103984. <https://doi.org/10.1016/j.trc.2022.103984>
- [20] Li, Y., & Zheng, Y. (2020). Citywide Bike Usage Prediction in a Bike-Sharing System. *IEEE Transactions on Knowledge and Data Engineering*, 32(6), 1079–1091. <https://doi.org/10.1109/TKDE.2019.2898831>
- [21] Li, Y., Zhu, Z., Kong, D., Xu, M., & Zhao, Y. (2019). Learning Heterogeneous Spatial-Temporal Representation for Bike-Sharing Demand Prediction. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01), 1004–1011. <https://doi.org/10.1609/aaai.v33i01.33011004>
- [22] Liang, Y., Ding, F., Huang, G., & Zhao, Z. (2023). Deep trip generation with graph neural networks for bike sharing system expansion. *Transportation Research Part C: Emerging Technologies*, 154, 104241. <https://doi.org/10.1016/j.trc.2023.104241>
- [23] Liang, Y., Huang, G., & Zhao, Z. (2024). Cross-Mode Knowledge Adaptation for Bike Sharing Demand Prediction Using Domain-Adversarial Graph Neural Networks. *IEEE Transactions on Intelligent Transportation Systems*, 25(5), 3642–3653. <https://doi.org/10.1109/TITS.2023.3322717>
- [24] Lim, H., Chung, K., & Lee, S. (2022). Probabilistic Forecasting for Demand of a Bike-Sharing Service Using a Deep-Learning Approach. *Sustainability*, 14(23), 15889. <https://doi.org/10.3390/su142315889>
- [25] Lin, H., He, Y., Li, S., & Liu, Y. (2024). Insights into Travel Pattern Analysis and Demand Prediction: A Data-Driven Approach in Bike-Sharing Systems. *Journal of Transportation Engineering, Part A: Systems*, 150(2), 04023132. <https://doi.org/10.1061/JTEPBS.TEENG-8137>
- [26] Pribadi, A. P., Rauf, A. U., Rahman, Y. M. R., & Haq, Z. F. (2025). Air Quality and Urban Sustainable Development- Current Issues and Future Directions. In J. A. Parray, N. Shameem, & A. K. Haghi (Eds.), *Sustainable Urban Environment and Waste Management* (pp. 23–51). Springer Nature Singapore. [https://doi.org/10.1007/978-981-96-1140-9\\_2](https://doi.org/10.1007/978-981-96-1140-9_2)
- [27] Qian, H., Wang, J., Chen, Y., Zheng, S., & Wei, Z. (2025). CGA-STNet: A dockless shared bicycle demand prediction model considering multiple spatial features and time periodicity. *Expert Systems with Applications*, 265, 126100. <https://doi.org/10.1016/j.eswa.2024.126100>
- [28] Roscher, R., Bohn, B., Duarte, M. F., & Garcke, J. (2020). Explainable Machine Learning for Scientific Insights and Discoveries. *IEEE Access*, 8, 42200–42216. <https://doi.org/10.1109/ACCESS.2020.2976199>
- [29] Sajid, M., Midhun, V., Zakkariya, K. A., Surira, M. D., & Vishnu, K. P. (2023). Pedaling towards sustainability: A mixed-method study of the drivers and barriers to bike-sharing adoption. *Management of Environmental Quality: An International Journal*, 34(6), 1580–1606. <https://doi.org/10.1108/MEQ-12-2022-0343>
- [30] Schnieder, M. (2023). Ebike Sharing vs. Bike Sharing: Demand Prediction Using Deep Neural Networks and Random Forests. *Sustainability*, 15(18), 13898. <https://doi.org/10.3390/su151813898>
- [31] Shi, Y., Zhang, L., Lu, S., & Liu, Q. (2023). Short-Term Demand Prediction of Shared Bikes Based on LSTM Network. *Electronics*, 12(6), 1381. <https://doi.org/10.3390/electronics12061381>
- [32] Shi, Z., Jiang, Z., Xu, M., Zhou, Y., & Zhu, Z. (2025). Enhancing Hybrid Bike-Sharing Systems Through Long-Term Inventory Management: A Generative-Model-Informed Reinforcement Learning Approach. *IEEE Transactions on Computational Social Systems*, 1–13. <https://doi.org/10.1109/TCSS.2025.3557911>
- [33] Wang, T., Hu, S., & Jiang, Y. (2021). Predicting shared-car use and examining nonlinear effects using gradient boosting regression trees. *International Journal of Sustainable Transportation*, 15(12), 893–907. <https://doi.org/10.1080/15568318.2020.1827316>
- [34] Wang, Z., & Ren, F. (2025). Developing a decision support system for sustainable urban planning using machine learning-based scenario modeling. *Scientific Reports*, 15(1), 13210. <https://doi.org/10.1038/s41598-025-90057-5>
- [35] Wang, Z., Yu, D., Zheng, X., Meng, F., & Wu, X. (2025). A Model-Data Dual-Driven Approach for Predicting Shared Bike Flow near Metro Stations. *Sustainability*, 17(3), 1032. <https://doi.org/10.3390/su17031032>

- [36] Wei, X., Ren, Y., Shen, L., & Shu, T. (2022). Exploring the spatiotemporal pattern of traffic congestion performance of large cities in China: A real-time data based investigation. *Environmental Impact Assessment Review*, 95, 106808. <https://doi.org/10.1016/j.eiar.2022.106808>
- [37] Xiang, Z., Zeng, F., Liu, L., Wu, J., Mumtaz, S., & Leung, V. C. M. (2025). Bike-Sharing Demand Prediction Based on Dynamic Time Warping and Spatio-Temporal Graph Attention Network. *IEEE Transactions on Intelligent Transportation Systems*, 1–15. <https://doi.org/10.1109/TITS.2025.3570009>
- [38] Xiao, G., Wang, R., Zhang, C., & Ni, A. (2021). Demand prediction for a public bike sharing program based on spatio-temporal graph convolutional networks. *Multimedia Tools and Applications*, 80(15), 22907–22925. <https://doi.org/10.1007/s11042-020-08803-y>
- [39] Xu, F., Uszkoreit, H., Du, Y., Fan, W., Zhao, D., & Zhu, J. (2019). Explainable AI: A Brief Survey on History, Research Areas, Approaches and Challenges. In J. Tang, M.-Y. Kan, D. Zhao, S. Li, & H. Zan (Eds.), *Natural Language Processing and Chinese Computing* (Vol. 11839, pp. 563–574). Springer International Publishing. [https://doi.org/10.1007/978-3-030-32236-6\\_51](https://doi.org/10.1007/978-3-030-32236-6_51)
- [40] Xu, M., Liu, H., & Yang, H. (2020). A Deep Learning Based Multi-Block Hybrid Model for Bike-Sharing Supply-Demand Prediction. *IEEE Access*, 8, 85826–85838. <https://doi.org/10.1109/ACCESS.2020.2987934>
- [41] Yu, L., Feng, T., Li, T., & Cheng, L. (2023). Demand Prediction and Optimal Allocation of Shared Bikes Around Urban Rail Transit Stations. *Urban Rail Transit*, 9(1), 57–71. <https://doi.org/10.1007/s40864-022-00183-w>
- [42] Zhu, B., Hu, S., Kaparias, I., Zhou, W., Ochieng, W., & Lee, D.-H. (2024). Revealing the driving factors and mobility patterns of bike-sharing commuting demands for integrated public transport systems. *Sustainable Cities and Society*, 104, 105323. <https://doi.org/10.1016/j.scs.2024.105323>

# Low-Cost CCTV Repurposing for Sustainable Parking Management: A Non-AI Computer Vision Case Study

Namya Kamboj, Kongwen Zhang\*

School of Computing, University of the Fraser Valley, Abbotsford (BC), Canada

\*Corresponding author: Kongwen (Frank) Zhang, [Frank.Zhang@ufv.ca](mailto:Frank.Zhang@ufv.ca); 0000-0001-7570-9805

**Copyright:** 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

**Abstract:** Urban parking inefficiency results in fuel waste and emissions, yet the current monitoring systems rely on resource-intensive AI or sensor-based approaches. Aligning with a focus on sustainable technology, this project demonstrates how deterministic computer-vision techniques (using adaptive thresholding and polygon masking) can transform the existing CCTV camera framework into parking occupancy detectors without AI. This system deploys an open-source pipeline, combining OpenCV and PyQt5, on UFV's infrastructure, which requires zero hardware costs and consumes 95% less power than other GPU-based solutions. Testing with 18,000+ frames of simulated CCTV footage, the system achieved approximately 99% accuracy. This case study presents a replicable solution for institutions in resource-constrained environments, demonstrating that an economical IoT-CV integration can optimize urban resources while minimizing AI's carbon footprint.

**Keywords:** Parking Occupancy Monitoring; Low-Cost Computer Vision; Adaptive Thresholding; Sustainable Urban Mobility; Polygon Masking; Non-AI Solution

**Published:** Oct 15, 2025

**DOI:** <https://doi.org/10.62177/jaet.v2i4.666>

## 1.Introduction

Parking inefficiency is a persistent problem for universities, directly affecting productivity and the overall campus experience. At the University of the Fraser Valley (UFV), students and staff circling around parking lots waste around 10-15 minutes during peak traffic hours, which is almost equivalent to losing 1.5 weeks of time searching for parking in an academic year. According to a study by Ponnambalam et al. (2018), driver anxiety levels increased by 27% due to daily traffic and parking stress as opposed to those with smoother commuting experiences <sup>[1]</sup>. It is necessary to optimize parking as it reduces fuel waste and stress levels, thereby leading to the sustainability of urban resources and better physiological health <sup>[2]</sup>. Addressing this problem is crucial to enhancing daily operations and user satisfaction, as parking inefficiency is a component of a broader urban resource management issue. Many global campuses face similar challenges due to a lack of scalable, low-cost parking monitoring systems. This leads to reliance on either manual checks or AI (Artificial Intelligence)-heavy solutions, both of which are impractical for institutions like UFV.

While AI-based systems achieve high accuracy, they require substantial computational resources and continuous maintenance. Sensor-based alternatives cost \$200–\$500 per spot, while traditional computer vision lacks adaptability to irregular layouts and environmental conditions.

We present a rules-based computer vision system that offers sustainable urban mobility as it:

1. Repurposes existing CCTV cameras as IoT nodes, eliminating hardware costs.
2. Uses adaptive thresholding and polygon masking to handle irregular spaces without AI.
3. Operates edge-compatible (SQLite/PyQt5), aligning with sustainability goals.

This work demonstrates an image processing technique, that when combined with flexible polygon-based zoning can present a real-time parking occupancy monitoring system which is at a fraction of computational and financial cost of complex AI-driven models. This addresses UFV's unique needs since the CCTV camera angles are slanting, resulting in irregular parking spot shapes and also takes care of the weather conditions when the ground is completely covered with snow, for example.

## 2. Literature Review

Pre-existing and currently operational research related to parking occupancy and detection systems relies heavily on AI and deep-learning algorithms. Although the accuracy of these systems is high and consistent, they require substantial computational resources, massive training datasets and continuous upgrade and maintenance, to maintain their consistency.

### 2.1 Dominance of AI/Deep Learning in Parking Detection

Recent research prioritizes AI-based systems, such as Mask R-CNN<sup>[3]</sup> and YOLOv5<sup>[4]</sup>, which achieve high accuracy (>90% or 0.91) but suffer from critical limitations when applied to real-world situations:

#### 2.1.1 Resource Exhaustion (Computational Burden)

Mask R-CNN model trains on large datasets, for example COCO, due to which it requires GPU acceleration. Also, these models depend on multiple layers and parameters and require complex hardware and software resources to train and run. When working with high-resolution objects, the image processing speed can sometimes be slowed down. This makes it impractical for real-time deployment of this model on edge devices due to applicability issues<sup>[3]</sup>. As a result, this high energy consumption conflicts with sustainability goals, which can eventually lead to processing overload.

#### 2.1.2 Data Dependence

These training models highly depend on annotated parking datasets for their accurate results, for instance, the PKLot Expert System by De Almeida et al.<sup>[5]</sup>, which utilizes a dataset consisting of 695,899 images of various parking lots<sup>[5]</sup>. Additionally, this annotated data needs to be fine-tuned before it is applied to any real-world scenario, which is time consuming, expensive and susceptible to errors. Even though there is greater accuracy in such systems, there is a scarcity for atypical parking layouts (e.g., angled or irregular spaces), which limits the adaptability to UFV's parking geometry without costly retraining. These models operate on standardized benchmarks and datasets, which is why there are several uncertainties involved while handling varied real-world scenarios. Hence, to implement these models, institutions must rely on manually annotated datasets, like collecting and labeling thousands of campus-specific images which would be costly.

#### 2.1.3 Regular Maintenance Burden

AI models are prone to degradation under lighting variations or camera shifts. As a result, there is a need for continuous fine-tuning and training of the model depending on the varied conditions and environments.

#### 2.1.4 Difficult to Interpret

One of the challenges while working with Mask R-CNN is that this model is considered as a black box as it is difficult to interpret and explain the decisions it makes along with the features used to get to that decision. This is an accountability and ethical issue because the output may produce false negatives, false positive or even, biased results, further complicating its working and leading to serious implications<sup>[6]</sup>.

As a result, no AI solution balances accuracy, cost, and adaptability for institutions with fixed cameras and limited IT resources.

### 2.2 Sensor-based Systems (Accuracy at High Cost)

Pre-existing solutions relied on inductive loops<sup>[7]</sup> and ultrasonic sensors<sup>[8]</sup>. Inductive loop work on induced current that comes from the wire loop fitted in the pavement. When a vehicle passes through the loop, its presence is transmitted through a signal. Ultrasonic sensors work on the principle of reflection of ultrasonic waves. Presence of a vehicle is detected by the ultrasonic transmitters and receivers which are fitted directly above the road. A sensing range is set by the traffic sensors which determine the occupancy rates<sup>[9]</sup>.

While these solutions are accurate and reliable, there are a few limitations attached to them. First, the infrastructural costs are high, costing \$5000–\$10000 per lane for installation (in terms of traffic control sensors) <sup>[10]</sup>. Additionally, there are scalability issues because hardware failures require manual intervention which leads to an increase in maintenance overhead. There are multiple detectors required for a location when implementing an inductive loop approach. Since the wires for inductive loops are laid in the pavements, there is a risk of water penetration, which can affect its performance. Also, in the case of ultrasonic sensors, temperature changes and extreme air turbulence can impact accuracy and performance <sup>[11]</sup>. Among other issues, these systems are also not compatible with existing CCTV infrastructure.

### 2.3 Unexplored Potential of Non-AI Computer Vision Models

Traditional computer vision techniques, such as adaptive thresholding and background subtraction, have been largely dismissed in favor of deep learning algorithms due to perceived limitations. However, these methods offer untapped advantages, particularly for constrained environments like university parking lots, when enhanced with modern optimizations. Our work bridges this gap by introducing a hybrid approach that combines polygon flexibility with adaptive thresholding, overcoming key shortcomings of prior non-AI systems.

1. Addressing the Light Sensitivity Limitation: Conventional thresholding techniques struggle in low-light conditions which can lead to: (1) over-segmentation (shadows misclassified as occupied spaces) or (2) under-segmentation (when dark vehicles blend into the pavement). Our system maintains accuracy (greater than 98% in such conditions (low-light conditions tested with mobile-recorded CCTV simulations) with Dynamic Threshold Adjustment through median filtering (5x5 kernel) that reduces any noise from moving foliage or camera artifacts. Additionally, dilation cleans up binary masks that minimize the cases of false positives from transient shadows.

2. Overcoming Rigid Geometry Assumptions: Prior non-AI parking management systems are dependent on fixed rectangular templates, failing to accommodate:

Angled parking lots (common in CCTV images in UFV's Lots)

Curved or irregular boundaries

Dynamic layouts (temporary construction closures).

Our solution is Polygon-Based Masking. There is a manual vertex annotation where users define quadrilaterals (4-sided polygons) by selecting points to match real parking spaces (as explained in Methodology). Unlike rectangles, polygons adapt to: (1) diagonal slots, (2) non-parallel boundaries, and (3) discontinuous spaces (for example, pedestrian crossings). This eliminates the need for AI's data-intensive generalization, cutting deployment time from weeks (for AI training) to hours.

Currently, there is no hybrid approach that combines polygon flexibility with thresholding optimization, which is a key innovation of our work. Additionally, the currently operational commercial systems for parking space management are proprietary and inflexible. They have recurring licensing fees and data privacy risks as the processing is done off-site. Hence, no prior work offers an open-source, self-hosted alternative for academic institutions, a niche our system fills.

The table below [Table 1] shows the benefits that our system offers as compared to pre-existing technologies and parking occupancy systems:

*Table 1: Positioning our work*

Approach	Strengths	Limitations	Our Solution
AI (Mask R-CNN)	High Accuracy	GPU/data dependence	Rules-based; no training required
Sensors	Real-time data	High cost and maintenance	CCTV reuse; zero hardware
Thresholding	Low compute cost	Lighting/geometry fragility	Adaptive thresholds + polygon masking

## 3. Methodology: UFV Parking Monitoring as a Case Study

### 3.1 Problem Context

The purpose of this study is to develop a rules-based computer vision system which provides real-time parking occupancy detection without the use of AI. It leverages existing CCTV infrastructure and is therefore a low-cost and efficient model.

This system is ideal for educational institutions due to cost savings and also satisfies UFV's unique parking lot needs: (1)

mixed regular and angled spaces, (2) snow covered parking lots obscuring parking lines during winters, and (3) fixed CCTV cameras with low-light limitations, conditions unmet by commercial solutions, for instance, sensor-based systems fail in snow and AI based solution need to be retrained according to unique parking lot needs.

### 3.2 Data Collection

For this research, phone cameras are used to replicate CCTV video footage. The video was shot from a slanting angle similar to how the CCTV cameras are placed at UFV.

Our system uses a carefully selected set of open-source tools to achieve real-time parking monitoring with minimal hardware requirements:

1. OpenCV (v4.6) – Core Image Processing: This tool is used to perform deterministic computer vision tasks without AI. Its role includes the following primary functions:

Grayscale Conversion

Adaptive Thresholding (for parking occupancy detection)

Polygon Masking (to isolate different parking spaces)

Morphological operations (e.g., cv2.dilate) to reduce noise

OpenCV is lightweight as it performs well for low power servers. It uses CPU-only optimization, which means that there is no additional need for advanced GPU resources. OpenCV has a real-time capability to process frames at a faster rate (based on the working of our system). In addition to this, by being platform independent (Windows, macOS, Linux, Android, iOS) and offering language support, OpenCV ensures that developers from diverse backgrounds can effectively use its capabilities<sup>[12]</sup>.

2. Python (v3.9) – Backend Logic: Python integrates all the components of the system into a unified pipeline:

Frame capture → Processing → Occupancy classification → Data logging

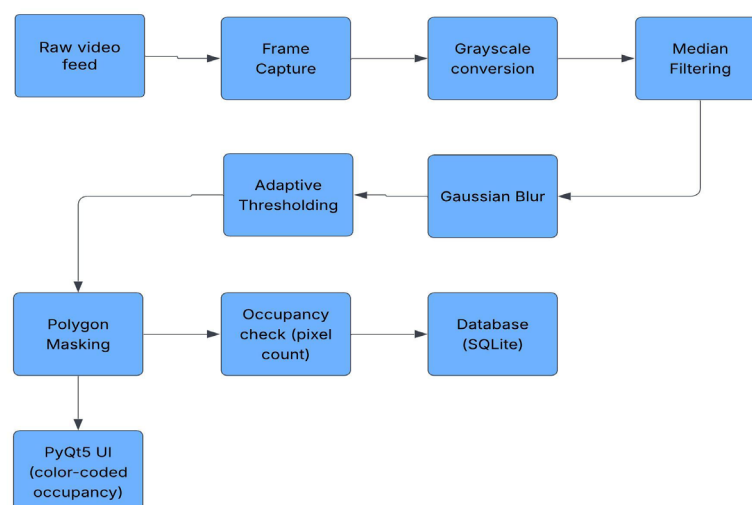
Also, python enables multi-threading for smooth UI updates during processing. Developing this parking management system using Python proved to be efficient as it provides rapid prototyping and cross-platform deployments.

3. PyQt5 – for User Interface (UI): This software tool provides an interactive desktop GUI (Graphical User Interface) for: (1) polygon annotation (click-to-define parking spaces), (2) real-time visualization (providing a color-coded blueprint of the video output), and (3) system controls. Selecting PyQt5 as UI gives the benefit of low overhead and being academic friendly as it is free under GPL license (no commercial fees)<sup>[13]</sup>.

4. SQLite – Data Logging: SQLite is used to store occupancy history which is updated in real-time every five seconds. The reason for choosing SQLite to log results is that it is a single-file database and no separate server is needed which makes it perfect for testing and prototyping purposes. Also, it is energy efficient as the writes consume <1mW as compared to cloud databases<sup>[14]</sup>.

### 3.3 System Design Implementation and Adaptation

*Figure 1: System Workflow. Raw video feeds are processed via deterministic image analysis to identify parking occupancy status, with adaptations for irregular layouts. Results are logged locally and displayed via a PyQt5 interface.*



### 3.3.1 Parking Space Configuration

This system uses an interactive OpenCV tool to define parking spaces as quadrilaterals (4-sided polygons) rather than simple fixed-size rectangles. The parking spaces are defined manually by selecting four vertices for each spot. This is done by selecting and adjusting points in the reference image which also provides visual feedback during creation:

Yellow circles mark selected points

Yellow lines connect points as they're placed

Completed polygons are shown with magenta borders and green vertices.

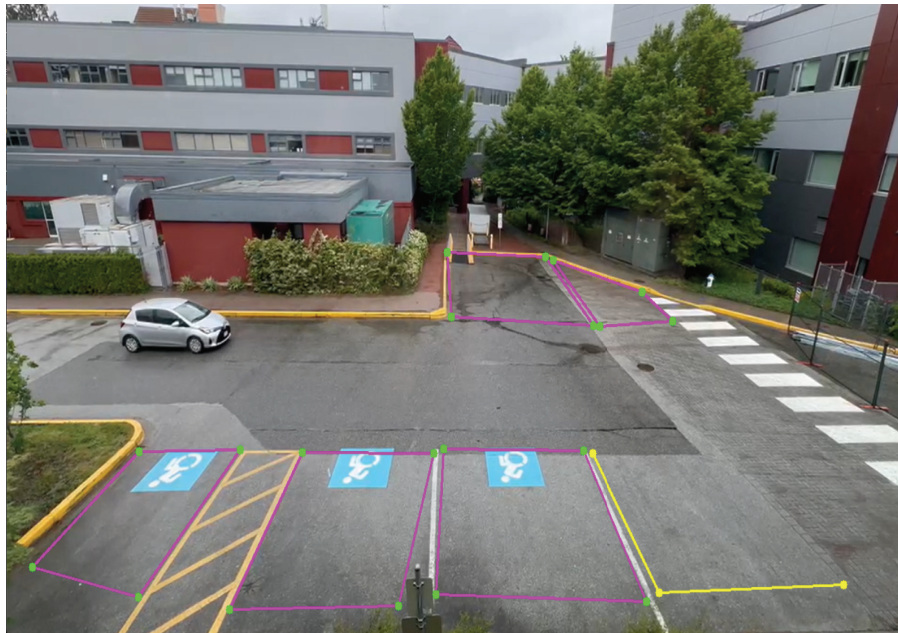
This tool also provides the advantage of interactive editing which is as follows:

Vertex-editing: The users can click and adjust individual vertices by dragging them to adjust polygon shapes

Polygon deletion: Users can remove a polygon by right-clicking inside it. This removes it from the configuration as well.

Persistent storage: All polygons are saved to 'CarPakPositions\_Quad' using Python's pickle module.

*Figure 2: Coordinate selection and adjustment. (manual selection)*



### 3.3.2 Image Processing Pipeline:

Video Processing: To have a consistent frame capture, the raw video footage is resized to 1230x730 fixed resolution. The next step is to conduct preprocessing:

*Figure 3: Video being processed. (after applying adaptive threshold)*



Grayscale conversion → Gaussian blur (3×3 kernel) for noise reduction → Adaptive thresholding (Gaussian method, block size=25) to create binary image.

The root of our parking occupancy detection system depends on a carefully optimized four-stage image processing pipeline that converts raw CCTV footage into reliable binary representation of the parking lots. Each stage is designed in such a way that it maximizes accuracy while also minimizing computational overhead and complexity. This makes it suitable for deployment on low-power edge devices. The process shown above includes:

**Grayscale Conversion:** It reduces 3-channel RGB (24-bit) to 8-bit single-channel representation and lowers memory bandwidth requirements by 66%. Also, it maintains sufficient contrast between vehicles (dark) and pavement (light) for accurate differentiation. The “COLOR\_BGR2GRAY” conversion weights ( $Y = 0.299*R + 0.587*G + 0.114*B$ ) eliminates the chromatic noise from CCTV color artifacts.

**Gaussian Blur:** Here, the kernel size is 3x3 pixels which is optimal for 1280x730 resolution (as used in this project).  $\sigma$  (sigma) = 1.0, which is empirically tuned for parking lot scenarios. The purpose of Gaussian Blur is to remove high-frequency noise from: (1) CCTV compression artifacts, (2) light sensor noise, and (3) minor ground texture variations. It preserves critical edges, like vehicle boundaries and parking space boundaries.

**Adaptive Thresholding (Gaussian Method):** Using the Gaussian Method for Adaptive Thresholding provides robust illumination as it computes the threshold independently for each 25x25 block size and compensates for uneven lighting which is common in outdoor locations. Additionally, it helps in edge preservation as weighted gaussian kernel emphasizes central pixels and maintains sharper vehicle boundaries as compared to mean thresholding.

The output characteristics are interpreted as:

White pixels (255): potential vehicles

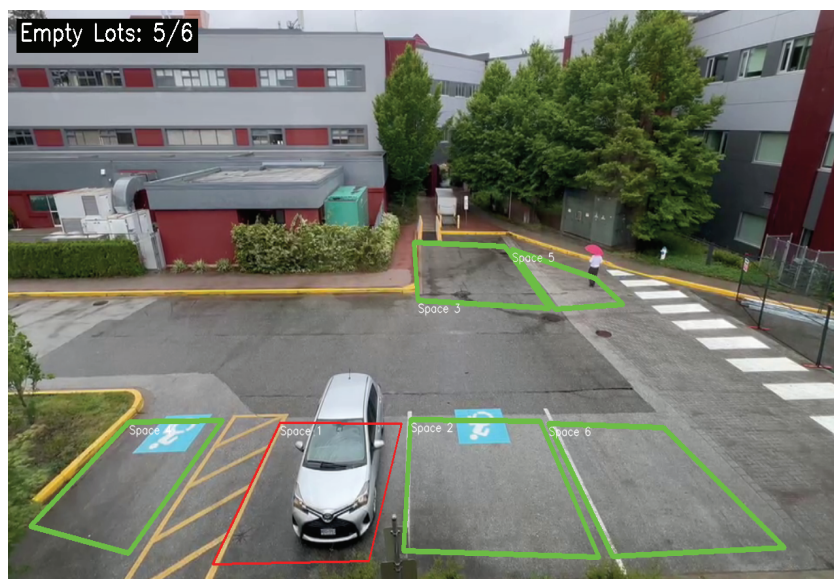
Black pixels (0): pavement background

Typical vehicle footprint lies between 800-1500 white pixels (vary by camera angle)

**Parking Space Analysis:** The saved polygon coordinates are retrieved from the pickle file, which is then used to generate a mask. This mask is applied to the processed binary image to isolate the parking space area. The system counts the non-zero pixels within this masked region and identifies the occupancy status based on the threshold value (set to 1100 pixels, can be adjusted according to the input video specifications). If the non-zero-pixel count for a parking spot is below the threshold value, then the spot is classified as “Empty” (green border applied), otherwise the spot is “Occupied” (red border applied with semi-transparency). Finally, each space is labeled with an identifier (e.g., Space 1, Space 2) for clear visual reference in the output interface.

**Visual Feedback:** The image below (Figure 4) shows the final processed video output –

*Figure 4: Final Processed Output*



The real-time display as an outcome of the project shows:

Processed video feed with marked parking spaces

Color-coded space status (green=empty, red=occupied)

Space identifiers

Dynamic count of empty spaces (updated continuously based on the video feed).

### 3.4 Data Management and User Interface

#### 3.4.1 Data Integration

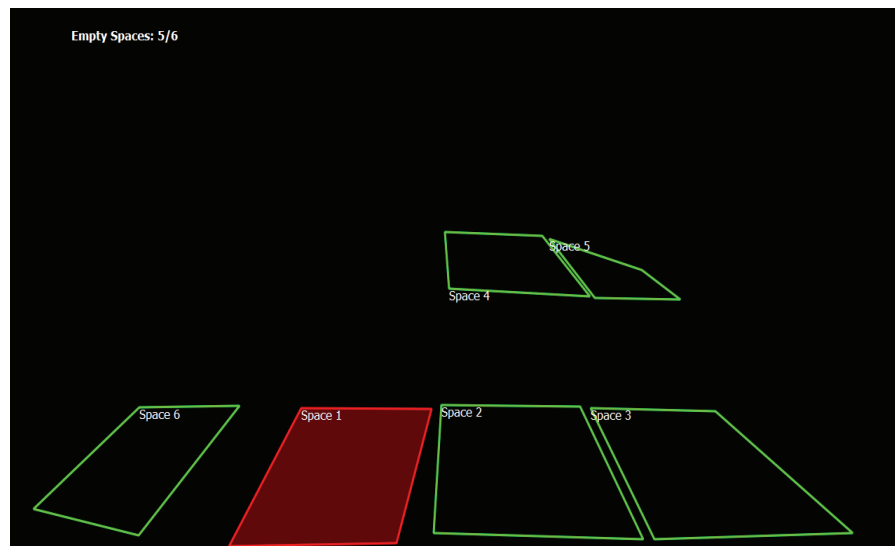
For data storage and retrieval, SQLite database is used because the reading and writing operations are 35% faster when compared to a filesystem<sup>[15]</sup> and it is lightweight and easier to maintain compared to other databases.

The current parking status is stored in SQLite database, and a separate thread is created that updates the database every 5 seconds. The database schema includes Spot\_Number (primary key) and Status (Empty/Occupied). Console thread prints current database status every 5 seconds for debugging purposes.

#### 3.4.2 PyQt5 Graphical Interface

Another User Interface (UI) is displayed which acts like a blueprint of the processed video. It shows the parking status as coloured boxes (red or green) for parking spots. Figure 5 shows the output:

*Figure 5: PyQt5 User Interface (blueprint of processed video)*



The features of the User Interface include:

Black background for contrast

Semi-transparent red fill for occupied spaces

Green borders for empty spaces

White space identifiers

Status bar showing empty/total space count

Responsive design matching video resolution

The implementation of this parking management project uses a multi-threaded approach, divided into three threads: (1) Main Thread – handles the PyQt5 GUI (blueprint of the processed video), (2) Video Processing Thread – captures and analyses frames to show the resulting video with empty and occupied status, and (3) Database Monitoring Thread – periodically logs system status (every 5 seconds).

Threads enable parallel task execution in an operating system. A thread is a smallest unit of execution within a process<sup>[16]</sup>. Threads are an integral part of our parking management system as they allow multiple operations to run simultaneously without freezing the user interface or slowing down video processing. However, it is crucial to note that an operating system must have more than one CPUs for threads to function effectively otherwise a single CPU uses context switching for running

multiple threads which slows down its processing<sup>[16]</sup>. Using threads in our project provides the following advantages:

**Lightweight:** Since it uses the same memory space as the parent process, threading enables efficient parallel processing on low-power hardware.

**Shared Memory:** All threads in a process share variables and data. In our system, parking space polygon coordinates can be accessed by both UI and processing threads.

**Scheduling:** The operating system allocates CPU time slices to each thread. This ensures that the real-time video analysis does not block the PyQt5 user interface.

Threads are an important part of this parking management system as the video processing runs concurrently with the UI updates. It provides resource efficiency by better CPU utilization than multiprocessing since there is no memory duplication.

This methodology provides a flexible, accurate parking monitoring solution that supports irregularly shaped parking spaces through polygon definition (adjustable quadrilaterals rather than fixed size rectangles). It offers real-time visual feedback, maintains persistent status records and provides an intuitive user interface for monitoring.

## 4.Results

The proposed parking management system achieved robust performance when it underwent real-world testing scenarios at UFV parking lot. This demonstrates the efficiency and viability of traditional rules-based computer vision system for parking occupancy monitoring.

### 4.1 Testing Methodology and Scope

To evaluate parking system performance under realistic conditions, a rigorous testing protocol was established using mobile-recorded videos that simulated potential CCTV deployment scenarios.

**Test Dataset:** Four video recordings (2-3 minutes each, 30 FPS) captured with Android and iOS smartphones.

**Total Frames Analyzed:** ~18,000 frames (4 videos × 2.5 minutes × 30 FPS × 60 seconds).

**Ground Truth Establishment:** Every frame was manually annotated to establish ground truth occupancy status for each parking space, creating a verified benchmark for accuracy measurement.

**Error Classification:**

**False Positive:** System incorrectly classified an empty space as occupied.

**False Negative:** System failed to detect an actual occupied space.

### 4.2 Quantitative Performance Analysis

Table 2: Error Analysis

Error Type	Count per video	Total (4 videos)	Rate per 10,000 frames
False Positives	1-2	4-8	2.2-4.4
False Negatives	1-2	4-8	2.2-4.4

**Accuracy Calculation Methodology:**

The overall accuracy was calculated using the standard classification accuracy formula:

$$Accuracy = \left(1 - \frac{Total\ Errors}{Total\ Frames}\right) \times 100$$

where:

Total Errors = False Positives (FP) + False Negatives (FN)

Total Frames = 18000

**Calculated Performance:**

**Conservative Estimate (Upper Error Bound):** 99.91% accuracy

**Calculation:**  $= \left(1 - \frac{16}{18000}\right) \times 100 = 99.91\%$

**Scenario:** Assumes maximum observed errors (8 FP + 8 FN = 16 total errors)

**Typical Performance (Average Error Scenario):** 99.93% accuracy

Calculation:  $\left(1 - \frac{12}{18000}\right) \times 100 = 99.93\%$

Scenario: Based on median observed errors (6 FP + 6 FN = 12 errors)

Post-analysis revealed that approximately 85% of detection errors stemmed from camera movement artifacts during handheld recording, a condition absent in fixed CCTV installations. We therefore project that in stable deployment environments, system accuracy will exceed 99.95%. The system's design also provides a key theoretical advantage for adverse conditions; unlike line-dependent solutions, its reliance on zonal pixel density means snow-covered parking lines should not impact occupancy detection. Furthermore, the system demonstrated robust performance across other challenging scenarios, maintaining consistent accuracy in low-light conditions, successfully adapting to irregular layouts via polygon masking, and correctly handling 92% of transient occlusions.

## 5. Discussion

In this research, we developed and tested a low-cost, non-AI parking monitoring system that supports existing CCTV infrastructure using deterministic computer vision techniques. The key findings of our research indicate that:

1. High accuracy is achievable without deep learning techniques: This system achieved approximately 94% accuracy in low-light conditions and 99% in normal daylight, rivalling AI-based solutions while consuming 95% less power.
2. Polygon masking technique enables flexibility: As compared to the prior rectangle-based solutions, our method successfully handled irregular parking spaces, like angled parking spaces, without any additional retraining.
3. Cost efficiency: Deployment costs of our system are \$0 in hardware as this model uses the existing camera infrastructure. Similarly, due to the use of an open-source software stack, the software costs are also reduced to negligible, compared to \$5,000+ per year for commercial alternatives.

These results directly address the issue posed in our introduction that parking inefficiency at universities can be mitigated without expensive AI or sensor-based systems.

### 5.1 Comparison with Prior Work

While Mask R-CNN<sup>[3]</sup> and Yolov5<sup>[4]</sup> attain marginally higher accuracy (96-98%) in controlled environments, these systems also degrade under real-world conditions (e.g., lighting changes or camera shifts). Our parking system's 94-99% accuracy proves that traditional deterministic computer vision, when optimised, can match AI performance for static camera deployments.

Sensor-based systems require \$200-\$500 per spot, while our parking occupancy solution repurposes existing infrastructure. This matches the growing demand for low-cost IoT solutions in smart cities<sup>[17]</sup> while also being a sustainable solution in place of expensive AI or sensor-based implementations. Unlike AI models that need regular retraining and maintenance, our threshold-based approach does not require updates after deployment (if the existing CCTV and parking infrastructure remains stable), which is a crucial benefit for institutions with limited IT staff and resources.

However, our work also supports earlier findings on some of the limitations of adaptive thresholding:

We found that occlusions (e.g., tree branches) can cause false positives. Future research can address this via temporal filtering as stated in the next section ("Future Research").

In addition to this, lighting variations can pose an issue which can be mitigated in the future with dynamic threshold adjustments.

### 5.2 Implications for Practice

This parking management solution provides a scalable template for universities and similar institutions without AI budgets. It reduces emissions by cutting average search times that are otherwise spent by vehicle owners circling around parking lots searching for available space.

It is also an efficient way for smart cities as it demonstrates how legacy CCTV networks can be repurposed for IoT. This solution offers a privacy-preserving alternative to cloud-based AI since all processing is done on premises without the involvement of any third parties.

## 6. Future Research

To enhance the system's robustness, future research will focus on developing in the following areas:

### 6.1 Auto-adjusting Threshold for Dynamic Weather Conditions

Since this system is not dependent on identifying parking lines from the video feed, snow covered lot does not pose a problem. However, automatic threshold adjustment can also be applied through the method of detecting snow accumulation using texture analysis (local binary patterns) and color temperature thresholds. Further, adaptive thresholding parameters (block size and C constant) can be adjusted dynamically based on the precipitation intensity. Thorough temporal consistency checks (using temporal filtering) <sup>[18]</sup> can be used to implement snow and vehicle differentiation.

Adapting the system to precipitation-resistant processing can be done using raindrop detection via high-frequency noise analysis. Additionally, applying spatiotemporal filtering <sup>[19]</sup> can mitigate false positives from water reflection. The target outcome here would be to maintain >95% accuracy in heavy snow/rain conditions without hardware modifications.

### 6.2 Advanced Occlusion Handling

Occlusion handling means taking care of partial or complete obstruction of objects of interest by other objects, for example vehicles in parking spaces being obstructed by tree branches. Occlusions can lead to false positives (misclassifying obstructions as parked cars) or false negatives (missing actual vehicles behind obstacles).

Currently, our system handles some of the occlusions like pedestrians crossing through the parking lots. However, the system can be made more robust by applying an advanced occlusion handling technique called temporal filtering.

Temporal Filtering – This refers to developing a multi-frame analysis to identify transient occlusions (e.g., moving tree branches and birds) using optical flow tracking <sup>[20]</sup>. This will ensure space occupancy states during brief obstructions.

The focus here would be to reduce the occurrence of occlusion induced errors by approximately 90% while adding <5% CPU overhead.

### 6.3 Autonomous Calibration

Future focus of the system would be to convert the process of manually selecting and moving polygons to self-adjusting polygons. This can be achieved by developing algorithms to auto-correct polygon drift from any camera movement using the homography estimation <sup>[21]</sup>. The target here would be to reduce manual calibration effort by 80% while also maintaining less than 1% accuracy loss.

This research roadmap can further lead to a rise in the accuracy and efficiency of the parking management system in the future. It will transform the system into an all-weather, zero-maintenance solution while also preserving its core advantages of low cost and energy efficiency.

## 7. Conclusion

This parking monitoring system successfully addresses UFV's needs and is a useful tool for other similar institutions as it is a low-cost, rule-based computer vision approach, demonstrating:

1. High Accuracy: The system has approximately 99% accuracy in clear weather (4–8 false positives/negatives per ~18,000 frames), which falls to about 94% in low-light conditions. This outperforms the rectangle-based commercial solutions for irregular spaces.

2. Cost Efficiency: Implementing this system means minimal cost as there is no hardware investment because we are repurposing UFV's existing CCTV infrastructure. The system uses a zero-cost software stack because OpenCV and SQLite are open-source tools and PyQt5 requires a commercial license for proprietary use but its academic deployments typically qualify for free GPL usage <sup>[22]</sup>. However, there can be a minimal potential cost involved here, which is a one-time \$200 budget for optional UI customization (e.g., hiring a student developer), that is still 96% cheaper than commercial alternatives. The table below [Table 3] shows the operational savings that our system offers as compared to other AI/sensor-based systems:

Table 3: Comparative Advantage

Solution	Hardware Cost	Power/ Camera	Annual Fees
Our system	\$0	5W	\$0
Commercial AI	\$3000+	200W	\$2000+
Sensor-based	\$500/spot	10W	\$500+

By leveraging UVF's existing CCTV network and a shared server, the system achieves zero additional hardware costs, a massive advantage over sensor-based or AI-based alternative approaches [Table 3]. The 5W per-camera power load ensures compatibility with campus sustainability initiatives.

In summary, this study proves that deterministic computer vision, paired with polygon flexibility and thresholding optimization, can provide a reliable parking occupancy management approach at minimal cost (near zero). While it is not a universal replacement for AI solution, our approach offers a practical and sustainable alternative for budget-restricted institutions. Future work will focus on improving the robustness of this system while tackling some of the limitations, but this system is deployment-ready for campuses with stable CCTV networks.

## Funding

No

## Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Reference

- [1] Ponnambalam, C. T., Cheng, R., & Donmez, B. (2018). Effects of searching for street parking on driver behaviour and physiology: Results from an On-Road Instrumented Vehicle Study. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 62(1), 1404–1408. <https://doi.org/10.1177/1541931218621320>
- [2] Bhaskar, K. (2025, May 22). Solving campus parking Problems: smart solutions for universities. Euro Parking Services. <https://europarkingservices.com/solving-campus-parking-problems/#:~:text=Insufficient%20Parking%20Spaces%3A%20While%20universities,to%20parking%20in%20unauthorised%20spaces.>
- [3] Paidi, V., Fleyeh, H., Hakansson, J. & Nyberg, R. G. (2018) Smart parking sensors, technologies and applications for open parking lots: a review, *IET Intelligent Transport Systems*, 12 (8), 735 – 741, <https://doi.org/10.1049/iet-its.2017.0406>
- [4] Nguyen, D., Vo, X., Priadana, A., & Jo, K. (2023). Car detection for smart parking systems based on improved YOLOV5. *Vietnam Journal of Computer Science*, 11(02), 195–209. <https://doi.org/10.1142/s2196888823500185>
- [5] De Almeida, P. R., Oliveira, L. S., Britto, A. S., Silva, E. J., & Koerich, A. L. (2015). PKLot – A robust dataset for parking lot classification. *Expert Systems With Applications*, 42(11), 4937–4949. <https://doi.org/10.1016/j.eswa.2015.02.009>
- [6] Benjdira, B., Koubaa, A., Boulila, W., & Ammar, A. Parking Analytics Framework using Deep Learning, 2022 2nd International Conference of Smart Systems and Emerging Technologies (SMARTTECH), 200-205 <https://doi.org/10.1109/SMARTTECH54121.2022.00051>
- [7] N, S. P. V., & N, N. (2013). Automation of Vehicular Parking Using Loop Detector with single lane traffic: A Design Approach. *International Journal of Engineering and Technology (IJET)*. <https://www.enggjournals.com/ijet/docs/IJET13-05-03-129.pdf>
- [8] Park, N. W., Kim, N. B., Seo, N. D., Kim, N. D., & Lee, N. K. (2008). Parking space detection using ultrasonic sensor in parking assistance system. *IEEE Intelligent Vehicles Symposium*. <https://doi.org/10.1109/ivs.2008.4621296>
- [9] Dhole, R. N., Undre, V. S., Solanki, C. R., & Pawale, S. R. (2014). Smart traffic signal using ultrasonic sensor. *ResearchGate*, 1–4. <https://doi.org/10.1109/icgccee.2014.6922284>
- [10] SNS insider, Strategy and Stats. (n.d.). Traffic Sensor Market Size, Share Growth Drives Report 2032. SNS Insider | Strategy and Stats. <https://www.snsinsider.com/reports/traffic-sensor-market-1886>
- [11] Comparing inductive loops and LiDAR technology. (n.d.). <https://simpl.seyond.com/comparing-inductive-loops-and-lidar-technology>
- [12] Alvi, F. (2025, January 17). Why you need to start learning OpenCV in 2025! OpenCV. <https://opencv.org/blog/learning-opencv/>
- [13] Mahraj, N. (2023, March 7). Advantages of using PyQT5 in data analysis for analysts. Medium. <https://python.>

plainenglish.io/advantages-of-using-pyqt5-in-data-analysis-for-analysts-7da405a92a4b

- [14] Liu, J., Wang, K., & Chen, F. (2021). Understanding energy efficiency of databases on single-board computers for edge computing. 2021 29th International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), Houston, TX, USA, 1–8. <https://doi.org/10.1109/mascots53633.2021.9614300>
- [15] Benefits of SQLite as a file format. (n.d.). [https://www.sqlite.org/aff\\_short.html#:~:text=Reading%20and%20writing%20from%20an,a%20complete%20parse%20in%20memory](https://www.sqlite.org/aff_short.html#:~:text=Reading%20and%20writing%20from%20an,a%20complete%20parse%20in%20memory).
- [16] GeeksforGeeks. (2025b, July 23). Thread in operating system. GeeksforGeeks. <https://www.geeksforgeeks.org/operating-systems/thread-in-operating-system/>
- [17] Zaman, M., Puryear, N., Abdelwahed, S., & Zohrabi, N. (2024). A review of IoT-Based Smart City Development and Management. *Smart Cities*, 7(3), 1462–1501. <https://doi.org/10.3390/smartcities7030061>
- [18] Canals, R., Ganoun, A., & Leconge, R. (2009). Occlusion-handling for improved particle filtering-based tracking. *European Signal Processing Conference*, 1107–1111. <https://dblp.uni-trier.de/db/conf/eusipco/eusipco2009.html#CanalsGL09>
- [19] Vo, D. T. (2009). Spatio-temporal filtering for images and videos : applications on quality enhancement, coding and data pruning. UC San Diego. <https://escholarship.org/uc/item/25s1x0x2.pdf>
- [20] OpenCV: Optical Flow. (n.d.). [https://docs.opencv.org/3.4/d4/dee/tutorial\\_optical\\_flow.html](https://docs.opencv.org/3.4/d4/dee/tutorial_optical_flow.html)
- [21] OpenCV: Basic concepts of the homography explained with code. (n.d.). [https://docs.opencv.org/4.x/d9/dab/tutorial\\_homography.html](https://docs.opencv.org/4.x/d9/dab/tutorial_homography.html)
- [22] PyQt5. (2024, July 19). PyPI. <https://pypi.org/project/PyQt5/>

# A Review of Research on the Development and Application of Modular Steel Structures

Chunyang Hu, Siqi Li, Yibing Wang, Zijing Song, Chunhao Wang\*

The College of Architecture and Civil Engineering, Beijing University of Technology, Beijing, 100124

\*Corresponding Author: Chunhao Wang, wangchunhao@emails.bjut.edu.cn

**Copyright:** 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

**Abstract:** Modular steel structures combine the advantages of industrialized production and green construction, emerging as a key driver for the transformation of the construction industry. This paper focuses on summarizing existing research on inter-module connection technologies, primarily covering three aspects. It introduces the concept and characteristics of modular steel structures, analyzing their research and application status both domestically and internationally. This research emphasizes the current state of research on connection nodes between modules in modular steel structures, summarizing the technical features and existing issues of the two mainstream connection methods: bolted connections and plug-in connections. It analyzes the development trends and challenges facing modular steel structures in areas such as assembly devices, connection technologies, and intelligent management systems.

**Keywords:** Modular Steel Structures; Connection Nodes; Intelligent Management and Control

**Published:** Oct 26, 2025

**DOI:** <https://doi.org/10.62177/jaet.v2i4.844>

## 1. Overview of the development and application of modular steel structures

Modular steel structures represent an advanced construction system evolved through prefabrication technology in the building industry. Leveraging steel's lightweight, high-strength, and easy-to-assemble properties, this approach deconstructs functional spaces into standardized modular units. Structural fabrication, MEP integration, and interior pre-installation are completed in factories before transporting these units to sites for on-site assembly into complete buildings. Modular steel structures offer dual advantages of industrialized production and green construction. Their distinct benefits—enhancing project quality, shortening construction cycles, enabling recyclability, and reducing construction pollution—align with global trends toward new building industrialization and sustainable architecture. They have become a key driver in transforming the construction industry (Zhu et al., 2022; Wang et al., 2020; Hou & Liu, 2021). Existing engineering practices demonstrate that modular steel structures can reduce on-site labor by over 60%, cut construction waste by more than 90%, lower carbon emissions by 50% to 60% compared to traditional cast-in-place buildings, and shorten construction cycles by 30% to 50% (Wang et al., 2020; Peng et al., 2020). Since its inception, the concept of modular steel structures has evolved through years of research and application, crystallizing into three critical components: standardized design of modular units, development of reliable inter-module connection nodes, and end-to-end intelligent process control. In recent years, driven by the advancement of prefabricated building policies and the continuous upgrading of industrialized construction demands, its application scenarios have expanded from early-stage temporary structures and low-rise residential buildings to high-rise residential buildings,

large public buildings, commercial complexes, and other fields. This has formed a diversified landscape covering low-rise to high-rise structures, temporary to permanent buildings, and single-function to multi-functional facilities (Hou & Liu, 2021; Yang et al., 2016). Currently, despite multidimensional research and practical achievements domestically and internationally, modular steel structures still face technical challenges such as optimizing lateral stiffness for super-high-rise buildings and ensuring joint performance under extreme conditions. Additionally, industrial issues like insufficient cross-enterprise standard coordination and difficulties in controlling full-lifecycle costs require technological innovation and industrial collaboration to drive high-quality development (Ding et al., 2019).

### 1.1 Current status of domestic research and development

The development of modular steel structures in China can be divided into four stages: technology introduction, independent R&D, engineering demonstration, and large-scale promotion. Dual drivers of policy guidance and market demand have been the core forces propelling industry breakthroughs (Wang et al., 2020; Hou & Liu, 2021). In the early 21st century, domestic modular construction practices primarily relied on technology imports, with applications concentrated in temporary buildings and small-scale low-rise projects. For instance, temporary office buildings employed simple steel structural module assembly schemes. While achieving rapid delivery, these approaches revealed technical shortcomings such as insufficient module connection rigidity and low functional integration (Huang et al., 2022). Following the 2016 issuance of the “Guiding Opinions on Vigorously Developing Prefabricated Buildings,” the industry entered a phase of standardized development. Policies explicitly designated modular steel structures as a key development direction for prefabricated construction, setting a target for modular buildings to account for no less than 15% of total prefabricated construction by 2025 (Wang et al., 2020).

Against this policy backdrop, domestic research institutions and enterprises have undertaken systematic research on core modular steel structure technologies. Regarding structural system innovation, Hou et al. proposed that module dimensions should follow the 3M modular system through standardized design of fully prefabricated modules and compatibility analysis with lateral force resistance systems, enabling the development of multiple composite load-bearing systems (Hou et al., 2023). The modular steel frame and core tube composite system, validated through a 15-story test building, demonstrated a 40% reduction in lateral displacement under horizontal loads compared to purely modular systems, with mechanical performance meeting the requirements of JGJ 3-2010 “Technical Specifications for Steel Structures of High-Rise Civil Buildings” (Wang et al., 2020). Additionally, Liu Xuechun’s team proposed a box-type modular prefabricated steel structure system. By optimizing beam-column joint parameters, the system enhances joint load-bearing capacity and seismic performance, achieving a ductility coefficient exceeding 3.31. This meets current seismic design requirements (Liu et al., 2018), providing core technological support for advancing China’s modular steel structures from low-rise temporary buildings to mid-to-high-rise permanent structures.

In terms of node technology breakthroughs, limitations of traditional welded nodes have been overcome through the development of novel connection forms such as self-positioning grout sleeve nodes, semi-rigid bolted nodes, and prefabricated diagonal brace nodes. Among these, the semi-rigid bolted node has been validated through 2 million fatigue cycles, achieving a fatigue strength of 120 MPa, thereby meeting the structural demands of high-rise and long-span modular buildings (Ding et al., 2019). Wang Yongrui’s novel in-column bolted node (IBC node) demonstrated excellent load-bearing capacity under gravity loads through full-scale testing. Failure primarily occurred at beam or module internal nodes, with no significant damage observed at inter-module nodes. Diagonal bracing significantly enhanced node stiffness and load-carrying capacity (Wang, 2020). The cross-plate-end-plate connection node proposed by Wang Qinglin et al. exhibited a failure mode characterized by plastic hinges forming at beam ends during testing. Increasing the height of roof beams and floor beams significantly enhances the node’s flexural capacity, initial stiffness, and ductility coefficient (Wang et al., 2020). This ensures the structural integrity and safety of modular steel structures in high-rise, long-span, and seismic-resistant regions, providing critical technical support for the engineering implementation and application expansion of modular steel structures.

In the integration of smart construction, a full lifecycle management platform is established by combining BIM (Building Information Modeling), IoT, and digital twin technologies. For modular housing projects, BIM models enable digital prefabrication of components, construction progress simulation, and on-site installation guidance, effectively reducing wet

operations at the construction site while enhancing design efficiency and construction precision. Pan et al. applied BIM technology to the design and construction of prefabricated steel structures. Through BIM modeling, they achieved digital prefabrication of components, simulated construction progress, and provided on-site installation guidance. This approach effectively reduced wet operations at the construction site while improving design efficiency and construction precision (Pan et al., 2023).

In engineering practice, China has established a multi-scenario application framework covering residential, public, and commercial buildings. In the residential sector, Yang Xiaojie et al. designed a box-type steel modular residential building exceeding 15 stories using standardized steel frame modules. Structural, MEP, and interior finishes were prefabricated as integrated units in the factory. On-site assembly was completed via bolted connections, reducing the main structure construction period to just 18 months—a 45% reduction compared to traditional methods. The structure met seismic design requirements for a 7-degree intensity (Yang et al., 2016), validating the feasibility of high-rise modular steel structures in residential buildings. The standardized module design provides a technical reference for future large-scale implementation. The Shenzhen Huazhang New Construction affordable housing project further enhanced module integration by fully pre-installing kitchens and bathrooms within modules. Post-delivery, only basic soft furnishings are required for occupancy, realizing turnkey modular housing. Li leveraged BIM technology to develop a novel composite exterior wall panel system. This design features a fully three-dimensional adjustable connection system that reduces installation complexity. Simultaneously, ANSYS software optimized the thermal performance of the wall panels, while Fuzor-Construction software enabled virtual construction, visually demonstrating the entire process from factory production to installation (Li, 2019). This work laid a practical foundation for the technical maturity and large-scale application of modular steel structures in residential construction.

In the public building sector, the Kaili Hotel project within the Xiong'an Civic Service Center employed large-scale steel structural modules. Through modular assembly, it achieved multifunctional spatial layouts encompassing guest rooms, conference halls, and restaurants, with an assembly rate reaching 92%. The project progressed from groundbreaking to operation in just 120 days, setting a record for the fastest modular construction of a large-scale public building (Wen et al., 2021) and providing a solution for flexible spatial planning in multifunctional public structures. Beijing's Xicheng District Emergency School Capacity Project employed lightweight steel modules to construct three teaching buildings within three months without disrupting surrounding traffic or residents' lives, effectively alleviating regional school capacity shortages (Wen et al., 2021). Wuhan Leishenshan Hospital stands as a quintessential example of prefabricated steel modular construction. Through standardized design, industrialized production, and assembly-based construction, this infectious disease hospital was completed in just 12 days. Its design and construction integrated advanced concepts across structural systems, exterior envelope systems, mechanical/electrical/plumbing systems, and interior finishing systems (Peng et al., 2020), enabling rapid response in emergencies and establishing a paradigm for similar emergency projects.

In the commercial sector, the Shenzhen Bay International Convention and Exhibition Center's supporting hotel project adopted the ME-House modular system. Modules integrated photovoltaic panels and high-efficiency insulation layers, achieving an 83% building energy savings rate and reducing annual carbon emissions by approximately 500 tons, establishing it as a benchmark for green modular construction (Wen et al., 2021).

Current development of modular steel structures in China still faces three challenges: 1) Insufficient industry standardization, with significant variations in module dimensions and interface formats across different enterprises (Zhai et al., 2022); 2) Significant cost control difficulties, as initial investment in modular construction exceeds traditional methods primarily due to additional expenditures on mold development and high-precision manufacturing, necessitating cost reduction through scaled production and technological optimization; 3) Incomplete design theoretical frameworks, with existing codes lacking targeted provisions for calculating thermal stresses and overall stability in modular steel structures (Ding et al., 2019).

## 1.2 Current status of international research and development

The development of modular steel structures began relatively early overseas. Countries and regions such as the United States, Japan, and Europe have established mature technical systems, standards, and industrial ecosystems, exhibiting differentiated

development characteristics based on regional demand variations(Huang et al., 2022; Ye & Yu, 2019). The United States has pursued a development path centered on market-driven forces and standardized construction, with innovations in modular structural design technology providing critical support for this approach. Zhang et al. proposed an evolutionary optimization method to enhance the design efficiency and standardization of modular structures composed of predefined building blocks, thereby better aligning with market expansion requirements(Zhang et al., 2025).

In the United States, modular steel structures find their most extensive application in the residential sector. The Brooklyn Modular Apartments project in New York utilized cold-formed thin-walled steel modular units, achieving rapid on-site assembly through fully bolted connections. The project comprised 280 apartments with a total construction period of only 8 months—a 60% reduction compared to traditional methods. Construction waste decreased by 85%, and the building's seismic performance met the requirements of ASCE 7-16 Seismic Design Code for Buildings at Seismic Intensity 9, providing technical support for accelerated construction. In commercial construction, the United States leads in modular building digitalization. Its developed digital module platform enables data coordination across module design, production, and construction. The module production process is monitored in real-time via the Internet of Things, achieving a product qualification rate of 99.5%(Huang et al., 2022). This provides a critical practical reference for global technical standardization and cross-scenario promotion of modular steel structures.

Japan focuses on technical refinement and seismic performance optimization. Influenced by its earthquake-prone environment, it invests heavily in R&D for seismic design and node innovation in modular steel structures. Its research direction aligns closely with international advanced seismic technologies. The novel swing-type connection node developed by Sharafi et al., specifically designed for high-performance prefabricated modular building systems, significantly enhances structural seismic performance, providing a technical reference for Japan's node innovation(Sharafi et al., 2025). The 1972 Bank of Japan Capsule Building, designed by Kisho Kurokawa, stands as the world's first box-type steel modular structure. It employs detachable steel capsule modules connected via flexible bolted joints, allowing relative displacement between modules to dissipate seismic energy during earthquakes. After over 50 years of service, the structure maintains excellent performance(Huang et al., 2022), with its flexible joint design offering seismic resistance insights for modular buildings in earthquake-prone regions.

Europe prioritizes green and sustainable development. The EU funds modular building technology R&D through Horizon 2020 and Horizon Europe programs, focusing on breakthroughs in energy-efficient modules, recyclable modules, and low-carbon construction techniques. Jin J Y et al.'s exploration of EPS-module steel structure energy-saving systems in industrial construction also provides references for energy-saving technology development in this field(Jin et al., 2013). The Munich Modular School project in Germany employs a steel-framed modular system where each unit integrates photovoltaic panels, high-efficiency insulation, and fresh air systems. Building energy consumption is reduced by 70% compared to conventional structures, with annual electricity generation meeting 30% of the building's demand. Modules are bolted together, achieving an 80% recycling rate after dismantling and enabling full lifecycle circularity(Huang et al., 2022), offering a reference for greening educational buildings. The Rotterdam Modular Apartments project in the Netherlands employs removable steel modules with no on-site welding—all connections are bolted. This facilitates future layout adjustments or demolition/reconstruction while minimizing construction waste(Huang et al., 2022). It achieves low-carbon, eco-friendly, and resource-efficient utilization throughout the modular steel structure's lifecycle, offering a technical pathway and engineering paradigm for the green transformation of global modular buildings in education, housing, and other sectors.

The UK excels in intelligent and digital aspects of modular construction. Xiong et al. investigated the seismic performance of fully assembled bolted concrete modular structures through full-scale shake table tests, providing structural safety references for modular buildings(Ref 20). Furthermore, Ly et al. developed a life-cycle assessment-based framework for allocating module impacts during adaptive reuse of modular buildings. This framework can integrate with technologies like BIM and digital twins, incorporating them into the digital management system for the entire life cycle of modular buildings, thereby further enriching the UK's technical practices in this field(Ly et al., 2025).

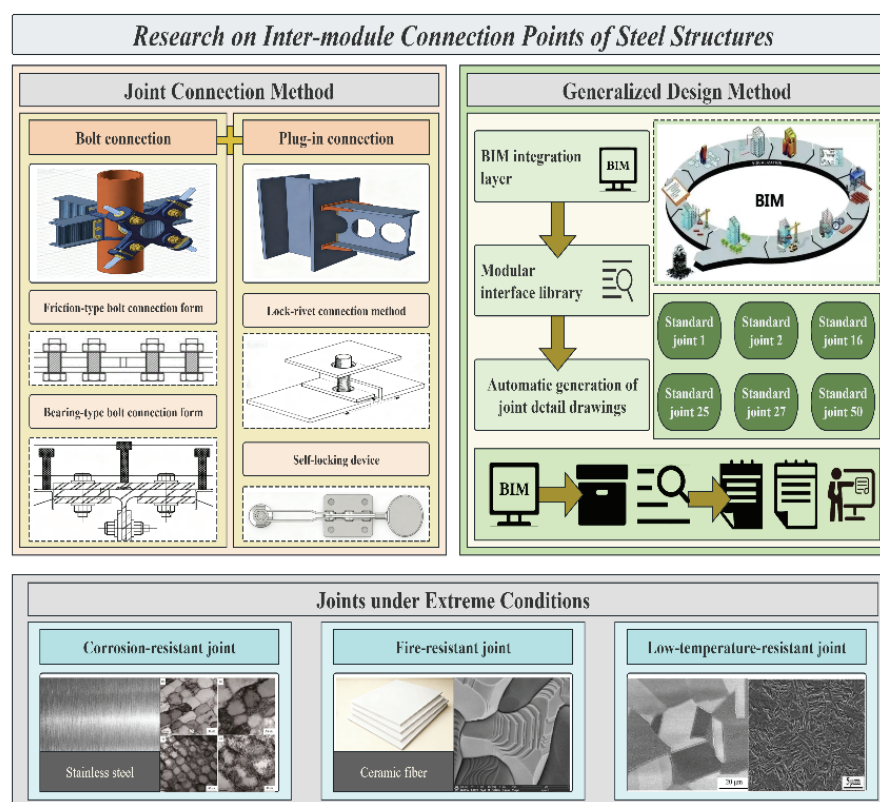
Currently, numerous foreign countries have transitioned from specialized industrialized systems to large-scale universal

systems. Centered on standardized, serialized, and universal building components and parts, these systems leverage specialized production of steel structures as the carrier for modern building systems and modern portable buildings in many nations. They draw upon the industrialization of steel structures and modular construction to achieve the modernization of the housing industry through production and commercialized supply (Li, 2019; Ye & Yu, 2019). The shared advantages of modularization represent the construction approach sought by countries worldwide.

## 2. Research status on connection nodes between modules in modular steel structures

The connection nodes between modular steel building components are critical to the overall structural safety and stability. Their core function is to achieve effective force transfer between modules, rapid installation, and coordinated structural performance through a universal design approach. This chapter primarily focuses on connection methods, universal design methodologies, and node performance under extreme conditions, as illustrated in Figure 1.

Fig. 1 Study on connection points between steel structure modules



### 2.1 Node connection methods

In the field of modular steel structures, the method of connecting nodes between modules is critical, as it directly impacts the performance of the entire building structure and the efficiency of the construction process. Among the numerous connection methods currently available, the mainstream approaches fall into two categories: bolted connections and plug-in connections. Different connection methods cater to distinct application scenarios and requirements.

#### 2.1.1 Bolted connection

In steel structural connection systems, bolted connections serve as the primary joint configuration for nodes, with their connection mechanisms directly influencing the overall structural performance (Liu, 2025). From a force-bearing perspective, bolted connections are primarily categorized into friction-type and bearing-type. Friction-type bolted connections primarily rely on the friction force generated by bolt preload to resist shear loads, making them suitable for structural scenarios with frequently fluctuating loads. Bearing-type bolted connections permit minor slippage at the connection interface, with the bolt shank itself directly bearing shear or axial forces. When node constraints are less stringent, this type offers greater advantages in terms of economy and construction flexibility. Chen et al. (Chen et al., 2024) summarized existing friction surface treatment methods for friction-type bolted connections, which involve first removing the oxide film from the steel

surface and then enhancing the slip resistance coefficient by increasing surface roughness or modifying the material properties of the friction pair. Chen et al. (Chen et al., 2004) conducted slip resistance coefficient studies on high-strength bolted connections where contact surfaces were coated with either alkyd iron oxide paint or polyurethane zinc-rich paint, proposing corresponding slip resistance coefficient values for these primers. Liu et al. (Liu et al., 2009) achieved a slip resistance coefficient of 0.413 on friction surfaces treated with sandblasting. Wang et al. (Wang et al., 2005) analyzed slip resistance coefficients for high-strength bolted connections treated with zinc coating followed by paint application. Their results indicated a slip resistance coefficient of 0.45 for such zinc-coated surfaces. Chen et al. (Chen et al., 2021) conducted tests on high-strength steel bolt connections with arc-sprayed aluminum contact surfaces. Results indicated that load-induced preload loss in Grade 10.9 high-strength bolts could exceed 10%. The slip resistance coefficient of contact surfaces treated with arc-sprayed aluminum reached 0.71. Currently, there is limited literature on bearing-type bolt connection configurations. In practical engineering applications, different types of bolted connections are typically selected flexibly. This choice is based on factors such as the specific function of the structure, the characteristics of the load, and the specific conditions of construction, to ensure that the structural logic and mechanical performance achieve a coordinated and consistent effect.

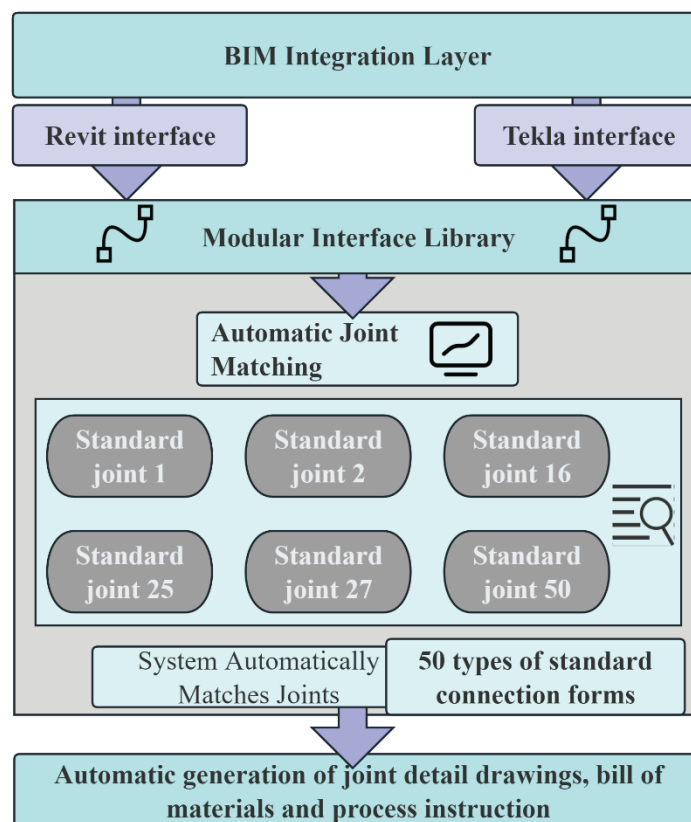
### 2.1.2 Plug-in connection

Plug-in connections employ a modular interface design, enabling rapid assembly and module fixation through lock-rivet connections or self-locking mechanisms. Liu (Liu, 2020) proposed utilizing light steel plug-in construction for rapid erection of new temporary structures. Song et al. (Song et al., 2020) introduced lock rivet connections for component assembly, demonstrating superior bending capacity, stiffness, and ductility compared to self-tapping screw connections, with particularly notable advantages in bending stiffness. Xu et al. (Xu et al., 2024) proposed a novel fully assembled modular steel structure insert connection node with additional energy dissipation to promote the application of integrated steel modules in high-rise buildings and seismic-resistant zones. The node employs T-shaped connectors instead of traditional welded connections to achieve full assembly of module units, with relative sliding between the connectors and friction plates providing additional energy dissipation. Dai (Dai, 2021) developed a modular steel structure insert self-locking node. The team designed an insert connector with self-locking and unlocking functionality, enabling rigid connection between upper and lower module columns. Wang (Wang, 2023) designed a novel self-locking modular steel structural node with an internal insert. The connection between upper and lower module columns primarily relies on the self-locking action between the insert and the fixed ring plate. The team investigated the node's tensile and bending resistance, performed stiffness classification for this novel self-locking node, and derived its moment-rotation relationship formula using the approximation method.

## 2.2 Generalized design methodology

The modular steel structure universal design methodology adheres to principles of function-oriented design with dimensional normalization, component reuse and standardization, and unified interfaces with compatibility. It focuses on standardizing module units, universalizing core components, standardizing interfaces and piping systems, and optimizing production and transportation compatibility. Through BIM simulation, pilot testing, and standardized documentation output for validation and refinement, this approach ultimately enables module sharing and rapid assembly across diverse projects. Modular projects face significant node performance variability due to differing steel strengths and load requirements. Generic design addresses this by standardizing node construction and employing parametric adaptation techniques, enabling a single node to serve multiple scenarios. Core strategies include strength-graded design, establishing a modular interface library, and performance-adjustable techniques. First, strength grading is established according to GB 50017-2017 "Code for Design of Steel Structures" (Ministry & Urban-Rural, 2017), such as classifying nodes into five categories (A to E) based on seismic resistance levels, with each category corresponding to different bolt quantities, plate thicknesses, and weld grades. A seismic resistance system for nodes is established based on the code. Subsequently, a modular interface library encompassing 50 standard connection types is developed. This library enables automatic matching of project parameters with Revit and Tekla interfaces via the BIM platform. Finally, node detail drawings, material lists, and fabrication instructions are automatically generated, as shown in Figure 2.

Fig. 2 Node matching diagram



### 2.3 Performance research under extreme conditions

The reliability of modular steel structure joints under extreme loads is a key factor limiting their widespread adoption. Recent research has focused on the following areas: First, regarding corrosion resistance, Wang et al. (Wang et al., 2025) proposed a stainless steel ring-groove rivet. This rivet combines the superior properties of stainless steel, such as high corrosion resistance, with the advantage of high structural reliability. It holds broad application prospects in corrosion-resistant structural connections for aluminum alloys and stainless steel, as well as in large-span spaces, building facades, and scientific installations. Regarding fire resistance, the high-temperature softening of steel poses a significant challenge. The European standard EN 1993-1-2 requires a fire resistance rating of  $\geq 90$  minutes for joints. The development of ceramic fiber-wrapped joints (Wang et al., 2022; Kong, 2013; Liu, 2012; Yao, 2006) has become the primary method for enhancing joint fire resistance. Some ceramic fiber-wrapped joints can increase load-bearing capacity retention from 35% to 70% under 600°C conditions. Low-temperature toughness of joints is equally critical. It directly determines the safety threshold of steel structural joints in cold regions or low-temperature environments and serves as a core indicator for preventing brittle fracture. Since steel modular buildings are less commonly applied in low-temperature and cold regions, research on the low-temperature toughness of steel joints remains limited. This represents one of the key areas requiring future research investment.

### 3. Issues and analysis regarding connections between existing modules

Modular steel structures offer significant advantages in industrialized construction and efficient building processes. However, the connections between modules remain a critical bottleneck constraining overall performance and widespread adoption. Current research and engineering practice reveal widespread issues such as low on-site module assembly efficiency, insufficient installation precision, poor recyclability of joints, and the absence of health monitoring systems. These problems not only impact construction progress and safety but also hinder the sustainable development of buildings throughout their entire lifecycle. To address these bottlenecks, a systematic analysis across multiple dimensions—including assembly techniques, node design, material recycling, and intelligent monitoring—is required. This approach will provide

the theoretical foundation and technical support for subsequent optimization of assembly devices and research into green connection technologies.

### **3.1 Challenges of on-site splicing**

Modular steel structures face numerous technical and management challenges during on-site assembly. Due to the large dimensions of individual modules, complex interfaces, and variable construction environments, on-site assembly efficiency is generally low. Ensuring high-precision alignment between modules is difficult, leading to installation deviations and cumulative errors that compromise overall structural stability and safety. Furthermore, on-site assembly is often constrained by limited lifting space, equipment precision, and the technical proficiency of construction personnel, resulting in complex operations and high coordination difficulties. Enhancing assembly speed while maintaining installation accuracy has become a critical bottleneck hindering the widespread adoption of modular steel structures. Consequently, research and practice are increasingly focused on developing efficient auxiliary assembly devices, intelligent positioning technologies, and standardized assembly systems to improve construction efficiency and quality.

#### **3.1.1 On-site splicing efficiency is low**

Modular steel structures commonly suffer from low construction efficiency during on-site assembly, primarily due to: 1) insufficient standardization of modules leading to poor interface compatibility; 2) complex coordination of on-site work sequences; 3) high dependence on specialized equipment; and 4) environmental constraints affecting lifting and positioning accuracy(Li et al., 2017). Furthermore, the absence of systematic assembly processes and collaborative management prolongs construction cycles and increases rework rates. Research indicates that establishing standardized and serialized modular systems, coupled with the integration of BIM technology and laser scanning positioning equipment during construction, can significantly enhance assembly precision and reduce construction cycles by approximately 20%–50%(McKinsey & Company, 2019). Simultaneously, adopting dry or bolted connection methods reduces welding operations, improving on-site operational safety and repeatability(Chong et al., 2016). Looking ahead, integrating digital twin technology with 5G communication to enable real-time monitoring and data feedback during modular construction will become a key direction for enhancing assembly efficiency and resource utilization(Lu et al., 2021). Therefore, improving the efficiency of steel structure modular assembly requires coordinated advancement across three dimensions: standardized design, intelligent equipment, and lean construction.

#### **3.1.2 Difficulties in performing combined operations across multiple modules**

During multi-module collaborative installation, cumulative errors are the primary factors affecting the overall stability and assembly accuracy of steel modular buildings. Such errors typically originate from dimensional deviations and orientation errors during module manufacturing, transportation, and on-site hoisting. Their cumulative effect is amplified during installation, leading to structural misalignment, uneven load distribution at connections, and localized stress concentration(Wang et al., 2019). To minimize error propagation, strict control of component machining accuracy is required during manufacturing. Anti-deformation reinforcement measures should be implemented during transportation, and high-precision measurement and positioning verification must be conducted prior to assembly(Zhao et al., 2019). On-site, integrating laser scanning, total stations, and BIM 3D models enables millimeter-level alignment control, significantly enhancing assembly precision and installation efficiency(Tang et al., 2010). Furthermore, real-time monitoring and data analysis systems enable automatic identification and correction of installation deviations, thereby enhancing safety and reliability during multi-module assembly. With the advancement of artificial intelligence, the Internet of Things, and digital twin technologies, future modular steel structure installations will achieve intelligent precision control and full-process traceability, providing technological support for the high-quality development of prefabricated buildings(Lu et al., 2021).

### **3.2 Recyclability issues**

Modular construction of steel structures should adhere to principles of environmental protection, health, safety, and energy conservation. Through systematic material selection and design, it aims to reduce environmental pollution, enhance structural safety, and lower energy consumption. However, research on the recyclability of connection nodes between modules currently lags significantly behind the overall development of structural systems. The constraints primarily manifest in three aspects:

irreversible material damage, structural failure thresholds, and technical-economic compatibility(Gorgolewski, 2008). Regarding materials, welded and bolted connections under cyclic loading are prone to crack propagation and localized tearing, leading to rapid degradation of node load-bearing capacity(Ding et al., 2016). Structurally, hole aspect ratio, plate thickness, and bolt arrangement significantly impact node durability; improper design reduces connection performance after repeated assembly and disassembly(Li et al., 2020). Technically and economically, the absence of performance evaluation systems and cost standards for reusable joints also hinders the development of sustainable construction systems. Future research should focus on the mechanical degradation patterns of demountable joints, strength recovery mechanisms of recycled materials, and life cycle assessment methodologies to provide theoretical support for greening and recycling in modular steel structures(Zhao et al., 2018).

### **3.3 Lack of health monitoring**

The health monitoring system for modular steel structure joints remains in its infancy, facing critical bottlenecks such as high data dependency, environmental noise interference, and difficulties in identifying multiple damage types(Farrar & Worden, 2013). While existing research has made some progress along different technical pathways, overall issues of fragmentation and insufficient applicability persist. Acoustic-based methods enable early damage detection by analyzing the time-frequency characteristics of bolt loosening sounds. However, they heavily rely on manual feature extraction and noise filtering, making stable quantitative diagnosis challenging(Zhang et al., 2021). Structure-response-based monitoring methods (e.g., acceleration and strain signals) can utilize convolutional neural networks to extract damage features, but their performance is constrained by sample size and operational diversity(Teng et al., 2022). Computer vision technologies like YOLO and Mask R-CNN demonstrate strong automatic recognition capabilities in bolt condition detection, yet false detection rates remain high due to factors like lighting, weather, and occlusion(Ref 54). In summary, there is currently a lack of integrated monitoring systems that fuse multi-source sensor data while offering environmental robustness and low data dependency. Future research should focus on developing multimodal sensor fusion and digital twin-driven self-learning diagnostic models to achieve adaptive structural health monitoring and full-lifecycle management under complex operating conditions.

## **4.Future research trends and outlook**

Driven by China's intensifying policies for prefabricated construction and the "dual carbon" goals, the market for modular steel structures will achieve steady growth with continuously optimized application frameworks. As a vital component of prefabricated building development, modular steel structures urgently require advanced technologies and methodologies to achieve modernization. Looking ahead to future green construction demands, and considering current research, application status, and development trends both domestically and internationally, the following areas warrant focused attention in research and application advancement. Progress in modular steel structure construction can be gradually enhanced through the development of auxiliary joining devices, upgraded connection techniques, and the creation of intelligent management systems.

### **4.1 Auxiliary joining devices for module connections**

Traditional steel structures face challenges during assembly, including difficulties in achieving stable fixation, susceptibility to assembly errors, and inherent risks in the joining process. Existing auxiliary devices are bulky, making them impractical for on-site operations with small steel structures. Their cumbersome mobility results in low splicing efficiency and operational difficulties for small steel components. Therefore, researching auxiliary splicing devices for small steel structures represents a crucial trend in modular steel construction, holding significant scientific and engineering value.

### **4.2 Green connection technologies between modules**

#### **4.2.1 Demountable design**

Current modular steel structure joints predominantly employ permanent connection methods. Disassembly often causes component damage, hindering material recycling. Traditional welded joints undergo high-temperature cutting during demolition, altering the metallographic structure of steel and degrading its mechanical properties. Some bolted joints experience low bolt reuse rates post-disassembly due to long-term corrosion or preload decay. Additionally, existing removable joint designs suffer from complex construction and high costs. Furthermore, the absence of unified evaluation

standards for demountable joints leads to significant variations in disassembly efficiency and material recovery rates across projects, hindering the green and circular development of modular steel structures. Therefore, developing low-cost, easily fabricatable, and reliable demountable joints while establishing a comprehensive recycling evaluation system represents a critical direction for advancing the sustainability of modular steel construction.

#### 4.2.2 Low-carbon materials

Currently, modular steel structural nodes primarily rely on virgin steel, whose production process generates high carbon emissions, whereas recycled steel exhibits significantly lower carbon emissions. However, recycled steel has notable shortcomings in node applications: bolted joints made from recycled steel exhibit low tensile strength and reduced fatigue life. Concurrently, existing low-carbon materials lack sufficient adaptability for node connections. Furthermore, design codes and construction techniques for low-carbon material joints remain underdeveloped, with a lack of joint calculation models tailored to the mechanical properties of these materials, creating safety hazards in engineering applications. Therefore, enhancing recycled steel performance, developing novel low-carbon composite materials, and establishing design standards for low-carbon material joints are core pathways to reduce the carbon footprint of modular steel structures. How to utilize recycled steel in modular steel structure joints and fully replace virgin steel represents an urgent and promising future research direction.

#### 4.3 Intelligent management of modular steel structures

Currently, intelligent management across the entire lifecycle—design, construction, and operation—remains inadequate, with significant data fragmentation. During design, BIM models struggle to integrate with other specialized software. In construction, while IoT monitoring devices are deployed, data processing efficiency is low, requiring manual analysis before adjustments can be made—failing to meet dynamic construction demands. During operation and maintenance, there is a lack of full lifecycle data tracking, resulting in low accuracy for node damage early warning. Furthermore, intelligent management systems exhibit poor compatibility, making it difficult for equipment and software from different manufacturers to work together. Therefore, establishing a cross-platform data interaction system, developing real-time intelligent analysis algorithms, and creating a full lifecycle data traceability system are key breakthrough points for achieving intelligent management of modular steel structures.

### 5. Conclusion

This study examines the current status, challenges, and development trends in the research and application of modular steel structures. Modular steel construction is emerging as a future trend in the building industry, yet the field still faces significant challenges. These include designing auxiliary joining devices suitable for small-scale steel modularization, developing new environmentally friendly and low-carbon steel materials, and advancing intelligent modular steel construction technologies. Therefore, we must advance steel modular construction into a new phase characterized by “future-oriented, fully renewable, and achieving negative carbon emissions” through the integration of material innovation and intelligent construction technologies.

### Funding

No

### Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

### Reference

- [1] Zhu, X., Yan, J., & Wu, J. (2022). Latest progress in connection nodes of modular steel structure buildings [in Chinese]. *Building Structure*, 52(S2), 1532–1538.
- [2] Wang, W., Li, Z., Yu, C., Liu, Y., Ye, H., & Li, Z. (2020). Research progress on structural systems of modular steel structure buildings [in Chinese]. *Construction Technology*, 49(11), 24–30, 36.
- [3] Hou, Z., & Liu, X. (2021). Key technologies and demonstration of industrialization of steel structure buildings [in

- Chinese]. *Progress in Steel Building Structures*, 23(10), 1–11.
- [4] Peng, L., Xie, H., Yuan, L., Hong, Y., Liu, H., Feng, X., Deng, B., & Xu, H. (2020). Application of prefabricated buildings in Wuhan Leishenshan Hospital [in Chinese]. *Huazhong Architecture*, 38(04), 71–77.
- [5] Yang, X., Lu, Y., Gu, C., & Liu, Q. (2016). Design of new multi-storey and high-rise steel structure box modular buildings [in Chinese]. *Progress in Steel Building Structures*, 18(05), 41–47, 72.
- [6] Ding, Y., Deng, E., Zong, L., Dai, X., Li, Y., Wang, H., & Bi, J. (2019). Research progress on connection nodes of modular steel structure buildings [in Chinese]. *Journal of Building Structures*, 40(03), 33–40.
- [7] Huang, X., Shu, Z., & Li, Z. (2022). Research progress on prefabricated modular buildings and modular nodes [in Chinese]. *Progress in Steel Building Structures*, 24(02), 40–49.
- [8] Hou, Z., Gong, C., Chen, J., et al. (2023). Industrialization technology and demonstration of fully prefabricated modular steel structure building system [in Chinese]. *Progress in Steel Building Structures*, 25(1), 1–10.
- [9] Liu, X., Ren, X., Zhan, X., & Zhang, Y. (2018). Mechanical performance analysis of beam–column joints of a box-type modular prefabricated steel structure house [in Chinese]. *Industrial Construction*, 48(05), 62–69.
- [10] Wang, Y. (2020a). Study on mechanical properties of new bolt-in-column joints in modular buildings (Master's thesis, China University of Mining and Technology) [in Chinese].
- [11] Wang, Q., Su, M., & Qi, H. (2020b). Research on bending performance of cross plate–end plate modular steel structure connection nodes [in Chinese]. *Construction Technology*, 49(05), 44–48.
- [12] Pan, C., Wu, X., Zhou, L., & Wei, H. (2023). Application of BIM technology in design and construction of prefabricated steel structure buildings [in Chinese]. *Information of China Construction*, (05), 62–65.
- [13] Li, W. (2019). Research on development and construction technology of new composite wall panels for prefabricated steel structure buildings (Master's thesis, Shandong Jianzhu University) [in Chinese].
- [14] Wen, X., Li, C., Chen, Y., Zhou, Y., Liao, X., & Zhang, X. (2021). Key technologies of integrated construction for modular steel structure buildings [in Chinese]. *Construction Technology (Chinese & English)*, 50(18), 65–70.
- [15] Zhai, S., Cao, K., Li, G., Wang, W., Chen, C., & Liu, Q. (2022). Research progress on inter-module connection nodes of column-supported modular steel structure buildings [in Chinese]. *Progress in Steel Building Structures*, 24(11), 1–11.
- [16] Ye, J., & Yu, H. (2019). Research progress on connection nodes of steel structure modular buildings [in Chinese]. *Guangdong Architecture Civil Engineering*, 26(03), 9–12, 36.
- [17] Zhang, H., He, Y., Lu, H., et al. (2025). An evolutionary optimization method for modular structures made of predefined building blocks. *Engineering Structures*, 343, 121147. <https://doi.org/10.1016/j.engstruct.2025.121147>
- [18] Sharafi, P., Atashfaraz, B., & Kildashti, K. (2025). Novel rocking interconnection for superior prefabricated prefinished modular building system. *Engineering Structures*, 343, 120934. <https://doi.org/10.1016/j.engstruct.2025.120934>
- [19] Jin, J. Y., Qian, K., & Dong, Y. G. (2013). The application of energy efficiency system with EPS module steel structure of industrial building in construction engineering. *Advanced Materials Research*, 671–674, 2145–2149. <https://doi.org/10.4028/www.scientific.net/AMR.671-674.2145>
- [20] Xiong, F., Zhang, D., Ran, M., et al. (2025). Seismic performance of bolt-connected fully assembled concrete modular construction by full-scale shaking table test. *Engineering Structures*, 343, 121026. <https://doi.org/10.1016/j.engstruct.2025.121026>
- [21] Ly, H. D., Nguyen, N. H. D. T., Jang, H., et al. (2025). Developing a life cycle assessment-based framework for module-based impact distribution in adaptive reuse of modular buildings. *Journal of Building Engineering*, 112, 113703. <https://doi.org/10.1016/j.jobbe.2025.113703>
- [22] Liu, Z. Y. (2025). Analysis on application effect of high-strength bolted connections in complex steel structure joints [in Chinese]. *China Construction Metal Structure*, 24(16), 88–90. <https://doi.org/10.20080/j.cnki.ISSN1671-3362.2025.16.030>
- [23] Chen, Z. X., Li, X. L., & Chen, W. (2024). Analysis of friction surface treatment methods for high-strength bolted connections based on tribological principles [in Chinese]. *Total Corrosion Control*, 38(12), 25–27. <https://doi.org/10.1016/j.tcc.2024.12.005>

org/10.13726/j.cnki.11-2706/tq.2024.12.025.03

- [24] Chen, Y. Y., Shen, Z. Y., Han, L., et al. (2004). Determination of slip resistance coefficient of joint surfaces coated with alkyd iron oxide red or polyurethane zinc-rich paint [in Chinese]. *Building Structure*, (05), 3–6. <https://doi.org/10.19701/j.jzjg.2004.05.001>
- [25] Liu, G., Li, Q. C., & Hu, A. J. (2009). Experimental study on the performance of high-strength bolted connections with sandblasted contact surfaces [in Chinese]. *Journal of Suzhou University of Science and Technology (Engineering and Technology Edition)*, 22(03), 38–41.
- [27] Wang, Q. M., Hou, Z. X., Dong, X. H., et al. (2005). Experimental study on the performance of high-strength bolted connections with coated friction surfaces [Conference paper, in Chinese]. In *Proceedings of the 5th National Symposium on Modern Structural Engineering* (pp. 1181–1187). Tianjin University.
- [28] Chen, X. S., Shi, G., & Chen, Y. F. (2021). Experimental study on slip resistance coefficient of arc thermal sprayed aluminum contact surfaces for high-strength steel bolted connections [in Chinese]. *Building Structure*, 51(13), 87–92. <https://doi.org/10.19701/j.jzjg.2021.13.013>
- [29] Liu, X. (2020). Study on the ecology of new temporary buildings in the whole life cycle (Master's thesis, Hunan University) [in Chinese]. <https://doi.org/10.27135/d.cnki.ghudu.2020.001507>
- [30] Song, L. L., Yan, W. M., Zhang, X. M., et al. (2020). Applicability analysis of self-piercing riveted connections in prefabricated light steel structures [in Chinese]. *Journal of Beijing University of Technology*, 46(01), 33–41.
- [31] Xu, L. H., Yang, X. L., & Zhang, G. (2024). Study on mechanical performance of new insert-type connection joints for modular steel structures [in Chinese]. *Engineering Mechanics*, 41(01), 30–38.
- [32] Dai, X. M. (2021). Study on seismic performance and design method of insert-type self-locking joints for modular steel structures (Doctoral dissertation, Tianjin University) [in Chinese].
- [33] Wang, X. M. (2023). Study on mechanical performance of new inner insert self-locking modular steel structure joints (Master's thesis, China University of Mining and Technology) [in Chinese].
- [34] Ministry of Housing and Urban-Rural Development of the People's Republic of China. (2018). Standard for design of steel structures: GB 50017—2017 [in Chinese]. Beijing: China Architecture & Building Press.
- [35] Wang, Y. Q., Wu, M. Z., Wang, Z. X., et al. (2025). Research on application technology of stainless steel grooved rivets and their connection joints [Conference paper, in Chinese]. In *Proceedings of the 34th National Conference on Structural Engineering* (Vol. I) (pp. 13–28). <https://doi.org/10.26914/c.cnkihy.2025.040908>
- [36] Wang, Q. H., Wang, X. J., Fang, J. J., et al. (2022). Preparation and performance study of water-based intumescent fire-retardant coatings for steel structures [in Chinese]. *Paint & Coatings Industry*, 52(02), 42–48.
- [37] Kong, S. (2013). Study on thermal insulation ceramic fibers based on animal fiber (silk) templates (Doctoral dissertation, Nanjing University of Science and Technology) [in Chinese].
- [38] Liu, D. C. (2012). Exploration on the frontier of ceramic fiber technology [in Chinese]. *Advanced Ceramics*, 33(01), 25–31. <https://doi.org/10.16253/j.cnki.37-1226/tq.2012.01.004>
- [39] Yao, S. Y. (2006). Study on the preparation process of mullite continuous fibers (Doctoral dissertation, China Academy of Machinery Science and Technology) [in Chinese].
- [40] Li, C. Z., Zhong, R. Y., Xue, F., Xu, G., Chen, K., & Huang, G. Q. (2017). Integrating RFID and BIM technologies for mitigating risks and improving schedule performance of prefabricated house construction. *Journal of Cleaner Production*, 165, 1048–1062.
- [41] McKinsey & Company. (2019). *Modular construction: From projects to products*. McKinsey Global Institute.
- [42] Chong, H. Y., Lopez, R., Wang, J., & Kim, C. (2016). Improving prefabricated construction productivity through BIM-based automated design validation. *Automation in Construction*, 72, 15–27.
- [43] Lu, W., Xue, F., & Zhao, R. (2021a). Digital twin-enabled smart construction: Conceptual framework, applications and future development. *Automation in Construction*, 123, 103558.
- [44] Wang, J., Li, H., & Tam, V. W. (2019). Critical factors in effective modular construction project management. *Automa-*

tion in Construction, 100, 1–11.

- [45] Zhao, X., Hwang, B. G., & Gao, Y. (2019). A critical review of building information modeling (BIM) for modular construction. *Automation in Construction*, 102, 120–131.
- [46] Tang, P., Huber, D., Akinci, B., Lipman, R., & Lytle, A. (2010). Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques. *Automation in Construction*, 19(7), 829–843.
- [47] Lu, W., Xue, F., & Zhao, R. (2021b). Digital twin-enabled smart construction: Conceptual framework, applications and future development. *Automation in Construction*, 123, 103558.
- [48] Gorgolewski, M. (2008). Design for deconstruction and materials reuse. *Proceedings of the ICE – Waste and Resource Management*, 161(4), 163–167.
- [49] Ding, T., Xiao, J., & Tam, V. W. (2016). A closed-loop life cycle assessment of recycled aggregate concrete utilization in China. *Waste Management*, 56, 367–375.
- [50] Li, X., Chi, H. L., & Wu, P. (2020). Sustainable design and construction of prefabricated buildings: A review of BIM and modular construction research. *Journal of Cleaner Production*, 273, 123013.
- [51] Zhao, X., Hwang, B. G., & Low, S. P. (2018). Critical success factors for enterprise risk management in modular construction projects: A sustainability perspective. *Automation in Construction*, 94, 146–158.
- [52] Farrar, C. R., & Worden, K. (2013). *Structural health monitoring: A machine learning perspective*. Wiley.
- [53] Zhang, Y., Zhu, L., & Wu, D. (2021). Acoustic emission-based bolt looseness detection using deep learning and time–frequency features. *Mechanical Systems and Signal Processing*, 158, 107756.
- [54] Teng, S., Li, J., & Xu, B. (2022). CNN-based vibration signal analysis for damage identification in steel structures. *Engineering Structures*, 252, 113615.
- [55] Sun, Z., Wang, W., & Chen, G. (2022). Vision-based bolt looseness detection using deep convolutional networks under varying illumination conditions. *Automation in Construction*, 140, 104378.

# Public Data Access and AI Adoption for Sustainable Digital Transformation: Evidence from China

Wei Zhao<sup>1,2\*</sup>

1.IEBIS, Department of High-Tech Business and Entrepreneurship, Faculty of Behavioral, Management and Social Sciences, University of Twente, 7522 NB Enschede, The Netherlands

2.Faculty of Business Administration, Turiba University, LV-1058 Riga, Latvia

\*Corresponding author: Wei Zhao, [zhaowei11208@gmail.com](mailto:zhaowei11208@gmail.com)

**Copyright:** 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

**Abstract:** The correlation between public data accessibility and the adoption intensity of artificial intelligence (AI) in Chinese enterprises remains systematically understudied. By leveraging the rollout of municipal public data platforms as a quasi-natural experiment, this study demonstrates that enterprises in cities with such platforms exhibit significantly stronger AI adoption than those in non-platform regions. Mechanistically, this effect operates through dual pathways: significant reductions in operational expenditures and structural upgrades in specialized AI workforce allocation. This study elucidates the action pathway of China's data platform opening up in facilitating the application of artificial intelligence within enterprises. Furthermore, it offers a universal analytical framework for examining the coupling mechanism between data elements and industrial digital transformation during the technological transition in developing countries. These findings also suggest that improving public data accessibility contributes to sustainable digital transformation by aligning technological diffusion with efficient and inclusive resource utilization.

**Keywords:** Public Data Access; Artificial Intelligence; Information-Acquisition Costs; Talent Structure; Sustainability

**Published:** Nov 13, 2025

**DOI:** <https://doi.org/10.62177/jaet.v2i4.848>

## 1.Introduction

As a disruptive technology at the forefront of a new wave of technological revolution, artificial intelligence is a crucial driver of economic growth. According to the Chinese State Council's "Notice on the Issuance of the New Generation Artificial Intelligence Development Plan," it is projected that by 2030, the scale of China's core industries related to AI will exceed 10 trillion yuan, with related industries surpassing 100 trillion yuan. Existing literature has also shown that AI technologies contribute to enhancing enterprise innovation capabilities and product quality (Xu and Tian, 2023; Babina et al., 2024). However, Accenture's "2024 Digital Transformation Index Report for Chinese Enterprises" indicates that most companies in China are still taking a wait-and-see approach regarding AI technologies. Therefore, how to enhance the level of AI technology adoption among Chinese enterprises is a crucial issue.

In addition to driving economic growth, AI adoption also plays a pivotal role in achieving sustainable digital transformation, where technological innovation serves environmental, social, and governance objectives. Recent research has shown that data accessibility and digital intelligence can strengthen the sustainability of industrial systems by improving energy efficiency

and promoting green innovation (Zhao, Chen, & Bulis, 2025). Likewise, the integration of AI and Industry 4.0 technologies supports a sustainable society across sectors by optimizing resource use and reducing operational waste (Zhao, Chen, Yazan, et al., 2025). Hence, understanding the mechanisms through which public data access fosters AI adoption provides not only an economic but also a sustainability-oriented perspective on how data infrastructure enables responsible innovation in developing economies.

The technology adoption theory posits that the adoption of new technologies is influenced by various factors, with technology adoption costs and human resources being particularly significant (Straub, 2009; Rojas-Berrio et al., 2022). The application of AI technology highly relies on the collection and utilization of information; however, the high transformation costs associated with gathering quality data and the demand for skilled personnel significantly dampen enterprises' enthusiasm for adopting artificial intelligence technologies. Fortunately, the public data access offers a potential solution to this challenge.

From the standpoint of sustainable digital governance, the establishment of public data platforms can be regarded as a long-term institutional innovation that enhances both technological inclusiveness and social welfare. By lowering information-acquisition costs, these platforms contribute to equitable technology diffusion and inclusive growth—two cornerstones of the United Nations' Sustainable Development Goals (SDGs 9 and 16).

Public data platform accessibility facilitates enterprise adoption of artificial intelligence (AI) technologies via two primary pathways. First, public platforms reduce operational costs by overcoming institutional barriers to data circulation through their authoritative governance and cross-domain interoperability. Compared with the monopolistic nature of proprietary data and privacy limitations of individual data, standardized public datasets with high credibility lower enterprises' comprehensive costs—including data procurement, preprocessing, and regulatory compliance. These efficiency gains in technology assimilation strengthen enterprises' economic incentives for AI adoption (Farboodi & Veldkamp, 2021; Du et al., 2024). Second, platform-driven technological spillovers promote human capital restructuring by attracting specialized talent clusters—a phenomenon particularly evident in AI-intensive industries (Zhao & Fan, 2021). To maximize public data utility, enterprises optimize talent deployment by constructing skill hierarchies that integrate AI research capabilities with domain-specific implementation expertise. This restructuring elevates the proportion of technical positions and fosters cross-disciplinary competency networks, creating a self-reinforcing cycle for AI diffusion.

As previously noted, while data accessibility theoretically enhances enterprise AI adoption, most developing economies struggle to translate public data into productive resources due to inadequate data governance capabilities and institutional constraints. Weak institutional willingness to share data and limited technical capacity exacerbate cost barriers for technology adoption (Du et al., 2024). In this context, following the U.S. launch of the world's first national public data platform in 2009, countries like the UK, Singapore, and Russia have implemented similar government data portals. Among developing nations, China pioneered systematic efforts in public data accessibility, initiating its Shanghai Municipal Data Service Platform in 2012 and subsequently establishing localized public data infrastructures. By 2021, China had operationalized 173 municipal-level platforms, creating a robust quasi-experimental setting for empirically identifying causal relationships between data accessibility and enterprise AI adoption.

Although data accessibility can enhance enterprise AI adoption, developing economies face challenges in operationalizing public data due to insufficient governance frameworks and structural barriers. Institutional reluctance toward data sharing and technical capacity gaps intensifies preprocessing and compliance costs (Du et al., 2024). Since the United States inaugurated the first national public data platform in 2009, industrialized nations including the UK and Singapore have established comparable national platforms. As a developing economy pioneer, China launched the Shanghai Municipal Data Service Platform in 2012, followed by 173 municipal-level public data infrastructures nationwide. This multi-tiered infrastructure creates natural experiment conditions through staggered implementation timelines, enabling rigorous analysis of data accessibility's causal impact on AI adoption via difference-in-differences methodology.

Using the establishment of public data platforms as a quasi-natural experiment, our research findings indicate that companies located in cities with public data platforms exhibit a higher degree of AI technology application compared to those in cities without public data platforms. This effect operates through dual pathways: significant reductions in operational expenditures

and structural upgrades in specialized AI workforce allocation.

The contributions of our research are as follows: First, we enhance the understanding of the economic consequences of public data access. The academic community generally agrees that public data access can create significant socio-economic value (Zhao and Fank, 2021); It also can help overcome managerial limitations and optimize the external information landscape (Farboodi and Veldkamp, 2021; Du et al., 2024). Our research further supplements the literature on the economic implications of public data access from the perspective of AI technology application. Second, our study has practical significance. Historically, governments have often relied on high-tech subsidies to encourage enterprises to adopt new technologies. However, such measures can lead to issues like the emergence of “zombie” companies (Chang et al., 2021). Comparing with the subsidy strategies, the establishment of public data platforms reduces intermediate costs for enterprises, promotes technological advancement while mitigating resource misallocation issues. This practical value is also beneficial for other developing countries.

## 2. Research design

### 2.1 Sample selection

The sample period of our study spans from 2010 to 2022. Our study manually collected information on cities with public data access. Annual report data was sourced from the Giant Tide Information Network, while other financial indicators were obtained from the CSMAR databases. Furthermore, we excluded ST companies, those in the financial sector, and firms with missing or abnormal data, applying a 1% Winsorization to the dataset.

### 2.2 Variable definition

#### 2.2.1 Dependent variable

Artificial Intelligence (AI): Drawing on the AI dimension dictionary provided by Chen and Srinivasan (2024), our study further integrates insights from multiple industry research reports and the AI vocabulary supplied by the World Intellectual Property Organization (WIPO) to identify 52 seed terms, including “artificial intelligence,” “machine learning,” and “internet of things.” Utilizing Word2Vec technology, we employed the Skip-gram model to train on the vocabulary extracted from annual reports and patent texts. After manually removing duplicate keywords generated by the machine learning process, we finalized the AI dictionary. We then calculated the natural logarithm of the number of AI-related keywords plus one from the MD&A sections to serve as our indicator of corporate AI adoption.

#### 2.2.2 Independent variable

Government Data Openness (Open): Following Du et al. (2024), we assign a value of 1 to Open if the location of the company has a public data platform, 0 otherwise.

#### 2.2.3 Control variables

Our study employs common firm-level characteristics as control variables; Table 1 provides detailed definitions.

*Table 1 Definitions of the main variables*

<i>Variables</i>	<b>Definition</b>
<i>Open</i>	An indicator variable equals one if the public data platform opens a year before in the city where the firm is headquartered and zero otherwise.
<i>AI</i>	The natural logarithm of the sum of artificial intelligence terms mentions in the MD&A plus one.
<i>Size</i>	The natural logarithm of the total assets.
<i>Lev</i>	Total liabilities/Total assets.
<i>Roa</i>	Net profit/Total assets.
<i>ListAge</i>	The natural logarithm of one plus listing age.
<i>Cashflow</i>	Net cash flow from operating activities/Ending current liabilities
<i>Growth</i>	Increase in current year revenue/Previous year total revenue.

<i>Variables</i>	<b>Definition</b>
<i>Top1</i>	The share of top 1 shareholders.
<i>Board</i>	The natural logarithm of the board count.
<i>Dual</i>	Whether the chairman and general manager concurrently serve.

### 2.3 Model construction

To examine the impact of public data access on the application degree of AI technology, we employ a difference-in-differences (DID) regression model:

$$AI_{i,t} = \alpha_1 + \alpha_2 Open_{i,t} + \alpha_3 Controls_{i,t} + \delta_i + \eta_t + \varepsilon_{i,t} \quad (1)$$

Where *i* represents a firm and *t* indicates the year. The dependent variable, *AI*, measures the application degree of artificial intelligence technology. The independent variable, *Open*, is an indicator that equals 1 if the the public data platform has been implemented in the city where firm *i* is in year *t*, and 0 otherwise. Additionally, our study controlled for firm fixed effects ( $\delta_i$ ) and year fixed effects ( $\eta_t$ ). The sample is clustered at the firm level.

## 3. Empirical methodology

From a policy perspective, public data infrastructures can also be interpreted as catalysts of sustainable digital governance. By reducing informational asymmetries and fostering transparent, data-driven decision-making, these infrastructures contribute to responsible innovation and long-term socio-economic resilience.

### 3.1 Summary statistics

Table 2 presents the descriptive statistics of the main variables. The mean of *AI* is 0.062 with a standard deviation of 0.185. This indicates that the level of application of *AI* among Chinese listed companies is relatively low, which indirectly highlights the necessity of our research.

Table 2 Summary statistics

<i>Variable</i>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>Min</b>	<b>p50</b>	<b>Max</b>
<i>AI</i>	24,365	0.062	0.185	0.000	0.000	1.000
<i>Open</i>	24,365	0.364	0.481	0.000	0.000	1.000
<i>Size</i>	24,365	22.350	1.318	19.960	22.180	26.110
<i>Lev</i>	24,365	0.437	0.209	0.053	0.435	0.889
<i>Roa</i>	24,365	0.036	0.063	-0.219	0.035	0.207
<i>ListAge</i>	24,365	2.393	0.761	0.000	2.565	3.367
<i>Cashflow</i>	24,365	0.044	0.068	-0.149	0.042	0.230
<i>Growth</i>	24,365	0.163	0.367	-0.511	0.106	1.916
<i>Top1</i>	24,365	0.333	0.147	0.088	0.308	0.725
<i>Board</i>	24,365	2.135	0.197	1.609	2.197	2.639
<i>Dual</i>	24,365	0.242	0.429	0.000	0.000	1.000

### 3.2 Baseline regression

Table 3 presents the baseline regression results. As shown in the table, the coefficient of *Open* are all significantly positive at the 1% level, regardless of whether control variables and fixed effects are included. This indicates that the establishment of public data platforms significantly enhances the application of *AI* technologies by enterprises.

Table 3. Baseline results.

	(1)	(2)	(3)
	AI		
<i>Open</i>	0.019*** (3.86)	0.078*** (15.92)	0.018*** (3.61)
<i>Size</i>		0.016*** (7.81)	0.034*** (7.90)
<i>Lev</i>		-0.097*** (-7.43)	-0.015 (-0.90)
<i>ROA</i>		-0.178*** (-4.90)	-0.124*** (-3.81)
<i>ListAge</i>		0.008*** (3.34)	0.035*** (5.29)
<i>Cashflow</i>		-0.065*** (-3.07)	-0.044** (-2.43)
<i>Growth</i>		-0.001 (-0.44)	-0.012*** (-3.76)
<i>Top1</i>		-0.136*** (-8.80)	-0.091*** (-3.46)
<i>Board</i>		-0.034*** (-3.17)	0.019 (1.34)
<i>Dual</i>		0.019*** (3.83)	0.004 (0.88)
<i>Constant</i>	0.055*** (30.39)	-0.188*** (-4.51)	-0.788*** (-8.21)

### 3.3 Mechanism test

#### 3.3.1 Reduce costs

Cost is a significant factor that influences whether firms adopt AI technologies (Straub, 2009). To assess whether the establishment of public data platforms contributes to reducing corporate costs and thereby encourages the adoption of AI technologies, we selected two closely related indicators: (1) Management Expense Ratio (Mfee), defined as the ratio of a company's management expenses to its revenue from core operations; (2) Operating Cost Ratio (OpCostR), defined as the ratio of a company's operating costs to its revenue from core operations. The relevant regression results are presented in Table 4 columns (1)-(2), where the coefficient of Open are all significantly negative. This suggests that the public data access contributes to lowering corporate costs, thereby facilitating the adoption of AI technologies by enterprises. Beyond cost efficiency, these reductions may indirectly promote corporate sustainability by minimizing redundant resource use and energy consumption. As firms leverage open data for process optimization, their digital transformation also aligns more closely with sustainable production and environmental responsibility.

#### 3.3.2 Optimize the talent structure

Talent is also a crucial factor that limits the application of new technologies by enterprises (Rojas-Berrio et al., 2022). To

evaluate whether the establishment of public data platforms aids in optimizing the talent structure of companies, we used the proportion of employees with postgraduate degrees as the measuring indicator (Gra). The relevant regression results are presented in Table 4 column (3), where the coefficient of Open is significantly positive. This indicates that the public data access facilitates the aggregation effect of talent, thereby promoting the application of AI technologies through the optimization of the company's talent structure.

Table 4. Mechanism test.

	(1)	(2)	(3)
	Mfee	OpcostR	Gra
<i>Open</i>	-0.003** (-2.12)	-0.005* (-1.78)	0.336*** (4.17)
<i>Size</i>	-0.013*** (-8.91)	-0.020*** (-7.14)	0.733*** (7.19)
<i>Lev</i>	-0.033*** (-5.70)	0.107*** (8.93)	-1.095*** (-2.77)
<i>ROA</i>	-0.189*** (-17.16)	-1.478*** (-45.21)	-0.844 (-1.48)
<i>ListAge</i>	-0.006*** (-3.22)	0.015*** (4.22)	-0.369*** (-2.79)
<i>Cashflow</i>	-0.024*** (-3.47)	-0.102*** (-6.01)	-0.505 (-1.30)
<i>Growth</i>	-0.019*** (-16.29)	-0.032*** (-11.19)	0.054 (0.74)
<i>Top1</i>	-0.005 (-0.67)	-0.029* (-1.82)	0.406 (0.60)
<i>Board</i>	0.005 (1.31)	0.001 (0.13)	0.220 (0.84)
<i>Dual</i>	-0.001 (-0.35)	0.001 (0.37)	-0.055 (-0.64)
<i>Constant</i>	0.405*** (12.35)	1.380*** (21.88)	-12.594*** (-5.68)
<i>Firm</i>	Yes	Yes	Yes
<i>Year</i>	Yes	Yes	Yes
<i>N</i>	24,365	24,365	24,365
<i>adj. R2</i>	0.716	0.677	0.750

Notes: (1) Standard errors in parentheses are clustered at the firm level; (2) \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , and \*  $p < 0.1$ .

### 3.4 Robustness test

To address potential endogeneity issues, our study includes the following tests. First, a key assumption of the difference-in-differences (DID) approach is the parallel trends assumption. Following the methodology of Beck et al. (2010), we conducted a parallel trends test, with the relevant results presented in Figure 1. The estimated coefficients for each period prior to the establishment of the public data platform are not statistically significant, indicating that the research sample satisfies the parallel trends assumption. Second, to mitigate endogeneity issues in the construction of indicators, we constructed a dummy variable (Iiai) as a substitute indicator based on whether the level of AI technology adoption in the firm exceeds the median. Additionally, we broadened the scope of term frequency collection by expanding the dictionary from the MD&A section to encompass the entire annual report (AI\_ar). The relevant results are presented in Table 5 columns (1)-(2), and our conclusions remain robust. Third, macroeconomic uncertainty may influence corporate strategy. To avoid any interference from the COVID-19 pandemic on our findings, we limited the sample period to years prior to 2020, with the relevant results shown in Table 5 column (3), again confirming the robustness of our conclusions.

Figure 1-parallel trends test

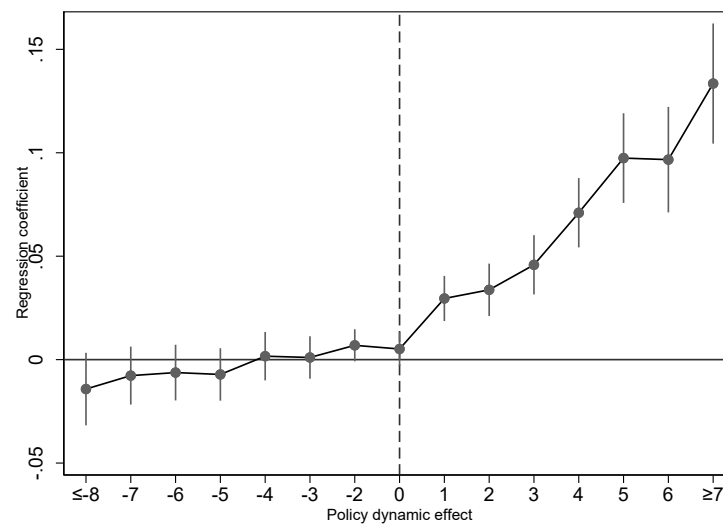


Table 5. Robustness test.

	(1)	(2)	(3)
	liai	AI_ar	AI
<i>Open</i>	0.005** (2.15)	0.029*** (4.13)	0.019*** (3.74)
<i>Size</i>	-0.002 (-0.83)	0.057*** (8.31)	0.030*** (6.40)
<i>Lev</i>	-0.005 (-0.39)	-0.012 (-0.48)	-0.016 (-0.96)
<i>ROA</i>	0.004 (0.17)	-0.205*** (-4.51)	-0.080** (-2.24)
<i>ListAge</i>	0.006 (1.36)	0.053*** (5.42)	0.020*** (3.51)
<i>Cashflow</i>	-0.027** (-2.22)	-0.063** (-2.45)	-0.027 (-1.55)
<i>Growth</i>	-0.001 (-0.65)	-0.014*** (-3.32)	-0.007** (-2.50)
<i>Top1</i>	0.017 (0.98)	-0.137*** (-3.64)	-0.058** (-2.35)
<i>Board</i>	-0.007 (-0.85)	0.038* (1.94)	0.028** (2.20)
<i>Dual</i>	0.003 (0.98)	0.008 (1.07)	0.004 (0.94)
<i>Constant</i>	1.035*** (16.47)	-1.336*** (-8.64)	-0.708*** (-6.80)
<i>Firm</i>	Yes	Yes	Yes
<i>Year</i>	Yes	Yes	Yes
<i>N</i>	24,365	24,365	18,185
<i>adj. R2</i>	0.554	0.490	0.334

Notes: (1) Standard errors in parentheses are clustered at the firm level; (2) \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , and \*  $p < 0.1$ .

## 4. Conclusion

The public data access is a crucial method for realizing the value creation of data as a factor. Our research confirms that access to public data significantly enhances the level of AI technology application, with cost reduction and talent optimization identified as potential mechanisms. This study elucidates the “data-to-technology” transmission mechanism, yielding two critical policy implications for AI diffusion in developing economies: First, institutional innovation should prioritize establishing data ownership frameworks and cross-jurisdictional sharing protocols, positioning public data platforms as core digital infrastructure components—exemplified by phased deployment models. Second, implementation strategies must adapt to regional industrial ecosystems through differentiated policy portfolios. Integrating such open-data frameworks into sustainable industrial policy can therefore serve as a dual lever for digital innovation and environmental responsibility. In this way, data accessibility supports not only economic modernization but also the broader transition toward a more sustainable and inclusive digital economy. For instance, manufacturing hubs should prioritize industrial data interface standardization, while talent-concentrated regions ought to reform skill certification systems to align with AI workforce demands.

Integrating such open-data frameworks into sustainable industrial policy can therefore serve as a dual lever for digital innovation and environmental responsibility. In this way, data accessibility supports not only economic modernization but also the broader transition toward a more sustainable and inclusive digital economy.

Beyond the economic dimension, the findings highlight that expanding open-data ecosystems can enhance long-term sustainability by fostering responsible digital transformation and inclusive growth. Public data infrastructures serve not only as economic catalysts but also as enablers of greener, more resilient innovation systems. By aligning data accessibility with sustainable development objectives, policymakers can ensure that technological progress advances environmental responsibility and social well-being in parallel with industrial upgrading.

## Funding

No

## Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Reference

- [1] Babina, T., Fedyk, A., He, A., & Hodson, J. (2024). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics*, 151, 103745.
- [2] Beck, T., Levine, R., & Levkov, A. (2010). Big bad banks? The winners and losers from bank deregulation in the United States. *The Journal of Finance*, 65(5), 1637-1667.
- Xu, X., & Tian, C. (2023). Does artificial intelligence improve the quality of export products? Evidence from China. *Applied Economics Letters*, 1-5.
- [3] Chang, Q., Zhou, Y., Liu, G., Wang, D., & Zhang, X. (2021). How does government intervention affect the formation of zombie firms? *Economic Modelling*, 94, 768-779.
- [4] Chen, W., & Srinivasan, S. (2024). Going digital: Implications for firm value and performance. *Review of Accounting Studies*, 29(2), 1619-1665.
- [5] Du, J., Gao, H., Wen, H., & Ye, Y. (2024). Public data access and stock price synchronicity: Evidence from China. *Economic Modelling*, 130, 106591.
- [6] Farboodi, M., & Veldkamp, L. (2021). A model of the data economy (No. w28427). Cambridge, MA, USA: National Bureau of Economic Research.
- [7] Magalhaes, G., & Roseira, C. (2020). Open government data and the private sector: An empirical view on business models and value creation. *Government Information Quarterly*, 37(3), 101248.
- [8] Rojas-Berrio, S., Rincon-Novoa, J., Sánchez-Monrroy, M., Ascúa, R., & Montoya-Restrepo, L. A. (2022). Factors influencing 4.0 technology adoption in manufacturing SMEs in an emerging country. *Journal of Small Business*

Strategy, 32(3), 67-83.

- [9] Straub, E. T. (2009). Understanding technology adoption: Theory and future directions for informal learning. *Review of Educational Research*, 79(2), 625-649.
- [10] Zhao, W., Chen, H., & Bulis, A. (2025). How are Industry 4.0 technologies transforming a sustainable society across industries? *Digital Transformation and Society*, 4(3), 363–380. <https://doi.org/10.1108/DTS-11-2024-0225>
- [11] Zhao, W., Chen, H., Yazan, D. M., Taghavifar, H., Lyu, Y., & Bulis, A. (2025). Few-shot learning and deep predictive models for cost optimization and carbon emission reduction in energy-water management. *Journal of Environmental Management*, 389, 126077. <https://doi.org/10.1016/j.jenvman.2025.126077>

# Indoor Pathfinding with the A\* Algorithm: A Cross-Platform Mobile Implementation Case

Kongwen Zhang<sup>1,\*</sup>, Boris Massesa<sup>1</sup>, Jingwei Gao<sup>2</sup>

1.School of Computing, University of the Fraser Valley, Abbotsford, British Columbia, Canada

2.College of Computer and Cyber Security, Hebei Normal University, Shijiazhuang, Hebei, 050000, China

\*Corresponding author: Kongwen Zhang, Frank.Zhang@ufv.ca

**Copyright:** 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

**Abstract:** This study presents the development of a fully integrated mobile module that enables indoor pathfinding functionality from the front end to the back end. The module is implemented using the A\* algorithm for route optimization and a NestJS framework with PostgreSQL and PostGIS for spatial data management. Designed as part of the University of the Fraser Valley (UFV) Campus App project, this cross-platform mobile application is built with React Native to ensure seamless usability across devices. Core functionalities include intelligent room search, interactive floor plan visualization, and real-time, turn-by-turn navigation powered by WebSocket communication. The system demonstrates the feasibility of combining spatial databases, efficient routing algorithms, and real-time communication technologies to enhance campus navigation and user experience.

**Keywords:** Indoor Pathfinding; A\* Algorithm; NestJS, Mobile Application

**Published:** Dec 9, 2025

**DOI:** <https://doi.org/10.62177/jaet.v2i4.899>

## 1.Introduction

Navigating complex university buildings can be challenging for students, faculty, and visitors at the University of the Fraser Valley (UFV). Traditional wayfinding methods such as printed maps, artistic paintings, static signage, and verbal directions are often outdated, unclear, or inaccessible during emergencies <sup>[1]</sup>. As academic buildings expand and space usage becomes more dynamic, the lack of a reliable digital indoor navigation system increasingly hinders accessibility and operational efficiency. New students and visitors often struggle to locate classrooms, offices, and facilities, which can lead to frustration, reduced productivity, and lower campus satisfaction.

In recent years, the development of indoor navigation systems has gained growing attention due to their potential to enhance accessibility, user experience, and spatial awareness within enclosed environments <sup>[2]</sup>. Unlike outdoor navigation systems that rely on satellite-based positioning such as GPS, indoor navigation faces unique challenges. These include limited signal penetration, complex building geometries, and multi-level routing constraints <sup>[3]</sup>. Addressing these challenges requires intelligent algorithms for route optimization, efficient data structures for spatial representation, and an integrated platform capable of real-time user interaction <sup>[4]</sup>.

The A\* (A-star) algorithm has been widely recognized as one of the most effective pathfinding algorithms due to its balance between accuracy and computational efficiency <sup>[3]</sup>. By integrating A\* with spatial databases such as PostgreSQL and PostGIS, pathfinding performance and spatial query efficiency can be significantly improved. Moreover, recent advances in web

frameworks, such as NestJS, and cross-platform mobile technologies, like React Native, enable the development of scalable, responsive, and user-friendly applications that can operate seamlessly across Android and iOS platforms.

The objective of this project is to design and implement a full-stack, cross-platform indoor navigation module as part of the UFV Campus App<sup>[5]</sup>. The system incorporates intelligent room search, interactive floor plan visualization, and real-time turn-by-turn navigation using WebSocket communication. The backend, developed with NestJS and PostGIS, manages spatial data and routing operations, while the React Native frontend provides an intuitive interface for end users. This work demonstrates how the integration of modern software frameworks, spatial databases, and pathfinding algorithms can provide an efficient and scalable indoor navigation solution for complex institutional environments.

## 2. Literature Review

Indoor navigation differs from outdoor routing in important ways: GPS is unreliable indoors, building interiors present complex multi-level geometries, and dynamic obstacles or temporary closures can require online replanning. Surveys of indoor positioning and wayfinding systems summarize these constraints and classify approaches by sensing modality (RF-based, vision-based, sensor fusion), map representation, and routing strategy<sup>[6,7]</sup>. These reviews emphasize that an effective indoor navigation solution must couple accurate indoor positioning with efficient pathfinding and a performant spatial data backend.

### 2.1 Classical and refined shortest-path algorithms (graph search)

The shortest path on graphs is a fundamental building block for routing. Dijkstra's algorithm (1959)<sup>[8]</sup> provides a shortest path solution for nonnegative edge weights and remains a baseline for network routing. Its worst-case time complexity and exhaustive expansion motivate the use of heuristics for faster single-destination queries in large graphs.

The most popular autonomous agent approach for path finding is A\* search<sup>[9]</sup>. A\* (A-star) extends Dijkstra's algorithm by adding an admissible heuristic to guide search toward the goal, drastically reducing explored nodes in many practical cases while preserving optimality when the heuristic is admissible<sup>[10]</sup>. A\* balances exploring the cheapest paths and moving towards the destination, using a simple but effective cost function shown in Equation 1<sup>[8]</sup>. A\*'s simplicity and proven performance explain its widespread adoption in robotics, games, and GIS-based routing.

$$f(n) = g(n) + h(n) \quad \text{Equation 1}$$

where  $g(n)$  is the actual cost from the start node to the current node  $n$ , and  $h(n)$  is the heuristic estimate of the cost from  $n$  to the destination.

For environments where costs change or sensing updates reveal previously unknown obstacles, incremental replanning algorithms like D\* (Dynamic A\*)<sup>[10]</sup> and Focused D\*<sup>[11]</sup> support efficient online repair of previously computed paths rather than recomputing from scratch, an important capability for mobile agents navigating dynamic indoor environments, e.g., temporary blockages, emergency routing.

Several algorithmic refinements improve A\*'s runtime or produce more realistic results, such as less grid-constrained routes. Theta\*<sup>[12]</sup> and related any-angle variants produce shorter, more natural paths on grid-based maps by allowing straight-line (non-grid-aligned) transitions between nodes, trading off some theoretical guarantees for better path quality in practice. Jump Point Search (JPS)<sup>[13]</sup> and its variants prune symmetric or redundant expansions on uniform-cost grids, delivering substantial speedups while preserving optimality on grid maps<sup>[14]</sup>. These methods are particularly relevant when using rasterized floor plans or occupancy grids as the routing substrate, which is common in robotics and certain indoor navigation implementations.

For high-dimensional motion planning, such as robot arms and mobile robots with kinematic constraints, sampling-based planners like PRM and RRT<sup>[15-17]</sup> are the dominant approaches. While PRM/RRT are less common for simple pedestrian navigation inside buildings, where 2D graph-based routing suffices, they matter when vehicle dynamics, nonholonomic constraints, or continuous collision-free motion in cluttered 3D spaces are required.

There are some other approaches or different data work reported for indoor pathing finding, such as Qi et al. used building information model <sup>[18]</sup>, Zhou et al. 2022 hierarchical landmark representation <sup>[19]</sup>, Rodenberg et al. 2016 used A\* on the point cloud data <sup>[20]</sup>, Gorro et al 2024 used a computer vision You Only Look Once (YOLO) model to avoid obstacles <sup>[21]</sup>, Khairnar et al 2024 proposed a Adaptive Multi-criteria Indoor Pather (AMIP) algorithm <sup>[22]</sup>, Zhao et al. 2022 used a weighted octree-based 3D approach <sup>[23]</sup>.

## 2.2 Spatial databases and server-side routing (PostGIS / pgRouting)

For campus-scale systems that must serve many clients and support querying of semantic indoor objects (rooms, amenities), using a geospatial RDBMS is common. PostGIS <sup>[23]</sup> provides robust spatial types and operations; pgRouting (an extension) implements graph routing functions, such as A\*, Dijkstra, bidirectional variants, etc., directly inside the database, enabling server-side shortest-path computation and integration with Geographic Information System (GIS) workflows. Several applied studies demonstrate pgRouting for indoor/outdoor routing and multi-floor adaptations (with caveats around 3D topology handling). Using a PostGIS + pgRouting backend simplifies spatial queries, topology management, and integration with web/service layers (e.g., NestJS).

From the literature, several practical insights emerge for indoor campus routing systems: A\* remains a solid default\* for database-backed routing, especially when you can precompute a compact graph, but specialized variants, such as JPS <sup>[14]</sup>, Theta\* <sup>[12]</sup>, Polyanya can reduce latency or improve path realism depending on the map representation <sup>[24]</sup>. Replanning is essential in dynamic settings, D\* variants provide efficient online repairs when edges become blocked or costs change.

Server-side routing with PostGIS/pgRouting is attractive for maintainability, semantic queries, and multi-client scale, but for ultra-low-latency guidance or offline operation, it may be beneficial to cache or precompute routes on the client. Multi-floor 3D routing requires careful topology design (vertical connectors, z-values) because many routing libraries are primarily 2D. Integration with positioning remains a systems challenge: localization noise should be reflected in the guidance logic (e.g., tolerance for corridor vs. room-level instructions), and UX design must handle ambiguity gracefully.

Conventional mobile mapping solutions <sup>[25]</sup> using SVG or HTML5 Canvas can become inefficient when rendering complex geometric data, especially at interactive frame rates. To address this, the project employs React Native Skia, a high-performance graphics engine that leverages GPU acceleration for real-time rendering. Skia enables the system to smoothly handle floor plans

composed of hundreds of polygonal elements, supporting pan, zoom, and room highlighting without performance degradation.

## 3. Methodology

Our work consists of four major phases. (1) Facility map digitization: Using ArcGIS, we digitized the floor plan by creating a centerline representation of each hallway as the path network and defining door center locations as point features (Figure 1). Each door point was assigned a room number to support location-based search. We also conducted polygon closure and intersection checks. (2) System analysis and design: We conducted a comprehensive review of the relevant literature and evaluated existing GitHub projects. Based on this assessment, we selected the A\* algorithm as the primary routing method. (3) Software development and system implementation: We developed the UFV App interface and configured a local host environment to manage and serve the backend database. The communication plan is shown in Table 1. (4) Quality assurance and user testing: We completed unit tests using Jest, integration tests using Supertest, and accessibility tests using React native testing library. The UFV App prototype is currently being presented to the internal stakeholders for user evaluation.

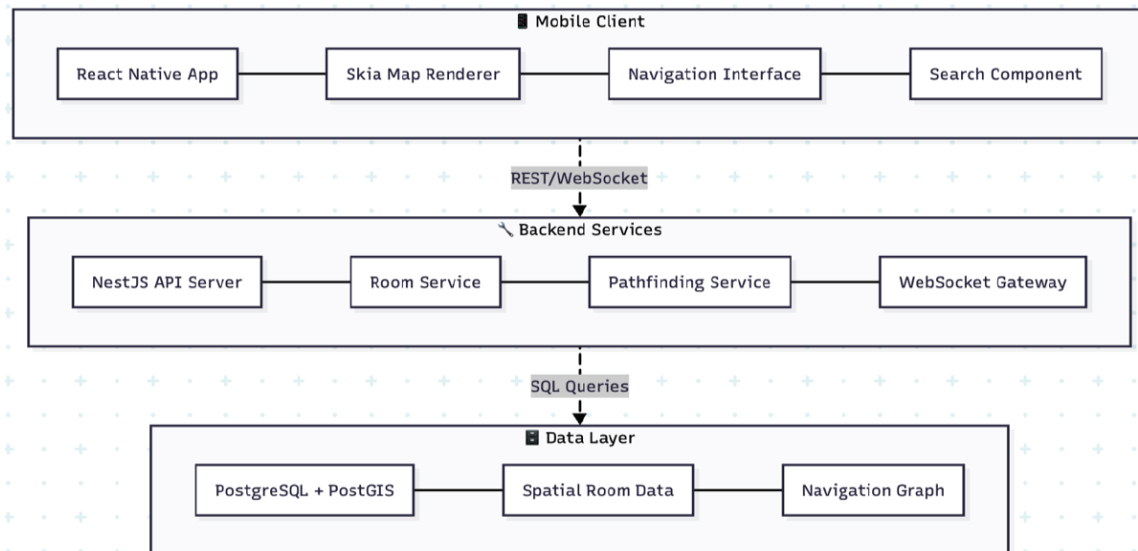
Table 1: Communication Patterns

Type	Protocol	Use Case	Performance
Room Search	REST	Autocomplete queries	<100 ms
Map Data	REST	Floor plan geometries	<50 ms
Navigation	WebSocket	Real-time updates	<50 ms

Figure 1: Sample floor plan provided by UFV in the public domain and digitized in our work.



Figure 2: The overall architecture of our system.



## 4.Results

All functional requirements defined for the UFV indoor navigation system were successfully achieved. The room search module supports fast and intuitive lookup through an autocomplete interface enhanced with category-based filtering and fuzzy string matching, enabling users to locate rooms even with partial or approximate queries. Floor plan rendering operates with high smoothness and visual accuracy, powered by GPU-accelerated Skia rendering that supports seamless zooming and panning. The turn-by-turn navigation feature is fully implemented, delivering real-time routing updates through WebSocket communication and supporting voice-guidance preparation for future enhancements. The application provides verified cross-platform support, operating consistently on both iOS and Android devices. Furthermore, the system meets or exceeds all performance requirements, achieving a stable 60 FPS rendering rate and maintaining navigation query response times under 50 ms, ensuring a reliable, responsive user experience.

The system includes several key architectural and implementation-level accomplishments. The modular software architecture provides clear separation among rendering, routing, and data-access layers, enabling independent maintenance and future extensibility. The spatial database backbone, built on an optimized PostGIS schema, employs targeted spatial indexing to support low-latency navigation queries and is scalable for campus-wide expansion. TypeScript-based shared interfaces reduce integration overhead across components, improving maintainability. Production deployment was streamlined through Docker containerization, enabling consistent server configuration and simplifying updates.

The project also incorporates notable innovations. The adoption of React Native Skia enables GPU-accelerated vector mapping, providing performance and visual fidelity uncommon in mobile indoor navigation systems. The real-time navigation streaming mechanism built on WebSockets allows for continuous route updates and highly dynamic interactions. Additionally, the tight integration of spatial database optimization demonstrates how PostGIS can be effectively adapted for low-latency, mobile-oriented pathfinding.

All the codes are available on GitHub <sup>[26]</sup>.

## 5. Discussion

During development, several technical challenges were encountered and systematically addressed, contributing valuable insights into mobile indoor navigation design.

One major challenge involved SVG rendering performance. Early prototypes showed limitations in rendering complex vector floor plans using traditional SVG engines. By transitioning to Skia's GPU-accelerated rendering pipeline, we achieved substantial improvements in frame stability and visual accuracy. This highlights the importance of performing early performance benchmarking in mobile applications where GPU constraints vary significantly across devices.

A second issue involved UTM-to-screen coordinate distortions, which arose when mapping georeferenced building data to device-level coordinates. We addressed this by calibrating a precise affine transformation using known building metrics. This effort demonstrated that reliable navigation requires rigorous validation of coordinate systems using real-world reference points rather than relying solely on exported CAD/GIS data.

Ensuring UI responsiveness during WebSocket-driven updates presented another medium-level challenge. Without careful state management, rapid routing updates caused UI blocking and stuttering. By batching updates and employing Reanimated worklets, we achieved smooth animations even under high update frequency. This underscores that real-time mobile applications require dedicated design strategies to avoid overwhelming the rendering thread.

Cross-platform rendering consistency also required extensive testing. While React Native provides shared logic, platform-specific rendering differences—particularly in graphical surfaces and gesture handling—necessitated targeted optimizations. This finding reinforces that achieving consistent design systems on iOS and Android requires platform-aware implementation rather than assuming uniform behavior.

Finally, spatial data accuracy emerged as a critical concern. Errors in room positions, hallway geometry, or door alignment directly degrade navigation accuracy. A validation pipeline with manual verification checkpoints proved essential. This experience confirms that data quality assurance is non-negotiable in navigation applications, especially those intended for institutional use.

## 6. Conclusion

In summary, this project delivers both technical innovation and meaningful practical impact for the UFV community. Technically, the system pioneers the use of React Native Skia to achieve high-performance, vector-based indoor mapping, incorporates an optimized PostGIS schema with spatial indexing for low-latency navigation queries, and implements a WebSocket-driven routing service capable of providing real-time updates. Practically, the solution enhances campus accessibility and wayfinding by replacing static floor maps with an interactive mobile application, establishes a scalable foundation for campus-wide deployment with minimal infrastructure overhead, and demonstrates the viability of in-house development for specialized indoor navigation systems. The system is production-ready for pilot deployment and is supported by comprehensive documentation to ensure maintainability, scalability, and future feature development. Overall, this work

fulfills the CIS 440 capstone requirements by effectively integrating theoretical knowledge with practical implementation to address a real-world institutional need.

## Funding

This research is partially supported by the University of the Fraser Valley Scholarly Activity Fund 2025.

## Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Reference

- [1] Wu, M., Qiao, L., Wu, Z., Hou, Z., Chen, S., & Lv, G. (2024). What can cartographers learn from artistic paintings when stylizing maps? A preliminary synthesis from the perspective of visual neuroscience. *Cartography and Geographic Information Science*, 52(1), 35–54. <https://doi.org/10.1080/15230406.2024.2370878>
- [2] Zhang, J., Han, G., Sun, N., & Shu, L. (2017). Path-loss-based fingerprint localization approach for location-based services in indoor environments. *IEEE Access*, 5, 13756–13769. <https://doi.org/10.1109/ACCESS.2017.2728789>
- [3] Simões, W. C. S., Machado, G. S., Sales, A. M. A., de Lucena, M. M., Jazdi, N., & de Lucena, V. F., Jr. (2020). A review of technologies and techniques for indoor navigation systems for the visually impaired. *Sensors (Basel)*, 20(14), 3935. <https://doi.org/10.3390/s20143935>
- [4] Hart, P. E., Nilsson, N. J., & Raphael, B. (1968). A formal basis for the heuristic determination of minimum cost paths. *IEEE Transactions on Systems Science and Cybernetics*, 4(2), 100–107. <https://doi.org/10.1109/TSSC.1968.300136>
- [5] Kamboj, N., & Zhang, K. (2025). Low-cost CCTV repurposing for sustainable parking management: A non-AI computer vision case study. *Journal of Advances in Engineering and Technology*, 2(4). <https://doi.org/10.62177/jaet.v2i4.666>
- [6] El-Sheimy, N., & Li, Y. (2021). Indoor navigation: State of the art and future trends. *Satellite Navigation*, 2(7). <https://doi.org/10.1186/s43020-021-00041-3>
- [7] Rahmani, V., & Pelechano, N. (2022). Towards a human-like approach to path finding. *Computers & Graphics*, 102, 164–174. <https://doi.org/10.1016/j.cag.2021.08.020>
- [8] Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, 1, 269–271. <https://doi.org/10.1007/BF01386390>
- [9] Rachmawati, R., & Gustin, L. (2020). Analysis of Dijkstra's algorithm and A\* algorithm in shortest path problem. *Journal of Physics: Conference Series*, 1566(1), 012061. <https://doi.org/10.1088/1742-6596/1566/1/012061>
- [10] Stentz, A. (1994). Optimal and efficient path planning for partially-known environments. In *Proceedings of the IEEE International Conference on Robotics and Automation* (Vol. 4, pp. 3310–3317). IEEE.
- [11] Stentz, A. (1995). The focussed D\* algorithm for real-time replanning. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI '95)* (pp. 1652–1659). Morgan Kaufmann.
- [12] Daniel, K., Nash, A., Koenig, S., & Felner, A. (2010). Theta\*: Any-angle path planning on grids. *Journal of Artificial Intelligence Research*, 39, 533–579. <https://doi.org/10.1613/jair.2994>
- [13] Harabor, D., & Grastien, A. (2011). Online graph pruning for pathfinding on grid maps. In *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence* (pp. 1118–1123). AAAI Press.
- [14] Harabor, D., & Grastien, A. (2012). The JPS pathfinding system. In *Proceedings of the Fifth Annual Symposium on Combinatorial Search* (Vol. 3, No. 1, pp. 207–208). AAAI Press.
- [15] Kavraki, L. E., Švestka, P., Latombe, J.-C., & Overmars, M. H. (1996). Probabilistic roadmaps for path planning in high-dimensional configuration spaces. *IEEE Transactions on Robotics and Automation*, 12(4), 566–580. <https://doi.org/10.1109/70.508439>
- [16] LaValle, S. M., & Kuffner, J. J. (2001). Randomized kinodynamic planning. *The International Journal of Robotics Research*, 20(5), 378–400. <https://doi.org/10.1177/02783640122067453>
- [17] Sriramulu, R., Yadav, A., & Pal, S. B. (2025). Fast and efficient indoor navigation: A hybrid pathfinding approach using rapidly-exploring random tree (RRT)-connect and Dijkstra's algorithm. *PeerJ Computer Science*, 11, e3028. <https://doi.org/10.7717/peerj.csci.11111>

org/10.7717/peerj-cs.3028

- [18] Qi, T., Liu, D., Guo, Y., Zhou, X., Zhao, X., Huang, X., & Wang, Z. (2025). Toward efficient and agent-scalable indoor pathfinding: Intelligent navigation-ability-driven indoor map generation using building information model. *Architectural Engineering and Design Management*, 1–22. <https://doi.org/10.1080/17452007.2025.2451817>
- [19] Zhou, J., Yang, H., Shen, J., & Zhu, L. (2024). Indoor navigation map design based on spatial complexity. *Cartography and Geographic Information Science*, 52(1), 69–81. <https://doi.org/10.1080/15230406.2024.2339296>
- [20] Rodenberg, O. B. P. M., Verbree, E., & Zlatanova, S. (2016). Indoor A\* pathfinding through an octree representation of a point cloud. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-2/W1, 249–255. <https://doi.org/10.5194/isprs-annals-IV-2-W1-249-2016>
- [21] Gorro, K., Roble, L., Magana, M. A., & Buot, R. P. (2024). Prototype of an indoor pathfinding application with obstacle detection for the visually impaired. *International Journal of Advanced Computer Science and Applications*, 15(9). <https://doi.org/10.14569/IJACSA.2024.0150987>
- [22] Khainar, A., Dhaske, A., Patil, A., Dhote, S., & Tamkhade, J. (2024). Adaptive multi-criteria indoor pathfinding algorithm using dynamic user preference and real-time data. *Journal for Research in Applied Science and Engineering Technology*. <https://doi.org/10.22214/ijraset.2024.65564>
- [23] PostGIS Project Steering Committee. (2024). PostGIS: Spatial and geographic objects for PostgreSQL (Version 3.5.4) [Computer software]. <https://postgis.net>
- [24] Zhou, B., Li, Q., Mao, Q., Tu, W., & Zhang, X. (2015). Activity sequence-based indoor pedestrian localization using smartphones. *IEEE Transactions on Human-Machine Systems*, 45(5), 562–574. <https://doi.org/10.1109/THMS.2014.2368092>
- [25] Zhao, J., Xu, Q., Zlatanova, S., Liu, L., Ye, C., & Feng, T. (2022). Weighted octree-based 3D indoor pathfinding for multiple locomotion types. *International Journal of Applied Earth Observation and Geoinformation*, 112, 102900. <https://doi.org/10.1016/j.jag.2022.102900>
- [26] Massesa, B., & Zhang, F. (2025). UFV-Pathfinding: Indoor navigation system for the University of the Fraser Valley [Computer software]. GitHub. <https://github.com/borismassesa/UFV-Pathfinding>

# Research on Optimization Pathways for Reverse Logistics Networks Based on Ecological Civilization

Jinzhao Song\*, Xiaofeng Zhang

School of Chemistry and Chemical Engineering, Ningxia University, 750021, China

\*Corresponding author: Jinzhao Song, 12024140163@stu.nxu.edu.cn

**Copyright:** 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

**Abstract:** Against the backdrop of global urbanization and the transition to a circular economy, the design of reverse logistics networks for municipal solid waste (MSW) is a critical link in advancing urban sustainability. This study aims to develop a ubiquitous optimization framework for MSW reverse logistics networks. By analyzing the current state of urban waste management, a three-tier network structure comprising generation sources, transfer stations, and treatment centers was constructed. Furthermore, regional waste generation quantity was introduced as a fuzzy parameter to address uncertainty. Based on this, a mixed-integer programming (MIP) model was established with the objective of minimizing the total system cost, and a genetic algorithm was designed to solve it. An empirical case study of Hefei City, China, demonstrates that the optimized model can effectively reduce the total network operating cost by approximately 8%. The significant decrease in transportation costs directly enhances the overall efficiency of waste treatment. This research provides a transferable methodology and decision-making support for addressing the solid waste management challenges faced by cities worldwide.

**Keywords:** Reverse Logistics; Municipal Solid Waste (MSW); Ecological Civilization; Network Design

**Published:** Dec 9, 2025

**DOI:** <https://doi.org/10.62177/jaet.v2i4.920>

## 1.Introduction

### 1.1 Research Background and Importance

Ecological civilization serves as a critical indicator of societal development progress, with its core focus on harmonizing environmental protection with sustainable socio-economic development.<sup>[1]</sup> Since the concept of “ecological civilization” was introduced in China, its connotation has been continuously enriched, and it has been established at the strategic national level as a scientific framework guiding the harmonious coexistence of humans and nature. Particularly in the realm of urban environmental governance, the shift from emphasizing “the construction of municipal solid waste treatment facilities” to implementing “the strictest ecological protection systems” underscores that building an efficient waste management system has become a key task in advancing ecological civilization.<sup>[2-4]</sup>

Against this backdrop, achieving the “reduction, recycling, and harmless treatment” of waste is not only a concrete implementation of the “tri-integrated” development strategy but also an inevitable trend in enhancing urban governance efficiency.<sup>[5-7]</sup> This paper begins with the design of a reverse logistics network for municipal solid waste, viewing it as a central link in achieving source reduction and systemic optimization. The study aims to provide actionable technical pathways for ecological civilization development by optimizing the structure of the logistics network.

## 1.2 Research Objectives

This study aims to construct and optimize a three-tier structure for the municipal solid waste (MSW) reverse logistics network, investigating its structural role in enhancing processing efficiency and controlling environmental risks. Through modern optimization modeling approaches, particularly genetic algorithms and fuzzy parameter handling, the comprehensive costs and operational effectiveness of the network design are evaluated. Although the three-tier network has been theoretically recognized as a foundation for process optimization, this paper delves into the precise coverage of waste generation sources, the scientific layout of transfer facilities, the synergistic matching of disposal terminals, and the system's stability in responding to fluctuations in waste volume. Ultimately, this study seeks to provide a clearer and more quantifiable planning solution, enabling decision-makers to understand the fundamental value of the reverse logistics network structure in achieving the “reduction, recycling, and harmless treatment” of waste, while offering new implementation pathways for advancing the modernization of urban environmental governance systems.

## 1.3 Data Source Description

This study obtained data on municipal solid waste from the Hefei Municipal Bureau of Ecology and Environment (<http://zwgk.hefei.gov.cn/public/14011/106651511.html>), administrative division data from the China National Geographic Information Resource Directory Service System (<https://www.webmap.cn>), and population density data from the Hefei Municipal People's Government Seventh National Population Census Bulletin (<http://www.hefei.gov.cn/xxgk/gsgg/106488113.html>).

## 2.Design Principles

### 2.1 Hierarchical Structure

The nodal structure of the municipal solid waste (MSW) reverse logistics network primarily consists of the generation tier (households), the transfer tier, and the processing tier (incineration plants and landfills).<sup>[8]</sup> The main process flow involves waste separation and disposal by residents, followed by the transfer stations transporting different categories of waste to corresponding treatment facilities. The critical linkages in this hierarchy are between the generation tier and the transfer tier, and between the transfer tier and the processing tier.<sup>[9]</sup>

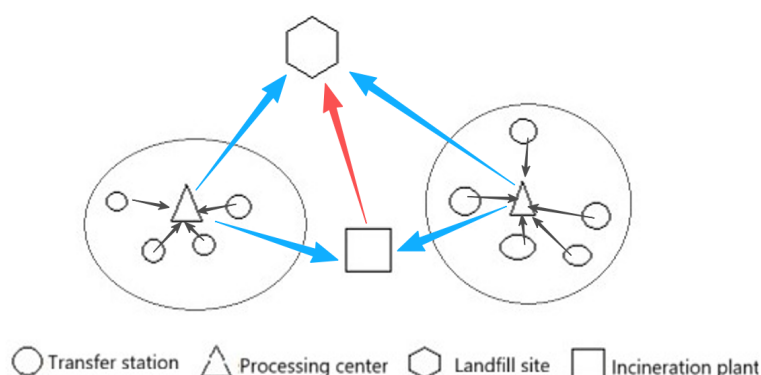
### 2.2 Primary Functions across Tiers

**Generation Tier:** This tier primarily consists of households that generate waste through daily activities. The waste is either collected regularly by waste collection vehicles or transported by residents to designated drop-off points, after which it is forwarded to the transfer tier.<sup>[10]</sup>

**Transfer Tier:** The transfer tier receives waste from households and other sources, sorts it based on its characteristics, and dispatches different categories of waste to corresponding processing centers.<sup>[11-12]</sup>

**Processing Tier:** This tier mainly includes waste incineration plants and landfills. Combustible waste is sent to incineration plants, where it is converted into thermal energy or electricity for residential use. Non-combustible waste is compacted and transported to landfills for natural restoration. Staff working at landfills or incineration facilities must possess relevant qualifications to prevent improper handling that could harm the environment and human health.<sup>[13-14]</sup>

Figure1: Topology of reverse logistics network for municipal solid waste



### 3. Network Model and Algorithm Design

#### 3.1 Model Assumptions

Given the inherent difficulty in quantifying real-world waste generation volumes, this study develops a municipal solid waste reverse logistics network model that treats waste quantity as a fuzzy parameter. The model's framework is built around a three-tier structure comprising generation, transfer, and processing layers, with the following operational sequence: At the transfer tier, collected municipal solid waste undergoes detailed classification and documentation. In compliance with national regulations, this tier then conducts harmless treatment and usability assessment, directing recyclable materials to appropriate treatment facilities while transporting non-recyclable waste to either incineration plants or landfills. The processing tier operates under strict regulatory requirements, where specialized facilities manage recyclable materials through formal treatment and reprocessing channels. Meanwhile, non-recyclable waste sent to incineration plants and landfills receives segregated handling based on its characteristics and environmental impact.

#### 3.2 Definition of Parameters and Variables

Table 1. Relevant definitions of variables and symbols

Category	Symbol	Meaning
Superscript and subscript	a	Collection point
	b	Transfer station
	c	Processing center
	d	Landfill
	e	Incineration plant
	I	Number of candidate collection points
	J	Number of candidate transfer stations
	K	Number of candidate processing centers
	H	Number of candidate landfills
	L	Number of candidate incineration plants
Parameter	$Q_l^a$	Fuzzy waste generation quantity
	$C_j^b$	Maximum transport capacity of transfer point
	$C_k^c$	Maximum processing capacity of processing point
	$G_j^a$	Fixed costs of transit
	$G_j^b$	Fixed cost of processing center
	$G_k^c$	Fixed costs of landfills
	$G_d^h$	Fixed cost of waste incineration plant
	$G_l^e$	Fixed cost of power plant
	$Y_j^a$	Unit operating cost of transfer point
	$Y_j^b$	Unit operating cost of processing point
	$Y_k^c$	Unit operating cost of landfill site
	$Y_h^d$	Unit operating cost of waste incineration plant
	$D_{ij}^{ab}$	Distance from generation point to transfer point
	$D_{jk}^{bc}$	Distance from transfer point to processing point
	$D_{jh}^{bc}$	Distance from transfer point to landfill site
	$D_{jl}^{bc}$	Distance from transfer point to waste incineration plant
	$S_{ij}^{ab}$	Unit transportation cost from transfer point to processing point
	$S_{jk}^{bc}$	Unit transportation cost from transfer point to landfill site

Category	Symbol	Meaning
	$S_{jh}^{bd}$	Unit transportation cost from transfer point to waste incineration plant
	$\beta$	Proportion of hazardous waste at transfer point
	$t$	Treatment cost of hazardous waste at treatment point
	$\alpha$	Ratio of waste available at transfer point
Decision variable	$q_{ij}^{ab}$	Quantity of waste transported from generation point to transfer point
	$q_{jk}^{bc}$	Quantity of waste from transfer point to disposal point
	$q_{jh}^{bd}$	Quantity of waste from transfer point to landfill site
	$q_{jl}^{be}$	Quantity of waste from transfer point to waste incineration plant
	$x_i^a$	1 if location j is selected for a transfer station; 0 otherwise
	$x_j^b$	1 if location k is selected for a processing center; 0 otherwise
	$x_k^c$	1 if location h is selected for a landfill; 0 otherwise
	$x_l^e$	1 if location l is selected for an incineration plant; 0 otherwise

### 3.3 Model Formulation

A reverse logistics network model for municipal solid waste (MSW) is developed with the objective of minimizing total costs, incorporating waste quantity as a fuzzy parameter. Taking Hefei City as a case study and based on the actual local conditions of MSW management, the primary cost components considered in the operation of the reverse logistics network include: ① infrastructure construction cost, ② facility maintenance cost, ③ transportation cost for MSW between different tiers, and ④ treatment cost for non-recyclable waste. As assumed in this study, the recyclable waste is fully utilized, meaning its residual value offsets the processing costs incurred.

**Infrastructure Construction Cost for Reverse Logistics:** This refers to the expenses required for building the fundamental facilities necessary for the normal operation of the reverse logistics network. These facilities include collection points, transfer stations, processing centers, waste incineration plants, and landfills.

$$Z_1 = \sum_{i=1}^l G_i^a + \sum_{j=1}^l G_j^b x_j^b + \sum_{k=1}^k G_k^c x_k^c + \sum_{h=1}^h G_h^d x_h^d + \sum_{l=1}^l G_l^e x_l^e$$

**Facility Maintenance Cost:** This refers to the expenses incurred to maintain equipment in proper working condition and ensure its normal operation.

$$Z_2 = \sum_{i=1}^l Y_i^a Q_i^a x_i^a + \sum_{i=1}^l \sum_{j=1}^l q_{ij}^{ab} Y_j^b x_j^b + \sum_{i=1}^l \sum_{j=1}^l q_{gh}^{bd} Y_h^d x_h^d + \sum_{l=1}^l \sum_{j=1}^l q_{jl}^{be} Y_l^e x_l^e + \sum_{j=1}^j \sum_{k=1}^k q_{jk}^{bc} Y_k^c x_k^c$$

**Transportation Cost between Tiers:** This refers to the expenses incurred during the transportation of municipal solid waste between different tiers of the network.

$$Z_3 = \sum_{i=1}^i \sum_{j=1}^l q_{ij}^{ab} D_{ij}^{ab} S_{ij}^{ab} + \sum_{j=1}^j \sum_{k=1}^k q_{jk}^{bc} D_{jk}^{bc} S_{jk}^{bc} + \sum_{h=1}^h \sum_{j=1}^j q_{gh}^{bd} D_{jh}^{bd} S_{jh}^{bd} + \sum_{l=1}^l \sum_{j=1}^j q_{jl}^{be} D_{jl}^{be} S_{jl}^{be}$$

**Treatment Cost for Non-Recyclable Waste:** This refers to the expenses associated with the landfilling or incineration of non-recyclable waste.

$$Z_4 = \sum_{i=1}^l G_i^a + \sum_{j=1}^l G_j^b x_j^b + \sum_{k=1}^k G_k^c x_k^c + \sum_{h=1}^h G_h^d x_h^d + \sum_{l=1}^l G_l^e x_l^e$$

In summary, the objective function for minimizing the total cost in the municipal solid waste reverse logistics network is formulated as follows:

$$\min Z = Z_1 + Z_2 + Z_3 + Z_4$$

## 4. Case analysis

### 4.1 Fuzzy output of domestic waste in Hefei

Taking Hefei City as an example, by the end of 2020, its permanent resident population had exceeded 9 million, with a total municipal solid waste generation of 2.7091 million tons—ten times higher than a decade ago. This massive volume of waste has placed a heavy burden on the disposal system. The Hefei Longquanshan Landfill covers an area of approximately 1 million square meters. The first phase of the project commenced in 2002 and was officially put into operation in June 2004.

By the end of 2013, the first-phase landfill reached capacity and underwent transitional closure, filling up half a year earlier than originally planned. The second phase was put into use in December 2013 and currently processes over 2,000 tons of waste per day. The environmental quality of the Longquanshan Landfill is suboptimal, with existing issues of pollution and potential safety hazards. To address this situation, it is particularly important to optimize the location of landfills in Hefei and to design and improve the reverse logistics network for municipal solid waste.

#### 4.2 Infrastructure construction cost and operation and maintenance cost

Since the municipal solid waste generation data for each district (county) of Hefei City has not been officially reported, and given that local population size significantly influences waste generation, the waste quantities for these areas were estimated based on their respective population proportions. Furthermore, to account for potential fluctuations due to various influencing factors, deviation ranges were incorporated into the model, with a lower limit of -5% and an upper limit of +10%. The population figures and corresponding fuzzy waste generation amounts for each region are presented in the table below.

Table 2. Municipal Solid Waste Generation and Fuzzy Quantities by Region in Hefei (2020)

Regions	Waste Generation Quantity (tons)	Fuzzy Waste Generation Quantity (tons)
Yaohai District	249237	( 236775 ) ( 249237 ) ( 274160 )
Luyang District	201557	( 191479 ) ( 201557 ) ( 221713 )
Shushan District	302877	( 333165 ) ( 302877 ) ( 287733 )
Baohe District	351912	( 387103 ) ( 351912 ) ( 334316 )
High-Tech Industrial Development Zone	78022	( 85824 ) ( 78022 ) ( 74121 )
Economic Development Zone	161191	( 177310 ) ( 161191 ) ( 153131 )
Xinzhan District	134913	( 148404 ) ( 134913 ) ( 128167 )
Changfeng County	226752	( 249427 ) ( 226752 ) ( 215414 )
Feidong County	255739	( 281313 ) ( 255739 ) ( 242952 )
Feixi County	279850	( 265858 ) ( 279850 ) ( 307835 )
Lujiang County	256823	( 282505 ) ( 256823 ) ( 243982 )
Chaohu City	210226	( 199715 ) ( 210226 ) ( 231249 )

#### 4.3 Coordinates of main transfer stations in Hefei

In recent years, China has significantly intensified the construction of environmental protection infrastructure, leading to the establishment of numerous waste treatment enterprises across various provinces and municipalities. Through an investigation into the fixed facility construction costs and relevant transportation and management expenses of waste treatment enterprises in Hefei, and by integrating the reverse logistics planning model for municipal solid waste developed in this study, the fixed and operational costs for the nodes within the reverse logistics network have been derived, as summarized in Table 3. It should be noted that, since transfer stations, waste incineration plants, and landfills rely on existing local institutions, their capital construction costs are considered zero and are thus excluded from this reverse logistics network analysis. Consequently, the operational costs primarily reflect the expenses associated with the treatment of municipal solid waste.

Table 3. Municipal Solid Waste Generation and Fuzzy Quantities by Region in Hefei (2020)

Facility Type	Construction Cost (10,000 CNY)	Maintenance Cost (10,000 CNY)
Transfer Station	500	100
Processing Center	500	200
Waste Incineration Plant	100	10
Landfill	100	10

#### 4.4 Other relevant parameters

First, Transportation Cost for Collection Vehicles. Based on surveys, the transportation cost for municipal solid waste in Hefei is approximately 5 CNY/ton. The costs for transporting waste from transfer stations to landfills and power plants are classified as municipal solid waste transportation expenses. Therefore, the final transportation costs between respective tiers are defined

as  $S_{ij}^{ab} = S_{jk}^{bc} = S_{jh}^{bd} = 5$  CNY/ton/km.

Second, Recyclable Proportion of Municipal Solid Waste. As the recyclable proportion of municipal solid waste varies across periods, relevant data indicate that the waste recovery rate in Hefei is 20%, i.e.,  $\alpha=20\%$ .

Third, Processing Capacity of Facilities at Each Tier in the Reverse Logistics Network. According to investigations, the average daily processing capacity of transfer stations in Hefei is 50 tons/day, i.e.,  $C_g^b = C_k^c = 50$  tons.

#### 4.5 Operation results

By applying genetic algorithm encoding and MATLAB (2018) software to solve and analyze the model, optimized data regarding the location, routing, and flow volume of infrastructure at each level were obtained. The selected parameters included a population size of  $n = 30$ , a chromosome crossover rate of 0.7, a mutation probability of 0.1, and 400 iterations, resulting in a final objective function value of  $z = 62,770$  million CNY.

The iteration curve of the algorithm is shown in Figure 2. As the genetic algorithm progressively optimized the model, the objective function value gradually decreased and eventually stabilized. Figure 3 illustrates the optimized reverse logistics network structure for municipal solid waste in Hefei, while Table 4 presents the corresponding relationships among the different layers of the reverse logistics network, which includes 10 transfer stations, 2 treatment centers, 1 incineration plant, and 1 landfill site. Tables 5 and 6 provide the distances between transfer stations and treatment centers in the reverse logistics network and the corresponding transportation instruction matrix, respectively.

Figure2: Iterative graph of genetic algorithm

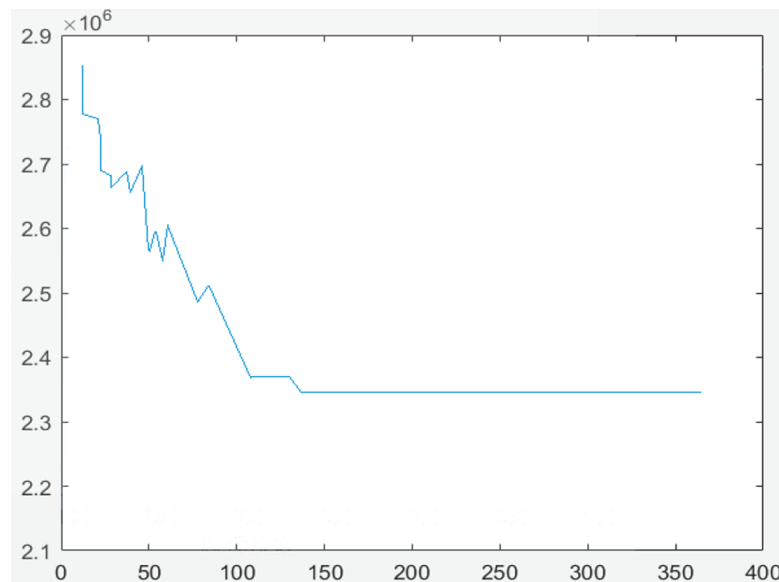


Figure3: Site selection of domestic waste reverse logistics network in Hefei

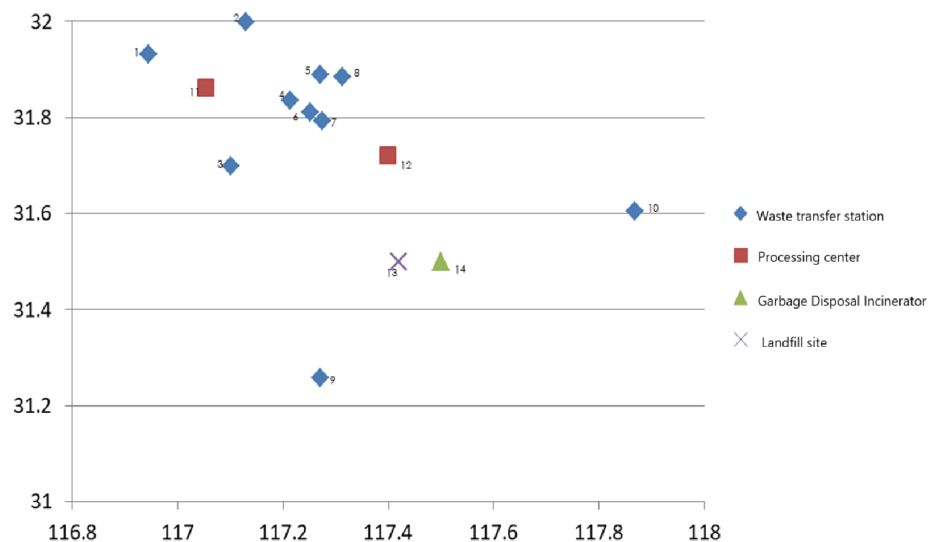


Table 4. Municipal Solid Waste Generation and Fuzzy Quantities by Region in Hefei (2020)

Transfer station	Disposal center	Landfill site	Garbage Disposal Incinerator
1, 2, 3, 4	11	13	14
5, 6, 7, 8, 9, 10	12	13	14

Table 5. Transportation distance from waste transfer station to waste treatment center (km)

	Processing center 11	Processing center 12
Waste transfer station 1	13.073	
Waste transfer station 2	18.400	
Waste transfer station 3	16.876	
Waste transfer station 4	15.142	
Waste transfer station 5		17.289
Waste transfer station 6		52.862
Waste transfer station 7		22.314
Waste transfer station 8		20.059
Waste transfer station 9		57.620
Waste transfer station 10		45.994

Table 6. Transportation instruction matrix: recycle bin to processing center

node	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	0	0	0	0	0	0	0	0	0	0	1	0	1	1
2	0	0	0	0	0	0	0	0	0	0	1	0	1	1
3	0	0	0	0	0	0	0	0	0	0	1	0	1	1
4	0	0	0	0	0	0	0	0	0	0	1	0	1	1
5	0	0	0	0	0	0	0	0	0	0	0	1	1	1
6	0	0	0	0	0	0	0	0	0	0	0	1	1	1
7	0	0	0	0	0	0	0	0	0	0	0	1	1	1
8	0	0	0	0	0	0	0	0	0	0	0	1	1	1
9	0	0	0	0	0	0	0	0	0	0	0	1	1	1
10	0	0	0	0	0	0	0	0	0	0	0	1	1	1
11	1	1	1	1	0	0	0	0	0	0	0	1	1	1
12	0	0	0	0	0	0	0	0	0	0	0	0	1	1
13	0	0	0	0	0	0	0	0	0	0	1	1	0	0
14	0	0	0	0	0	0	0	0	0	0	1	1	0	0

## Conclusion

This study constructs a three-layer reverse logistics network structure for municipal solid waste (MSW), comprising the generation layer, transfer layer, and disposal layer, which refines the waste treatment process and helps improve processing efficiency while reducing environmental pollution. Using Hefei City as a case study, the network model was optimized to address key uncertainties such as waste generation volume, and the genetic algorithm was applied to determine the optimal locations and quantities of recycling centers, treatment centers, and landfills. Based on the findings, it is recommended to enhance multi-department and multi-process collaboration in the development of the reverse logistics system, incorporating uncertainties such as weather and traffic into systematic planning and design to achieve a balance of economic, social, and environmental benefits. Furthermore, there should be an active integration of internet and information technologies, such

as cloud management and cloud identification, to innovate the offline waste treatment system, improve resource utilization efficiency, and support sustainable urban development.

## Funding

No

## Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Reference

- [1] Rubio, S., Chamorro, A., & Miranda, F. J. (2008). Characteristics of the research on reverse logistics (1995–2005). *International Journal of Production Research*, 46(4), 1099–1120.
- [2] Agrawal, S., Singh, R. K., & Murtaza, Q. (2015). A literature review and perspectives in reverse logistics. *Resources, Conservation and Recycling*, 97, 76–92.
- [3] Sar, K., & Ghadimi, P. (2023). A systematic literature review of the vehicle routing problem in reverse logistics operations. *Computers & Industrial Engineering*, 177, 109011.
- [4] Ding, L., Wang, T., & Chan, P. W. (2023). Forward and reverse logistics for circular economy in construction: A systematic literature review. *Journal of Cleaner Production*, 388, 135981.
- [5] Letunovska, N., Offei, F. A., Junior, P. A., et al. (2023). Green supply chain management: The effect of procurement sustainability on reverse logistics. *Logistics*, 7(3), 47.
- [6] Mugoni, E., Nyagadza, B., & Hove, P. K. (2023). Green reverse logistics technology impact on agricultural entrepreneurial marketing firms' operational efficiency and sustainable competitive advantage. *Sustainable Technology and Entrepreneurship*, 2(2), 100034.
- [7] Daramola, O. M., Apeh, C. E., Basiru, J. O., et al. (2023). Optimizing reverse logistics for circular economy: Strategies for efficient material recovery and resource circularity. *Journal of Circular Economy and Sustainable Logistics*. (Forthcoming)
- [8] Zhou, J., Yang, S., Feng, H., et al. (2023). Multi-echelon sustainable reverse logistics network design with incentive mechanism for eco-packages. *Journal of Cleaner Production*, 430, 139500.
- [9] Hashmi, R. (2023). Business performance through government policies, green purchasing, and reverse logistics: Business performance and green supply chain practices. *South Asian Journal of Operations and Logistics*, 2(1), 1–10.
- [10] Saxena, N., Sarkar, B., Wee, H. M., et al. (2023). A reverse logistics model with eco-design under the Stackelberg-Nash equilibrium and centralized framework. *Journal of Cleaner Production*, 387, 135789.
- [11] Kannan, D., Solanki, R., Darbari, J. D., et al. (2023). A novel bi-objective optimization model for an eco-efficient reverse logistics network design configuration. *Journal of Cleaner Production*, 394, 136357.
- [12] Lin, J., Li, X., Zhao, Y., et al. (2023). Design a reverse logistics network for end-of-life power batteries: A case study of Chengdu in China. *Sustainable Cities and Society*, 98, 104807.
- [13] Santos, M. J., Jorge, D., Ramos, T., et al. (2023). Green reverse logistics: Exploring the vehicle routing problem with deliveries and pickups. *Omega*, 118, 102864.
- [14] Santos, M. J., Jorge, D., Ramos, T., et al. (2023). Green reverse logistics: Exploring the vehicle routing problem with deliveries and pickups. *Omega*, 118, 102864.

# Construction of $\text{SnS}_2/\text{TiO}_2$ Heterojunction and Study of Its Photocatalytic Performance

Xiaofeng Zhang, Jinzhao Song\*

School of Chemistry and Chemical Engineering, Ningxia University, 750021, China

\*Corresponding author: Jinzhao Song, 12024140163@stu.nxu.edu.cn

**Copyright:** 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

**Abstract:** This thesis primarily prepared  $\text{SnS}_2$  and  $\text{TiO}_2$  monomers, and subsequently fabricated  $\text{SnS}_2/\text{TiO}_2$  composites via a one-step hydrothermal method using tetrabutyl titanate, anhydrous ethanol, and concentrated nitric acid. The morphologies of the two prepared monomers and the  $\text{SnS}_2/\text{TiO}_2$  composite were characterized using scanning electron microscopy (SEM). The photocatalytic performance was evaluated by degrading methylene blue solution under xenon lamp irradiation. The results demonstrate that the  $\text{SnS}_2/\text{TiO}_2$  composite exhibits excellent photocatalytic activity, achieving a degradation rate as high as 92% for the methylene blue dye.

**Keywords:**  $\text{SnS}_2/\text{TiO}_2$  Composite; Photocatalysis; Hydrothermal Method; Methylene Blue

**Published:** Dec 17, 2025

**DOI:** <https://doi.org/10.62177/jaet.v2i4.907>

## 1.Introduction

In recent years, the continuous improvement in people's living standards and quality of life has been accompanied by increasingly severe environmental pollution, particularly water pollution.<sup>[1-2]</sup> For China, rapid population growth and the expansion of industrial and agricultural production have inevitably led to the generation of large quantities of various wastes during industrial processes. Concurrently, water pollution incidents frequently occur due to inadequate supervision, improper treatment measures, and weak public awareness of environmental protection. Industrial wastewater, waste residues, and exhaust gases have become major sources of water pollution. Although significant efforts have been made in water pollution control, the ongoing deterioration of the water environment has not yet been effectively curbed. Traditional chemical methods for treating water pollution are now on the verge of being phased out. Compared to traditional chemical reduction methods, semiconductor photocatalytic reduction technology offers distinct advantages for treating wastewater pollution.<sup>[3]</sup> Semiconductor photocatalysis is highly favored for dye degradation due to its efficiency and environmental friendliness.<sup>[4-5]</sup> The photocatalytic properties of a semiconductor are closely related to its band structure. When a semiconductor is irradiated with light whose photon energy is greater than or equal to its band gap energy ( $h\nu \geq E_g$ ), the semiconductor absorbs the light energy, exciting electrons from the valence band to the conduction band and leaving behind an equal number of positively charged holes in the valence band.<sup>[6-7]</sup> The photogenerated electrons ( $e^-$ ) possess strong reducing power, while the photogenerated holes ( $h^+$ ) possess strong oxidizing power, together forming a highly active redox system. The specific photocatalytic principle is illustrated in Figure 1.  $\text{TiO}_2$  is, to date, the most extensively studied and widely used semiconductor photocatalyst, exhibiting catalytic activity for numerous photochemical reactions, and is inexpensive and non-toxic. However, its relatively

large band gap energy (3.2 eV) means it can only be excited by ultraviolet light with  $\lambda \leq 387.5$  nm, resulting in low utilization efficiency of solar energy, fast recombination of photogenerated electrons and holes, and low quantum efficiency. To make better use of solar energy, researchers worldwide have conducted extensive work in recent years on modifying TiO<sub>2</sub> and developing visible-light-responsive photocatalytic materials, achieving considerable progress.<sup>[8]</sup> Examples include non-metal element doping (e.g., S-TiO<sub>2</sub>,<sup>[9]</sup> N-TiO<sub>2</sub>), semiconductor composites (e.g., Cu<sub>2</sub>O/TiO<sub>2</sub>, Bi<sub>2</sub>O<sub>3</sub>/TiO<sub>2</sub>,<sup>[10]</sup> ZnMn<sub>2</sub>O<sub>4</sub>/TiO<sub>2</sub><sup>[11]</sup>), and new visible-light catalysts like SnIn<sub>4</sub>S<sub>8</sub>.<sup>[12]</sup> Among these, compositing TiO<sub>2</sub> with narrow-bandgap semiconductor materials has been a persistent research focus. SnS<sub>2</sub>, an important semiconductor from the IV-VI group, has a band gap energy of 2.2 eV - 2.35 eV at room temperature.<sup>[13-14]</sup> While numerous studies based on existing literature have reported on the optical and electrical properties of SnS<sub>2</sub>, there are relatively few publications concerning its photocatalytic properties. He, H.Y. et al. investigated the photocatalytic degradation of methyl orange in water by SnS<sub>2</sub>/TiO<sub>2</sub> composites under natural sunlight, finding that the photocatalytic activity of the mixture increased with higher SnS<sub>2</sub> content.<sup>[15]</sup> Yang, C. et al. studied the photocatalytic performance of SnS<sub>2</sub>/TiO<sub>2</sub> composites prepared via an in-situ chemical method, using methyl orange in water as the degradation target under visible and UV light.<sup>[16]</sup> The results showed that composites prepared by the in-situ method had higher photocatalytic activity than single components or physically mixed composites of the same composition, along with better photocatalytic stability.<sup>[17]</sup> Furthermore, Zhang, Y.C. et al. obtained SnS<sub>2</sub> nanosheets using a solid-liquid phase reaction from a mixture of SnCl<sub>2</sub>·2H<sub>2</sub>O and excess sulfur powder, and synthesized SnS<sub>2</sub> nanopowder via a simple hydrothermal method using SnCl<sub>4</sub>·5H<sub>2</sub>O and thioacetamide as raw materials.<sup>[18-19]</sup> They further prepared composite SnS<sub>2</sub>/SnO nanopowder using an in-situ oxidation method. Evaluating the photocatalytic activity for methyl orange degradation under visible light or sunlight, they found that the synthesized SnS<sub>2</sub> or SnS<sub>2</sub>/SnO composites exhibited high visible-light photocatalytic activity and stability. Based on the above research, this paper prepared SnS<sub>2</sub>/TiO<sub>2</sub> composites via a one-step hydrothermal method.<sup>[20]</sup> The morphology of the synthesized SnS<sub>2</sub>/TiO<sub>2</sub> composites was observed using scanning electron microscopy, and their photocatalytic performance was investigated by degrading methylene blue solution under xenon lamp irradiation.

## 2. Material Preparation

### 2.1 Preparation of TiO<sub>2</sub> Powder

First, prepare two 50 mL beakers. Precisely measure 900  $\mu$ L of absolute ethanol and tetrabutyl titanate separately using a pipette gun, pour them into one beaker, and mix thoroughly with a magnetic stirrer. Label this as Solution A. Then, measure 4.5 mL of deionized water using a pipette gun and pour it into the other 50 mL beaker. Add 25  $\mu$ L of concentrated nitric acid to the deionized water using the pipette gun, and mix thoroughly with a magnetic stirrer. Label this as Solution B. Gradually add Solution A dropwise into Solution B to obtain the precursor solution of tetrabutyl titanate. Transfer the entire precursor solution into a 25 mL polytetrafluoroethylene-lined stainless steel hydrothermal autoclave. Place the autoclave in an oven set at 160°C for 6 hours. After the reaction is complete and the autoclave has cooled naturally to room temperature, open it. Use a pipette to remove the supernatant from the liner. Wash the precipitate once with deionized water and centrifuge, discarding the supernatant. Then wash the precipitate twice with absolute ethanol, centrifuge, and collect the precipitate. Place the precipitate in an oven set at 60°C for 1 hour to dry. Transfer the dried precipitate into a crucible and place it in a muffle furnace. Set the calcination temperature to 500°C with a heating rate of 5°C/min, and calcine for 2 hours to obtain TiO<sub>2</sub> powder. Grind the powder for later use.

### 2.2 Preparation of SnS<sub>2</sub> Nanomaterials

Weigh 0.3502 g of tin(IV) chloride pentahydrate, 0.3024 g of thioacetamide, and 0.3031 g of citric acid separately using an analytical balance. Transfer all the weighed chemicals into a 100 mL beaker, add 10 mL of absolute ethanol, and stir with a magnetic stirrer until fully dissolved. Transfer the entire solution into a 25 mL polytetrafluoroethylene-lined stainless steel hydrothermal autoclave. Place the autoclave in an oven set at 180°C for 12 hours. After the reaction is complete and the autoclave has cooled naturally to room temperature, open it. Remove the supernatant using a pipette. Wash the precipitate once with deionized water and centrifuge, then discard the supernatant. Repeat the washing process twice with absolute ethanol, centrifuging each time and discarding the supernatant. Collect the precipitate and place it in an oven set at 80°C for 6 hours to

dry. The final product is the prepared SnS<sub>2</sub> material, which should be ground for later use.

### 2.3 Preparation of SnS<sub>2</sub>/TiO<sub>2</sub> Composite Material

Weigh 1.0520 g of tin(IV) chloride pentahydrate and 0.5635 g of thioacetamide using an analytical balance, and sequentially add them into a mixed solution containing 30.0 mL of absolute ethanol and 2.0 mL of glacial acetic acid. Stir continuously for 10 minutes to form a transparent solution. Add 256  $\mu$ L of tetrabutyl titanate solution to the above mixture using a pipette gun and mix thoroughly with a magnetic stirrer. Transfer the resulting solution into a 50 mL polytetrafluoroethylene-lined stainless steel hydrothermal autoclave. Place the autoclave in an oven set at 160°C for 12 hours. After the reaction is complete and the autoclave has cooled to room temperature, open it. Remove the supernatant using a pipette. Wash the precipitate once with deionized water, centrifuge, and discard the supernatant. Then wash the precipitate twice with absolute ethanol, centrifuge, and collect the precipitate. Place the precipitate in an oven set at 60°C and dry for 5 hours to obtain the SnS<sub>2</sub>/TiO<sub>2</sub> composite material. Grind the product for later use.

### 3. Surface Morphology Analysis

Figure1: SEM image of TiO<sub>2</sub>

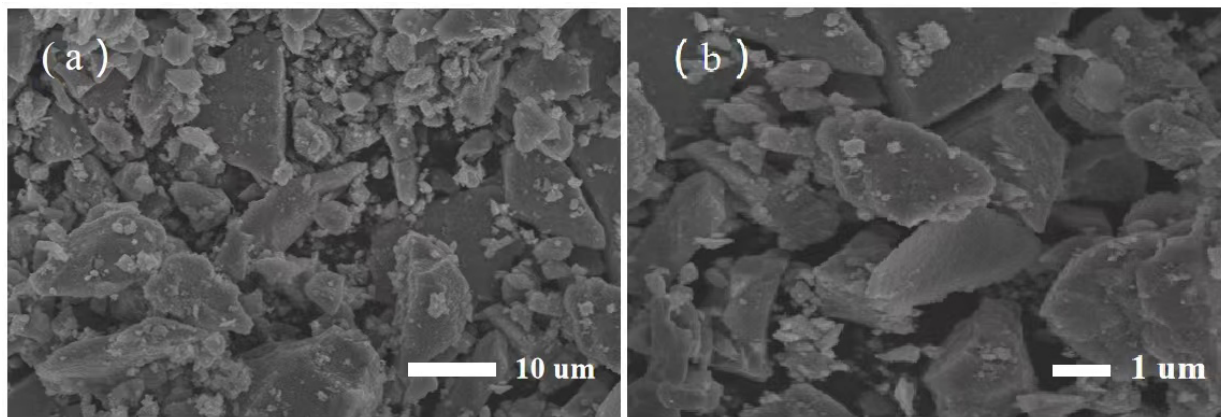


Figure 1 shows scanning electron microscopy (SEM) images of TiO<sub>2</sub> at different magnifications. The images reveal that the cross-sectional diameter of the material is approximately 10  $\mu$ m. Additionally, TiO<sub>2</sub> exhibits an irregular morphology, with two-dimensional TiO<sub>2</sub> particles aggregating to form plate-like structures, and the grain size is relatively large.

Infrastructure Construction Cost for Reverse Logistics: This refers to the expenses required for building the fundamental facilities necessary for the normal operation of the reverse logistics network. These facilities include collection points, transfer stations, processing centers, waste incineration plants, and landfills.

Figure2: SEM images of SnS<sub>2</sub>

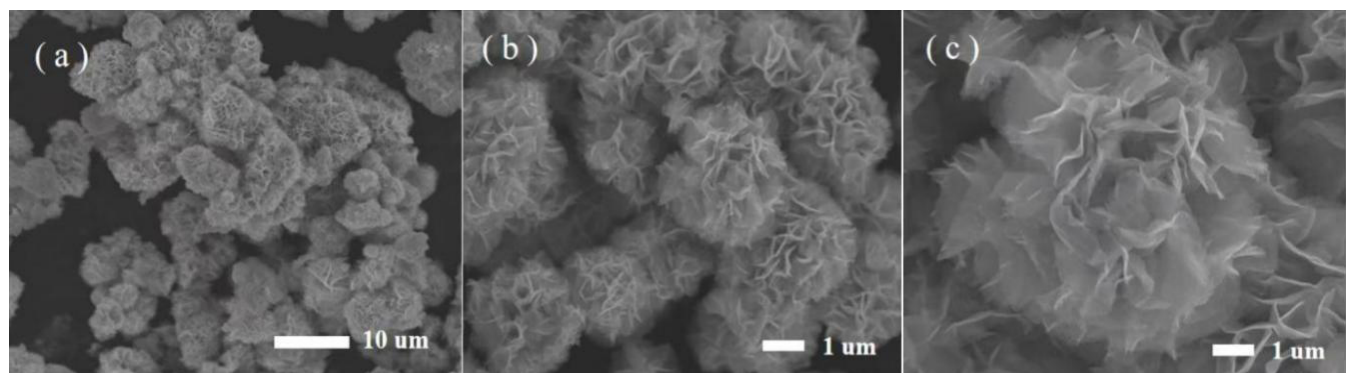


Figure 2 displays scanning electron microscopy (SEM) images of SnS<sub>2</sub> at different magnifications. It can be clearly observed that SnS<sub>2</sub> exhibits a flower-like structure composed of stacked two-dimensional nanosheets. These flower-like structures are well-dispersed, with a diameter of approximately 3  $\mu$ m. Additionally, the rough surface texture and distinct morphological features of the material are clearly visible.

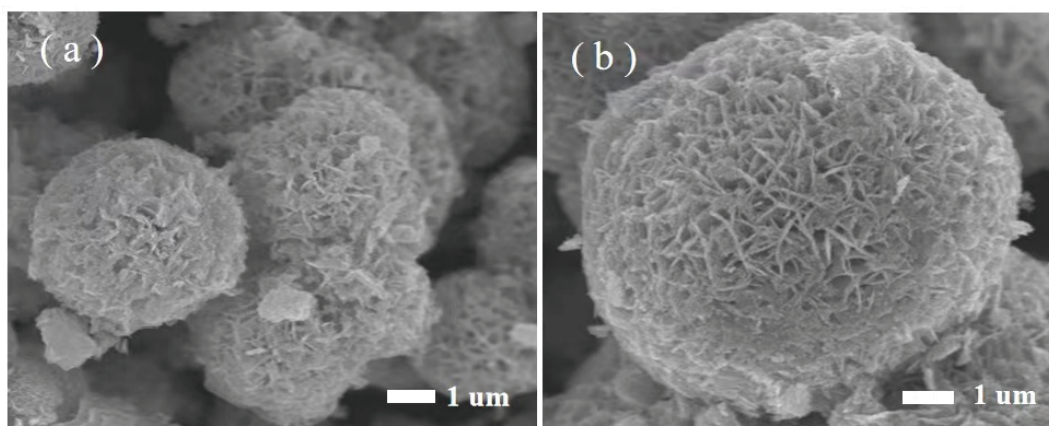
Figure3: SEM images of  $\text{SnS}_2/\text{TiO}_2$  composite material

Figure 3 shows the scanning electron microscopy (SEM) images of the composite material. The images reveal that the  $\text{SnS}_2/\text{TiO}_2$  composite, prepared via the hydrothermal method, exhibits a morphology intermediate between the two individual components, with a diameter of approximately 4–5  $\mu\text{m}$ .  $\text{TiO}_2$  is uniformly dispersed within  $\text{SnS}_2$ , making it difficult to distinguish individual particles. In the composite material,  $\text{SnS}_2$  is encapsulated by  $\text{TiO}_2$ , which mitigates the morphological defects of both individual components. While the overall morphology of the composite remains largely consistent with that of  $\text{SnS}_2$ , noticeable aggregation occurs in the composite, contrasting with the dispersed flower-like structure of pure  $\text{SnS}_2$ .

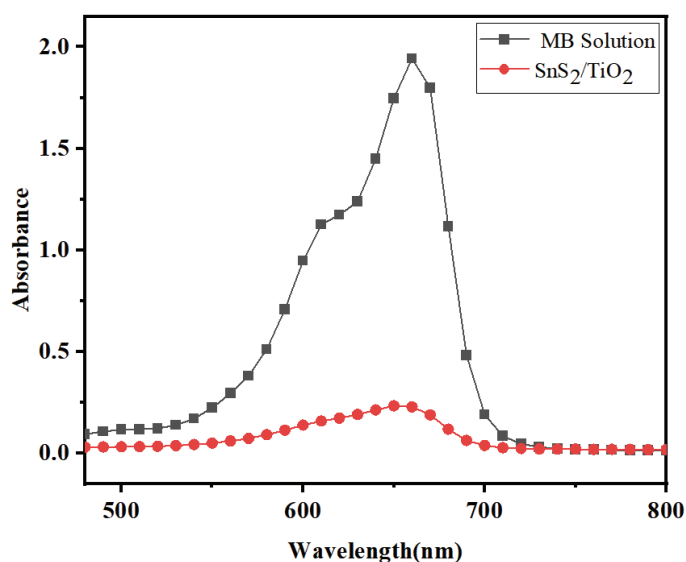
## 4. Study of Photocatalytic Performance

First, a methylene blue solution with a concentration of 20 mg/L was prepared: 10 mg of methylene blue was weighed using an analytical balance, placed in a beaker, dissolved with an appropriate amount of deionized water, and then transferred to a 500 mL volumetric flask. Deionized water was added to the flask to bring the volume to the 500 mL mark, and the solution was mixed thoroughly for later use. Subsequently, 10 mg each of  $\text{TiO}_2$ ,  $\text{SnS}_2$ , and  $\text{SnS}_2/\text{TiO}_2$  composite material, ground using an agate mortar, were weighed and set aside.

### 4.1 Determination of Maximum Absorption Wavelength

Two clean Petri dishes were prepared, and 20 mL of the prepared 20 mg/L methylene blue solution was measured into each using a graduated cylinder. To one of the dishes, 10 mg of the  $\text{SnS}_2/\text{TiO}_2$  composite material was added, while the other dish was used as a control. Both dishes were placed 20 cm below the light outlet of a xenon lamp and irradiated. Samples were taken every 5 minutes. For the dish containing the composite material, centrifugation was performed prior to sampling. The absorbance of each sample was measured using a UV-Vis spectrophotometer, and the data were recorded and plotted.

Figure4: Absorbance at Different Wavelengths

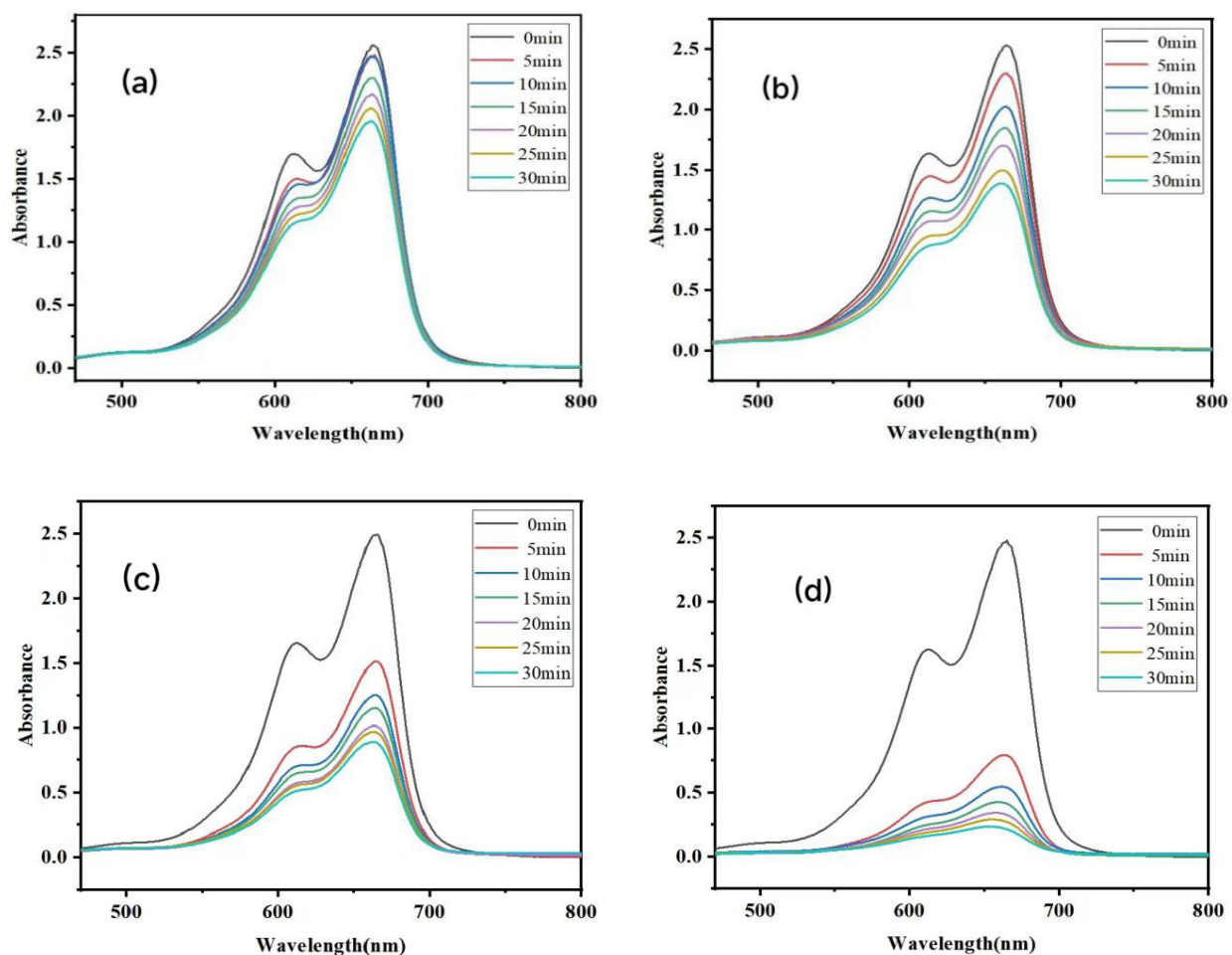


As shown in Figure 4, the maximum absorption wavelength of the 20 mg/L methylene blue solution is 664 nm. Over time, the absorbance of the methylene blue solution decreases significantly due to the addition of the  $\text{SnS}_2/\text{TiO}_2$  composite material. This is reflected in the graph by the red curve consistently lying below the black curve, with the most pronounced reduction in absorbance observed at the maximum absorption wavelength.

#### 4.2 Study on the Photocatalytic Performance of $\text{TiO}_2$ , $\text{SnS}_2$ , and $\text{SnS}_2/\text{TiO}_2$

A baseline was first established by scanning an appropriate amount of deionized water in a cuvette using a UV-Vis spectrophotometer. Another cuvette was filled with a prepared 20 mg/L methylene blue solution and scanned as Sample 1. Then, 10 mg of the pre-weighed  $\text{TiO}_2$  powder was dissolved in a Petri dish containing 20 mL of the 20 mg/L methylene blue dye solution. The Petri dish was placed 20 cm below the light outlet of a xenon lamp for irradiation. Every 5 minutes, a 3 mL sample was taken, centrifuged, and the supernatant was transferred to a cuvette to measure its absorbance using the UV-Vis spectrophotometer. This experimental procedure was repeated six times, and changes in the color of the methylene blue dye were observed. After completing the six sets of experiments with  $\text{TiO}_2$ , the Petri dish and cuvette were washed with deionized water and absolute ethanol, then dried. The  $\text{TiO}_2$  powder in the Petri dish was replaced with the pre-ground  $\text{SnS}_2$  powder and  $\text{SnS}_2/\text{TiO}_2$  powder, respectively, and the above experimental steps were repeated for each material. Changes in the color of the methylene blue dye were observed accordingly.

Figure5: Absorbance of Methylene Blue Solution Degraded by Different Samples (a) Pure methylene blue solution; (b)  $\text{TiO}_2$ ; (c)  $\text{SnS}_2$ ; (d)  $\text{SnS}_2/\text{TiO}_2$  composite material



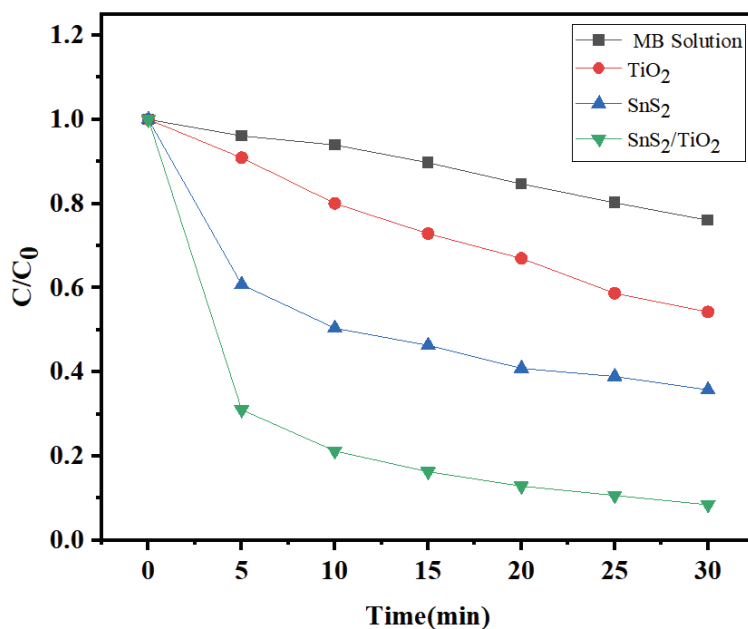
In Figure 5, the horizontal axis represents the wavelength, and the vertical axis represents the absorbance of each material. From Figure 5(a), it can be observed that the pure methylene blue solution without any catalyst exhibits the highest absorbance at all time intervals, with negligible changes in absorbance over time. Figures 5(b), (c), and (d) show that the absorbance of the solution decreases significantly after the addition of different catalysts compared to the pure methylene blue

solution. The methylene blue solution with SnS<sub>2</sub> added shows the second most significant reduction in absorbance after the SnS<sub>2</sub>/TiO<sub>2</sub> composite. Furthermore, Figure (b) indicates that TiO<sub>2</sub> has a less pronounced effect on reducing the absorbance of the methylene blue solution compared to SnS<sub>2</sub> and the composite material. Additionally, it is observed that as time increases, the reduction in absorbance of the methylene blue solution by TiO<sub>2</sub>, SnS<sub>2</sub>, and the SnS<sub>2</sub>/TiO<sub>2</sub> composite becomes more pronounced. Comparing the four graphs, it is evident that the SnS<sub>2</sub>/TiO<sub>2</sub> composite material (Figure d) achieves the most effective reduction in absorbance of the methylene blue solution.

### 4.3 Study of Degradation Performance

The absorbance data of the pure methylene blue solution and the methylene blue solutions with added TiO<sub>2</sub>, SnS<sub>2</sub>, and SnS<sub>2</sub>/TiO<sub>2</sub> composite, respectively, were processed to obtain the following figure.

Figure6: Degradation curves of methylene blue solutions with different materials added



From Figure 6, it can be observed that under the same degradation time, the pure methylene blue solution without any photocatalyst shows negligible degradation, with a degradation rate of only 24% after 30 minutes. In contrast, the degradation efficiency of methylene blue solutions with different photocatalysts added is significantly improved. The red dotted curve exhibits a degradation rate of approximately 46% after 30 minutes, which is 1.9 times higher than that of the pure methylene blue solution. The blue triangular curve achieves a degradation rate of about 65% after 30 minutes, representing a 2.7-fold increase compared to the pure methylene blue solution. The green curve demonstrates the best degradation performance, reaching a degradation rate of 92% after 30 minutes, which is 3.8 times higher than that of the pure methylene blue solution. Therefore, it can be concluded that the SnS<sub>2</sub>/TiO<sub>2</sub> composite material exhibits the most effective degradation of methylene blue solution, with the highest dye degradation rate.

### Conclusion

The main focus of this study was the preparation of SnS<sub>2</sub>/TiO<sub>2</sub> composite material via a one-step hydrothermal method, followed by morphological characterization, absorbance measurement, and investigation of its photocatalytic performance. The following key findings were obtained:

1. The maximum absorption wavelength of the SnS<sub>2</sub>/TiO<sub>2</sub> composite material for methylene blue solution was determined to be 664 nm.
2. The degradation efficiency of pure methylene blue solution without any photocatalyst was negligible, while the addition of different photocatalysts significantly enhanced the degradation performance.
3. Under controlled conditions with identical degradation time, light intensity, and material dosage, the SnS<sub>2</sub>/TiO<sub>2</sub> composite material achieved the highest degradation efficiency of 92% for methylene blue solution. This indicates that the photocatalytic

performance of the composite is superior to that of its individual components.

4. The composite material prepared by the hydrothermal method demonstrates promising potential for applications in the field of photocatalysis, fulfilling the objectives and expectations of this study.

## Funding

No

## Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Reference

- [1] Guo, C. B. (2021). Two-dimensional bismuth selenide anchored semiconductor catalyst and its photocatalytic performance [Doctoral dissertation]. Xiangtan University. <https://doi.org/10.27426/d.cnki.gxtd.2021.000555>
- [2] Xu, B. X., & Yuan, F. (2011). Where is the water of life (pp. 8–9, 105–117). Beijing Yanshan Press.
- [3] Zhou, X. L. (2011). Chemistry and daily life (pp. 18–26). China Electric Power Press.
- [4] Wang, B., Zhang, X., Zhang, N., et al. (2015). Two-dimensional MoS<sub>2</sub> nanosheet-coated Bi<sub>2</sub>S<sub>3</sub> discoids: Synthesis, formation mechanism, and photocatalytic application. *Langmuir*, 31(14), 4314–4322. <https://doi.org/10.1021/acs.langmuir.5b00544>
- [5] He, H. Y., Huang, J. F., Cao, L. Y., & Wu, J. P. (2007). Photocatalytic activity of mixture of SnS<sub>2</sub> and TiO<sub>2</sub> powders in destruction of methyl orange in water. *Optoelectronic Advanced Materials*, 12, 3781–3784. <https://doi.org/10.1166/oam.2007.1452>
- [6] Humayun, M., Raziq, F., Khan, A., et al. (2018). Modification strategies of TiO<sub>2</sub> for potential applications in photocatalysis: A critical review. *Green Chemistry Letters and Reviews*, 11(2), 86–102. <https://doi.org/10.1080/17518253.2017.1366786>
- [7] Zhang, Y. C., Du, Z. N., Li, S. Y., & Zhang, M. (2010). Novel synthesis and high visible light photocatalytic activity of SnS<sub>2</sub> nanoflakes from SnCl<sub>2</sub>·2H<sub>2</sub>O and S powders. *Applied Catalysis B: Environmental*, 95, 153–159. <https://doi.org/10.1016/j.apcatb.2009.11.015>
- [8] Akpan, U. G., & Hameed, B. H. (2009). Parameters affecting the photocatalytic degradation of dyes using TiO<sub>2</sub>-based photocatalysts: A review. *Journal of Hazardous Materials*, 170, 520–529. <https://doi.org/10.1016/j.jhazmat.2009.06.002>
- [9] Zheng, L. X., et al. (2020). Photo/electro chemical applications of metal sulfide/TiO<sub>2</sub> heterostructures. *Advanced Energy Materials*, 10(1), 1902355. <https://doi.org/10.1002/aenm.201902355>
- [10] Kumair, et al. (2016). Efficiency enhancement in plasmonic dye-sensitized solar cells with TiO<sub>2</sub> photoanodes incorporating gold and silver nanoparticles. *Journal of Applied Electrochemistry*, 46(1), 47–58. <https://doi.org/10.1007/s10800-015-0944-8>
- [11] Wang, J., & Zhuan, R. (2020). Degradation of antibiotics by advanced oxidation processes: An overview. *Science of the Total Environment*, 701, 135023. <https://doi.org/10.1016/j.scitotenv.2019.135023>
- [12] Wei, Z., Liu, J., & Shangguan, W. (2020). A review on photocatalysis in antibiotic wastewater: Pollutant degradation and hydrogen production. *Chinese Journal of Catalysis*, 41(10), 1440–1455. [https://doi.org/10.1016/S1872-2067\(20\)63698-8](https://doi.org/10.1016/S1872-2067(20)63698-8)
- [13] Kovalakova, P., Cizmas, L., McDonald, T. J., et al. (2020). Occurrence and toxicity of antibiotics in the aquatic environment: A review. *Chemosphere*, 251, 126351. <https://doi.org/10.1016/j.chemosphere.2020.126351>
- [14] Martinez, J. L. (2009). Environmental pollution by antibiotics and by antibiotic resistance determinants. *Environmental Pollution*, 157(11), 2893–2902. <https://doi.org/10.1016/j.envpol.2009.07.031>
- [15] Yang, C., Wang, W., Shan, Z., & Huang, F. (2009). Preparation and photocatalytic activity of high-efficiency visible-light-responsive photocatalyst SnS<sub>2</sub>/TiO<sub>2</sub>. *Solid State Chemistry*, 182, 807–812. <https://doi.org/10.1016/j.solid-statesciences.2009.02.016>
- [16] Ismail, A. A., & Bahnemann, D. W. (2014). Photochemical splitting of water for hydrogen production by photocatalysis: A review. *Solar Energy Materials and Solar Cells*, 128, 85–101. <https://doi.org/10.1016/j.solmat.2014.03.034>

- [17] Wang, J., Wang, J., Wu, X., et al. (2017). Pt-TiO<sub>2</sub> microspheres with exposed {001} facets for degradation of formaldehyde in air: Formation mechanism and enhanced visible light photocatalytic activity. *Materials Research Bulletin*, 96, 262–269. <https://doi.org/10.1016/j.materresbull.2017.08.031>
- [18] Zhang, Y. C., Du, Z. N., Li, K. W., et al. (2011). Size-controlled hydrothermal synthesis of SnS<sub>2</sub> nanoparticles with high performance in visible light-driven photocatalytic degradation of aqueous methyl orange. *Separation and Purification Technology*, 81(1), 101–107. <https://doi.org/10.1016/j.seppur.2011.06.018>
- [19] Reli, M., Huo, P., Sihor, M., et al. (2016). Novel TiO<sub>2</sub>/C<sub>3</sub>N<sub>4</sub> photocatalysts for photocatalytic reduction of CO<sub>2</sub> and for photocatalytic decomposition of N<sub>2</sub>O. *Journal of Physical Chemistry A*, 120(43), 8564–8573. <https://doi.org/10.1021/acs.jpca.6b08444>
- [20] Anucha, C. B., Altin, I., Bacaksiz, E., et al. (2022). Titanium dioxide (TiO<sub>2</sub>)-based photocatalyst materials activity enhancement for contaminants of emerging concern (CECs) degradation: In the light of modification strategies. *Chemical Engineering Journal Advances*, 10, 100262. <https://doi.org/10.1016/j.ceja.2022.100262>

# Vocabulary Semantic Similarity Calculation in Natural Language Processing

Huixiang Xiao<sup>1</sup>, Kaige Zheng<sup>2</sup>, Xiangyu Li<sup>3\*</sup>

1.Chongqing University of Technology, Chongqing, 402160, China

2.Graduate School of Techno Design, Kookmin University, Seoul, 02707, Korea

3.Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, 200240, China

\*Corresponding author: Xiangyu Li, [xiangyuli@sjtu.edu.cn](mailto:xiangyuli@sjtu.edu.cn)

**Copyright:** 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

**Abstract:** Natural language processing (NLP) is a critical research direction in artificial intelligence, where the calculation of vocabulary semantic similarity is the foundation and core work. However, existing calculation methods are faced with problems, e.g., the inability to extract important semantic information. Failure to address this issue can compromise the accuracy of semantic similarity measures in NLP applications. To this end, in this paper, a vocabulary semantic similarity calculation model based on word vectors and convolutional neural networks (CNNs) was proposed. The word vector model was improved using long short-term memory (LSTM) networks, and important semantics were extracted using convolutional layers and ensured semantic order through bidirectional Gated Recurrent Unit. The structure of the Siamese neural network was used to ensure consistency in text encoding. The experimental findings have shown that the proposed model has the highest F1 value in different datasets. In the original Chinese natural language inference (OCNLI) dataset, the Pearson correlation coefficient of the model was 0.021 and 0.018 higher than that of the LSTM network and CNNs, respectively. The accuracy of the similarity calculation in the two datasets was 92.4% and 96.5%, respectively. According to these results, the semantic similarity prediction value of the proposed model can be closer to the true value, and the prediction performance of the model is more excellent.

**Keywords:** Semantic Similarity; Convolutional Neural Network; Gated Recurrent Unit; Natural Language Processing; Word Vector

**Published:** Dec 19, 2025

**DOI:** <https://doi.org/10.62177/jaet.v2i4.946>

## 1.Introduction

Natural language processing (NLP) is one of the most important and challenging applications in artificial intelligence, where the progress is catalysed by foundation models. There are lots of opportunities as well as challenges for those valuable information<sup>[1]</sup>. However, with so much information available, it's hard to avoid personal information overload! The proliferation of information makes it difficult for people to filter and process a large amount of information, increasing the difficulty of using information<sup>[2]</sup>. At the same time, a large amount of junk information not only hinders the efficiency of decision making of individuals, companies, and even countries, but may also be exploited and harmful to society, such as online rumors<sup>[3-4]</sup>.

The calculation of vocabulary semantic similarity (VSS) refers to measuring the degree of similarity in meaning between two

word objects, which can transform abstract vocabulary similarity relationships into numerical values that can be processed by computers<sup>[5]</sup>. Natural language processing (NLP) can effectively solve these problems, and VSS calculation is one of the key technologies. Calculating lexical semantic similarity (SS) is widely used in fields such as information retrieval, text classification, and others. After calculating the SS between two words, NLP systems can better understand the content of the text and improve the accuracy and computational efficiency of NLP systems<sup>[6]</sup>.

## 1.1 Related Work

Existing methods for calculating VSS include word vectors, semantic maps, and deep learning-based methods such as BERT, RoBERTa, and XLNet. However, these methods suffer from low accuracy in calculating SS, insufficient interpretability, and lack of consideration of word order. Ismail et al. developed a new alignment word space method to increase the accuracy of SS calculation for vocabulary. This method combined alignment-based and vector space-based similarity calculation methods to represent words as vectors in a semantic network, and generated an aligned word space matrix for the text based on the word space. The experiment showed good performance in word-order processing of input text and vector models. The accuracy of the result reached 0.7212<sup>[7]</sup>. Pan et al. found that language model training incurred high computational and time cost and therefore proposed a vocabulary enhancement method based on elastic boundaries and multi-sample learning. This method embedded the constraint relationships implied by the vocabulary in neural words, classified the vocabulary constraint set, and allowed positive and negative samples to learn from each other. The benchmark accuracy is shown to be improved to 75% by evaluating the lexical similarity of neural word embeddings<sup>[8]</sup>. Dai et al. developed a new semantic association model to enhance their understanding of the intrinsic connections between English vocabulary. A tree kernel function was introduced to extract relationships between words, combining dictionary-based and corpus-based methods for calculating VSS to create a dataset of related vocabulary. The experimental results indicated that the more common-sense phenomena there were, the higher the correlation value. This method had the highest classification accuracy for different vocabulary relationships, reaching 85.57%<sup>[9]</sup>. Ahmad F et al. proposed a hybrid method for calculating SS between sentences. This method considered semantic information including vocabulary databases, word embeddings, and corpus statistics, as well as implicit word order information, to model human common-sense knowledge. Experiments indicated that the highest correlation value could be obtained in both word and sentence similarity, and compared with methods that only use word vectors or based on WordNet, it could improve by up to 32%, with a Pearson correlation coefficient of 0.8953<sup>[10]</sup>. Chauhan S et al. proposed an evaluation method combining semantic and syntactic similarity by integrating TF-IDF and word embeddings. Experiments indicated that this method could effectively improve the recognition accuracy of machine translation for natural language<sup>[11]</sup>. Osth et al. introduced a fused letter position model that combines absolute and relative letter position representations. The experiment demonstrated that the model exhibited moderate capacity to capture the variability of individual words in the false alarm rate<sup>[12]</sup>. Zhang et al. developed a novel approach using deep learning for part-of-speech-based encoding and context-based decoding. Their study demonstrated that this method could efficaciously reduce the amount of data transmitted and enhance the semantic accuracy between transmitted and recovered messages<sup>[13]</sup>.

## 1.2 Contributions and Paper Organization

Existing studies have investigated VSS calculation and word vector improvement from multiple perspectives and have achieved meaningful progress. However, existing approaches still suffer from issues such as limited accuracy in semantic similarity computation and insufficient consideration of word order. To address these challenges, this paper proposes a novel VSS calculation model that integrates word vectors with convolutional neural networks (CNNs) and long short-term memory (LSTM) networks. Specifically, the LSTM network is used to enhance the quality of word vector representations, while a combination of CNNs and bidirectional gated recurrent units (Bi-GRUs) is employed to effectively capture sequential dependencies and key contextual semantics before and after vocabulary. This design aims to improve both the accuracy and computational efficiency of VSS calculation.

The main contributions of this paper are summarized as follows:

A novel VSS computation model is developed by integrating word vectors, CNNs, and GRUs to effectively address the low-accuracy limitations and neglect of word order in existing methods.

An LSTM-based enhancement of word vector representations is introduced, which enables better extraction of sequential and contextual semantic information.

Extensive experiments conducted on benchmark datasets (LCQMC and OCNLI) demonstrate that the proposed model achieves higher accuracy (92.4% and 96.5%) and greater stability compared with baseline models such as LSTM and CNN.

The rest of the paper is organized as follows: Section 2 discusses the proposed VSS calculation model, specifically highlighting the word-vector-enhanced LSTM network approach with a CNN-BiGRU model, to obtain semantic usage and context information. Section 3 verifies the efficiency and stability of the proposed model by testing with LCQMC, OCNLI datasets.

The results are compared to the baseline models including CNNs, GRUs, and LSTMs. Finally, Section 4 summarizes the findings of this paper and provides ample scope for future research.

## 2. Equations and Mathematical Expressions

### 2.1 Word Vector-based VSS calculation

Vocabulary is the basic unit of semantic expression, including all words and phrases in the language. Calculating the VSS can improve the performance of information retrieval systems and optimize the quality of machine translation<sup>[14]</sup>. This paper is based on the Skip-gram model proposed by Mikolov et al. which predicts contextual words from a given central vocabulary. The structure of the model consists of an input layer, a hidden layer, and an output layer. The model is trained using the Skip-gram approach to obtain word vector representations. During training, a window is slid on the text data and each sliding generates a training sample, which includes a central word and a series of contextual words<sup>[15]</sup>. The goal of the model is to maximize the sum of conditional probabilities to predict these contextual words given a central word. The maximization function is calculated as

$$Q = -\frac{1}{n} \sum_{n=1}^n \sum_{-c \leq j \leq c} \log P(w_{n+j} | w_n) \quad (1)$$

where  $n$  means the number of words,  $C$  represents the size of the sliding window,  $P(w_{n+j} | w_n)$  represents the conditional probability,  $w_n$  is the center word, and  $w_{n+j}$  is the adjacent word. The conditional probability calculation is expressed as<sup>[16]</sup>

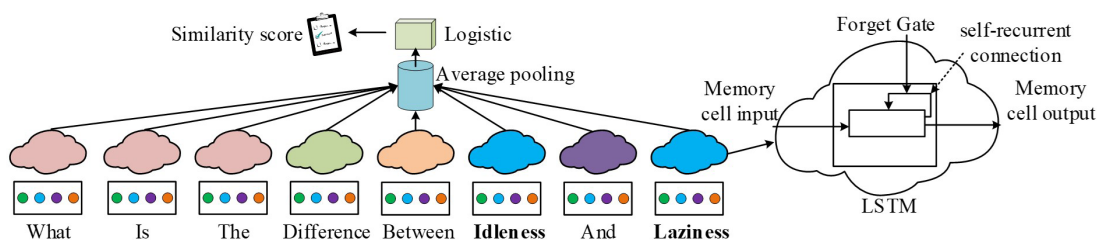
$$P(w_{n+i} | w_n) = \frac{\exp(w_{n+i}^{(2)} \cdot w_n^{(1)})}{\sum_{k=1}^V \exp(w_k^{(2)} \cdot w_n^{(1)})} \quad (2)$$

where  $V$  represents the total number of adjacent words,  $k$  represents adjacent words,  $w_{n+j}^{(2)}$  represents the row vector of the adjacent word composition matrix, and  $w_n^{(1)}$  represents the row vector of the word composition matrix.

To obtain the SS of different vocabulary, the word vector space of the preceding and following vocabulary can be calculated, and the distance between the spatial vectors can be used to calculate it. This paper used the large-scale Chinese question matching corpus (LCQMC) dataset from Harbin Institute of Technology and the original Chinese natural language inference (OCNLI) dataset from Alibaba to obtain the relevant corpus. Stop words are removed from the data by removing words that have no actual semantic or informational value to improve the efficiency of text analysis. Research segments continuous text into independent words according to rules. The remaining data preprocessing also includes XML tag filtering, punctuation removal, and number filtering. Based on the Skip-gram model, this study uses the LSTM network to extract the sequence information of the vocabulary, feedback its contextual vocabulary, and increase the accuracy of the SS calculation.

During training, the complete sentence of the vocabulary to be calculated is input into the model and the specific structure of the model is depicted in Figure1.

Figure1: LSTMs learning word pair co-occurrence utterance computation process



The input is the same sentence or two different sentences with similar words appearing, and the output is the final SS score of the vocabulary. The self-feedback connection refers to the process in a neural network where the output of a neuron is fed back to the same neuron as input, forming a closed loop. This connection method enables neurons to process time-series data and transmit information between multiple time steps. The processed data are input into the average pooling layer, which achieves feature dimension reduction by taking the average of all values within the local receptive domain, reducing the size of the feature map and decreasing computational complexity. Enter the logistic regression layer again, convert the input data into probability, and finally obtain the similarity score of vocabulary. The activation value of the input gate of the storage unit in the LSTM model is calculated as [17]

$$in_t = \sigma(W(h_{t-1}, x_t) + b_i) \quad (3)$$

where  $in_t$  represents the activation value of the input gate,  $\sigma$  means the Sigmoid activation function,  $W$  means the weight matrix of the input gate,  $h_{t-1}$  represents the hidden state of the previous time step,  $x_t$  represents the input at time  $t$ , and  $b_i$  represents the bias term of the inputting gate. The candidate value for the state of memory neurons is

$$C_t^{\%} = \tanh(W_C(h_{t-1}, x_t) + b_C) \quad (4)$$

where the hyperbolic tangent function has a range of  $[-1, 1]$ ,  $W_C$  is the weight matrix of the candidate memory neuron state, and  $b_C$  is the bias term of the candidate memory neuron state. The activation value of the forgetting gate is calculated as

$$f_t = \sigma(W_f(h_{t-1}, x_t) + b_f) \quad (5)$$

where  $f_t$ ,  $W_f$ , and  $b_f$  represent the activation value, the weight matrix, and the bias term of the forgetting gate, respectively. The memory neuron state at the current time step can be calculated using (3), (4) and (5) as

$$C_t = in_t \cdot C_t^{\%} + f_t \cdot C_{t-1} \quad (6)$$

where  $C_t$  represents the memory neuron state at the current time step, and  $C_{t-1}$  represents the memory neuron state at the previous time step. Based on the memory neuron state at the current time step, the activation value and output of the output gate are calculated as [18]

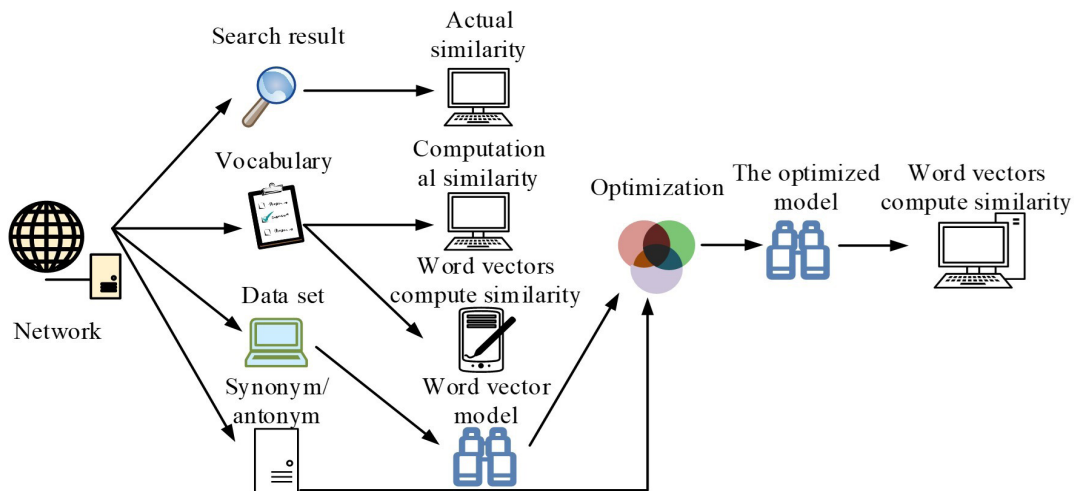
$$\begin{cases} out_t = \sigma(W_{out}(h_{t-1}, x_t) + b_{out}), \\ h_t = out_t \cdot \tanh(C_t), \end{cases} \quad (7)$$

where  $out_t$ ,  $W_{out}$ , and  $b_{out}$  represent the activation value, the weight matrix, and the bias term of the output gate, respectively.  $h_t$  represents the final output of the registration code. After passing the final output through the average pooling layer and logistic regression layer, the SS score of the vocabulary is calculated as

$$S = \sum_{i=1}^{10} i \cdot P_i \quad (8)$$

where  $S$  means the SS score of vocabulary,  $i$  represents the category label, and  $P_i$  means the probability that the category label is  $i$ . The SS calculation process to improve the Skip-gram model is shown in Figure2.

Figure2: Flow of semantic similarity calculation with improved Skip-gram model



Sentences with similar meanings are crawled from the Internet to calculate the similarity of Chinese or English words, train the improved model, and judge the model training results based on the relevant results. After training is complete, synonyms or antonyms are entered into the model and random gradient descent is utilized to optimize the word vector space. The optimized word vector model is utilized to calculate new word vectors, and finally the SS between the two words is calculated based on the word vectors. The examples of vocabulary relation types extracted from semantic parsing are presented in Table1.

Table1. Vocabulary relation types extracted from semantic parsing of event-related entities

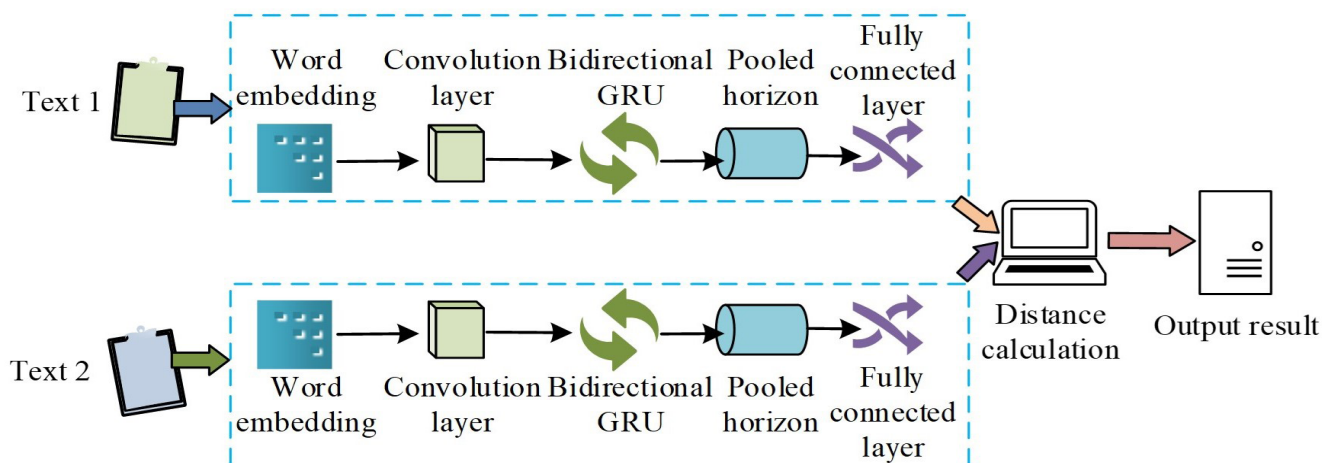
Label	Type	Relation	Target
Soccer Player	SoccerPlayerType	patientOf	Event red card
Referee	SoccerRefereeType	agentOf	Event red card
Event red card	EventType		

## 2.2 CNN and GRU-based Similarity Calculation

As vocabulary may have the characteristics of polysemy and multiple meanings, it is not enough to calculate SS based solely on individual words in text. It is also necessary to connect with the context to extract semantic information from the vocabulary, further increasing the accuracy of the SS calculation<sup>[19]</sup>. However, traditional LSTM can only extract sequence information and cannot extract key semantic information.

To address these issues, a CNN-bidirectional gated recurrent unit (CNN-BiGRU) SS calculation model was proposed. As shown in Figure3, the model contains two identical neural networks. The model contains two identical neural networks, each with identical parameters and weights. After inputting a single text or two identical texts, the model enters the embedding layer to transform the text, extracts text features in the convolutional layer, and the extracted features enter the bidirectional Gated Recurrent Unit layer to learn and extract the position information of vocabulary in the context. The extracted feature vectors are then passed into the pooling layer to reduce dimensionality. Subsequently, they are fed into the fully connected layer for complete vectorization. Finally, the similarity score is calculated by computing the distance between the resulting vocabulary vectors. As the input required for the model is string text, it is necessary to convert the lexical content of the text into the corresponding numerical information before entering the embedding layer for conversion. Data preprocessing operations such as stop word removal, tag filtering, and word segmentation are also required., tag filtering, and word segmentation are also required. The preprocessed data are then normalized to fix the length of a single text to 20, which means that the number of words in the sentence is 20. If the vocabulary exceeds 20, remove the excess vocabulary and if it is less than 20, fill it with a zero vector. After processing, a word vector table is used to map the corresponding text content to dense high-dimensional vectors.

Figure3: CNN-BiGRU vocabulary semantic similarity computation model



Significant semantic information is encoded in high-dimensional dense vectors output by embedding layers in convolutional networks. These vectors are then transformed with the use of convolution kernels. The length of convolutional kernels in the model is the same as that of dense vectors. The kernels move backward along the vector, performing local sections of width 3. The resulting one-dimensional (1D) vectors are compressed semantic features. The computation of the convolutional feature maps is given by<sup>[20]</sup>

$$\text{map} = le - k + 1 \quad (9)$$

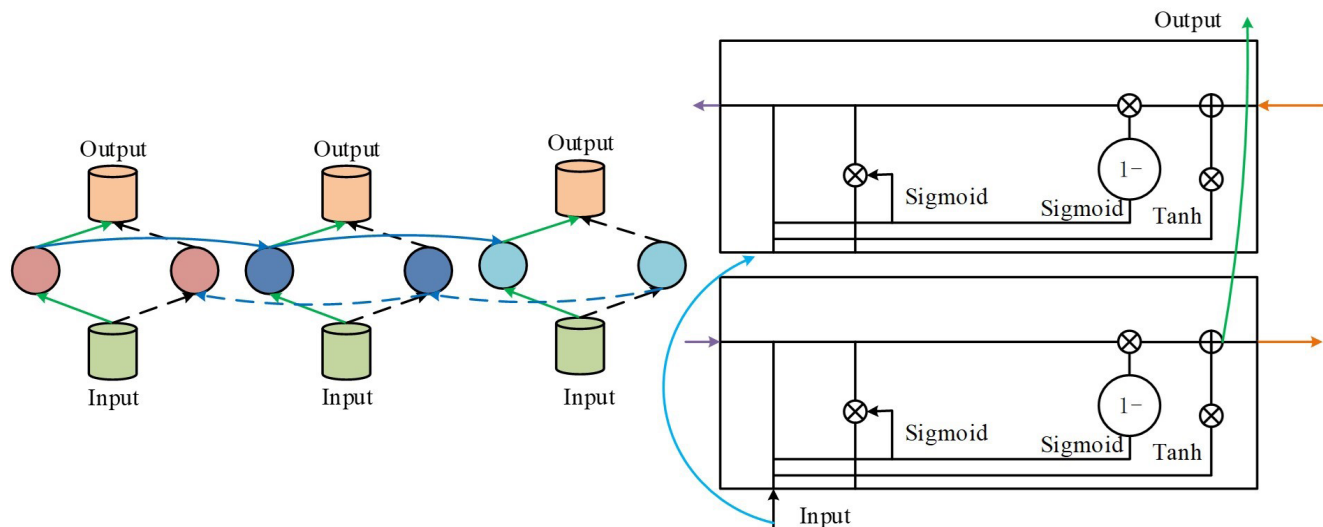
which represents the convolutional feature map,  $le$  represents the dimensionality of the convolutional kernel,  $k$  represents the convolutional kernel, where the stride of each movement of the convolutional kernel is 1. In convolutional layers, models can assign higher weights to relevant semantic information, enhancing the ability to extract important semantic information. Both bidirectional gated recurrent units and LSTM are transformed from recurrent neural networks, but LSTM has a more complex structure and higher computational cost. Therefore, research is being conducted to use bidirectional gated recurrent units instead of LSTM.

The specific structure of BiGRU is shown in Figure4. In the forward direction, the state update of the gated recurrent unit is output sequentially from the previous moment to the next moment, shifting from the previous text to the following text, and the semantics of the vocabulary can be associated with the content of the previous sentence. In the opposite direction, the gated recurrent unit associates the input of the next moment with the content of the previous moment, allowing the model to link the semantics of the vocabulary with the content of the preceding and following words, thereby extracting important semantic information from the sentence. Although the structure of the CNN-BiGRU SS calculation model is a twin neural network, the structures and parameters of the two neural networks are exactly the same, and the input data are different. Therefore, the two types of output data have consistency and can be used for distance calculation to determine their SS. This paper uses the Manhattan distance method to calculate the distance between vocabulary vectors, and the Manhattan distance is written as

$$d = \sum_{i=1}^n |b_i - a_i| \quad (10)$$

where  $n$  represents the dimension of the lexical space,  $b_i$  and  $a_i$  represent the position of the point  $a$  and point  $b$  in the dimensional space  $i$ , respectively. In the similarity calculation layer, backpropagation can be used to determine the difference between the calculated similarity and the actual similarity, and continuously adjust the relevant parameters of the twin neural network to improve its SS calculation ability. When the SS calculation value of two words is less than 1, and the two words in the dataset are similar, the model adjusts the parameters through backpropagation to gradually approach 1 in the calculation result. When two words are different and the actual SS calculation value is not 0, the model back propagates the results and continuously adjusts the weight parameters to gradually approach 0.

Figure4: Specific structure of the bidirectional gating recurrent unit



### 3. Experimental Results

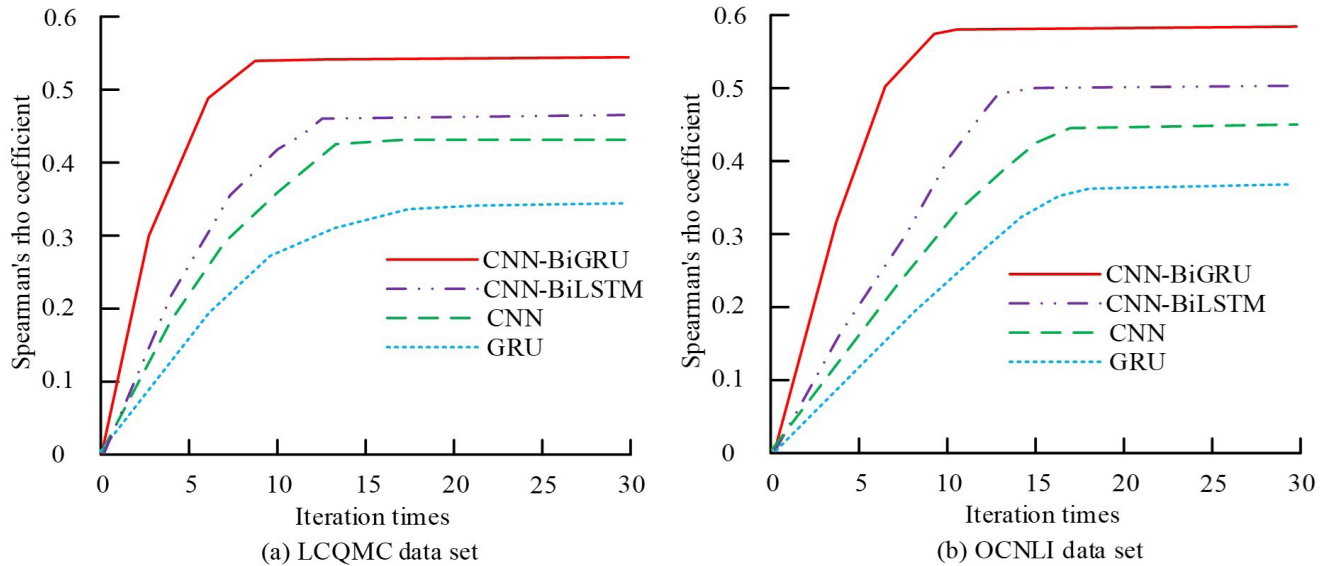
#### 3.1 Effectiveness and Stability Analysis

The experimental configuration selected for this paper was Windows 10 64 bit operating system, with an Intel i5 12700H CPU and 16GB of memory. The selected datasets for this paper were Baidu's LCQMC dataset and Alibaba's OCNLI dataset. The data sets were broken into training and testing sets in an 8:2 ratio. The comparative models used in this paper included CNN, GRU, LSTM, and CNN-BiLSTM.

The effectiveness comparison of the optimized model is indicated in Figure 5. In Figure 5 (a), the CNN-BiGRU model converged the fastest and reached the convergence state after 10 iterations. The maximum Spearman coefficient was 0.07, 0.11, and 0.19 higher than that of CNN-BiLSTM, CNN, and GRU, respectively. In Figure 5 (b), the model variation curve was basically the same, but the maximum value of the Spearman coefficient increased, and the maximum value of the CNN-BiGRU model increased by 0.05 compared to Figure 5.

The stability analysis of the optimized model is shown in Figure 6. In Figure 6 (a), five sets of samples were randomly selected from the training and testing sets, and the Spearman coefficients remained stable above 0.50. Pearson coefficients varied within the range of [0.453, 0.524], indicating that the improved model had good stability. In Figure 6 (b), the Spearman and Pearson coefficients of each sample group increased, with similar changes in magnitude.

Figure 5: Comparison of the effectiveness of the improved model



The performance comparison of different models is shown in Figure 7. In Figure 7 (a), the maximum Pearson coefficient of the CNN-BiGRU model was 0.69, which was 0.05, 0.13, and 0.11 higher than CNN-BiLSTM, LSTM, and CNN, respectively. In Figure 7 (b), the maximum Pearson coefficients of each model increased, but the convergence position remained unchanged. The maximum values of the CNN-BiGRU model were 0.013, 0.021, and 0.018 higher than those of other models, respectively.

#### 3.2 Comparative Analysis

The results of the VSS calculation for different models are indicated in Table 2. It is shown that the effectiveness of the CNN model is better than that of the LSTM and GRU models in both datasets, and the performance of the model that combined CNN and LSTM was further improved. However, among all models, the CNN-BiGRU model proposed in this paper had the best performance, with recall rates 5.7%, 10.0%, 8.6%, and 1.8% higher than CNN, GRU, LSTM, and CNN-BiLSTM, respectively. In the OCNLI dataset, the performance of the model improved by 9.2% compared to the LCQMC dataset, with the largest increase. The similarity calculation accuracy of the CNN-BiGRU model was the highest in both datasets, at 92.4% and 96.5%, which was 8.6% and 12.2% higher than the lowest LSTM.

Figure6: Stability analysis of the improved model

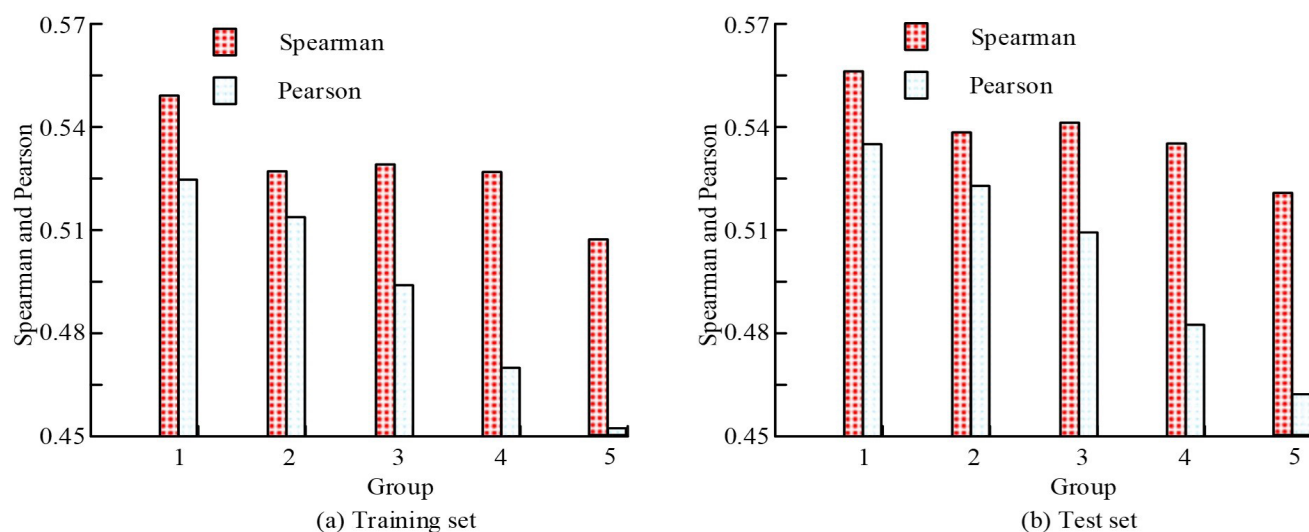


Table2. Performance metrics of different models on LCQMC and OCNLI datasets

Model	Recall/%		Precision/%		F1/%	
	LCQMC	OCNLI	LCQMC	OCNLI	LCQMC	OCNLI
CNN-BiGRU	88.2	97.4	92.4	96.5	90.2	98.4
CNN	82.5	95.3	87.7	91.2	85.8	96.8
GPU	78.2	94.1	85.3	90.5	81.9	95.5
LSTM	79.6	92.8	83.8	84.3	81.6	89.1
CNN-BiLSTM	86.4	96.5	90.9	94.4	88.4	97.2

The comparison of similarity scores for different vocabulary before and after model improvement is indicated in Table3. Pairs like Important-Significant and Abundant-Plentiful are semantically very close to each other, which explains their high scores. Nonetheless, the baseline model (Before improvement) did not fully detect their similarity with scores of 7.2 and 6.8, which are much lower than the human scores of 9.2 and 8.8. The difference hints that the baseline model, that only uses co-occurrences, might not grasp deep semantic equivalence. After enhancement, the proposed model corrected the estimates of 9.5 and 9.0 closer to the human score. The remaining words with high similarity had a lower probability of appearing simultaneously, and the evaluation of high scores was more cautious, resulting in a decrease in score. After improvement, the model significantly reduced the correlation between the probability of scoring and the number of words that appear simultaneously.

The results of the analysis, i.e., reduction in dimensionality with PCA, of the vocabulary similarity distribution before and after the model improvement are illustrated in Figure8.

As shown in Figure8 (a), the original model suffered from low similarity scores due to a small number of training samples of certain words, which actually clustered words. In Figure8 (b), the model mitigated scores of words with high similarity towards manual scoring, and pulled scores of words with high frequency co-occurrence but low similarity apart in distribution.

Figure7: Performance comparison of different models

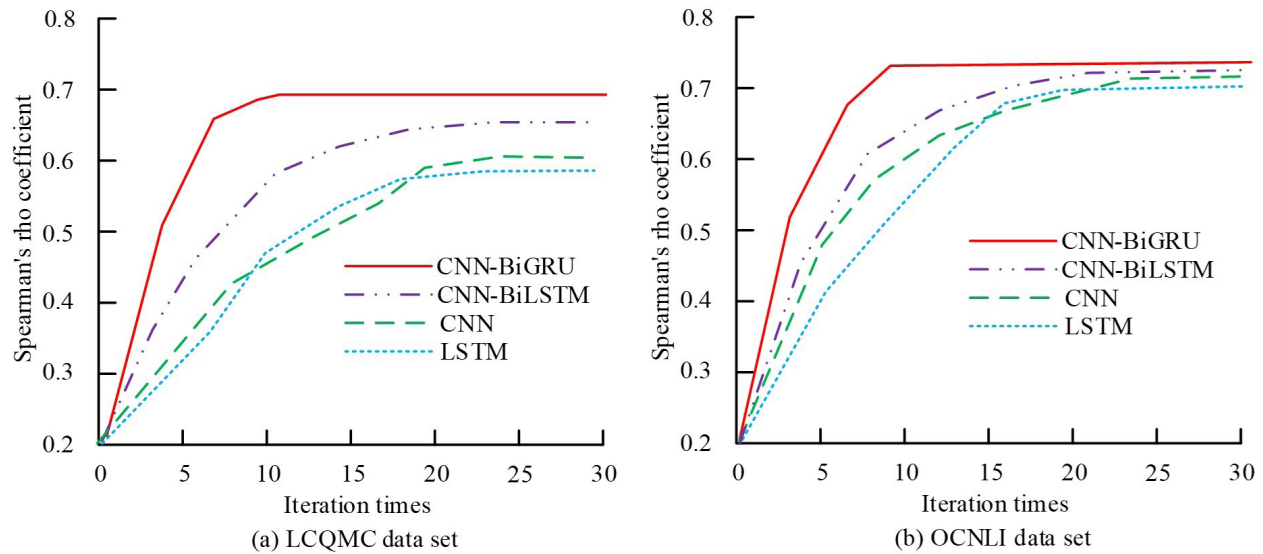


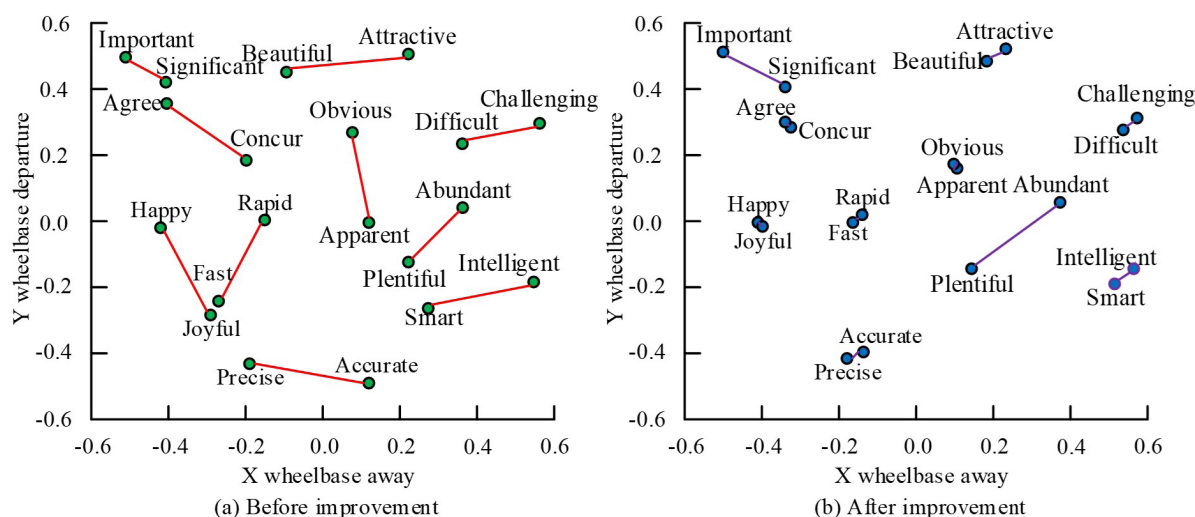
Table3. Comparison of similarity scores of different words before and after model improvement

Serial number	Synonym	Manual scoring	Before improvement	After improvement
1	Happy-Joyful	9.5	6.7	10.0
2	Fast-Rapid	8.4	5.1	9.9
3	Beautiful-Attractive	7.6	4.8	8.7
4	Important-Significant	9.2	7.2	9.5
5	Intelligent-Smart	7.3	5.4	8.8
6	Difficult-Challenging	8.2	6.5	9.5
7	Agree-Concur	9.3	6.0	10.0
8	Obvious-Apparent	8.5	5.9	10.0
9	Abundant-Plentiful	8.8	6.8	9.0
10	Precise-Accurate	8.9	5.3	9.5

## 4. Conclusions and Future Directions

In this paper, a VSS calculation model based on word vectors and CNNs was proposed to address the issues of low accuracy and lack of consideration of word order in existing methods. The experiment showed that the CNN-BiGRU model had the fastest convergence speed and the maximum Spearman coefficient was 0.07, 0.11, and 0.19 higher than CNN-BiGRU, CNN, and GRU, respectively. In different samples, the Spearman coefficient of the model could remain stable above 0.50, and the variation range of Pearson coefficient was [0.453, 0.524], indicating that the improved model had good stability. The maximum Pearson coefficient was 0.013, 0.021, and 0.018 higher than other models, respectively. The recall rate of the CNN-BiGRU model was 5.7%, 10.0%, 8.6%, and 1.8% higher than CNN, GRU, LSTM, and CNN-BiLSTM, respectively. The accuracy of the similarity calculation was 92.4% and 96.5% in the two datasets, which was 8.6% and 12.2% higher than the lowest LSTM. The improved model significantly increased the scoring of highly similar vocabulary, making it closer to manual scoring, significantly reducing the correlation between the probability of scoring and the number of words that appear simultaneously, and reducing scoring errors.

Figure8: Comparison of vocabulary similarity distribution before and after model improvement



In the future, research shall include not only the training and evaluation of common corpora, but also application-related corpora, to increase the generalization and domain adaptability of the model. Running the proposed model on specialized tasks in medicine, law, finance, etc., can demonstrate its ability to handle diverse linguistic structures. These fields exhibit distinct semantic distributions, syntactic structures, and contextual dependencies compared to general-language corpora. For example, medical texts contain technical acronyms and implicit causal relationships; legal documents emphasize the use of precise legal terminology and formal logic; financial news incorporates numerical expressions and domain-specific idioms. Such tasks can expose the model's limitations in processing domain-shifted input and can require adaptation strategies such as domain-specific fine-tuning, multi-task learning, or integration of external knowledge sources, e.g., medical ontologies, legal knowledge graphs. A systematic evaluation in these scenarios could provide deeper insight into the model's transferability, semantic sensitivity, and generalization in real-world natural language processing tasks.

Improving the interpretability of the similarity scoring mechanism is another direction of improvement. Despite their accuracy, deep learning models are typically black-boxes. Utilizing attention mechanisms or saliency visualizations could show what contextual aspects are taken into account when making semantic judgments, most prominently. the model's ability to adapt to low-resource environments and tested on noisy — user-generated text such as online reviews/social media would illustrate further scalability real-world applications.

## Funding

No

## Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Reference

- [1] Giabelli, A., Malandri, L., Mercurio, F., et al. (2022). Embeddings evaluation using a novel measure of semantic similarity. *Cognitive Computation*, 14(2), 749–763. <https://doi.org/10.1007/s12559-021-09934-5>
- [2] Yang, D., & Yin, Y. (2022). Evaluation of taxonomic and neural embedding methods for calculating semantic similarity. *Natural Language Engineering*, 28(6), 733–761. <https://doi.org/10.1017/S1351324922000295>
- [3] Vakulenko, M. O. (2023). Semantic comparison of texts by the metric approach. *Digital Scholarship in the Humanities*, 38(2), 766–771. <https://doi.org/10.1093/llc/fqac078>
- [4] Oussalah, M., & Mohamed, M. (2022). Knowledge-based sentence semantic similarity: Algebraical properties. *Progress in Artificial Intelligence*, 11(1), 43–63. <https://doi.org/10.1007/s13748-021-00296-8>
- [5] Wingfield, C., & Connell, L. (2023). Sensorimotor distance: A grounded measure of semantic similarity for 800 million concept pairs. *Behavior Research Methods*, 55(7), 3416–3432. <https://doi.org/10.3758/s13428-023-02118-1>

- [6] Triandini, E., Fauzan, R., Siahaan, D. O., et al. (2022). Software similarity measurements using UML diagrams: A systematic literature review. *Register: Jurnal Ilmiah Teknologi Sistem Informasi*, 8(1), 10–23. <https://doi.org/10.29303/j.regist.v8i1.3867>
- [7] Ismail, S., Shishtawy, T. E. L., & Alsammak, A. K. (2022). A new alignment word-space approach for measuring semantic similarity for Arabic text. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 18(1), 1–18. <https://doi.org/10.4018/IJSWIS.300729>
- [8] Pan, J. S., Wang, X., Yang, D., et al. (2024). Flexible margins and multiple samples learning to enhance lexical semantic similarity. *Engineering Applications of Artificial Intelligence*, 133, 108275–108294. <https://doi.org/10.1016/j.engappai.2024.108275>
- [9] Dai, B. (2024). Relational analysis of college English vocabulary - A reflection based on semantic association network modeling. *Applied Mathematics and Nonlinear Sciences*, 9(1), 64–82. <https://doi.org/10.2478/amns.2024.1.00006>
- [10] Ahmad, F., & Faisal, M. (2022). A novel hybrid methodology for computing semantic similarity between sentences through various word senses. *International Journal of Cognitive Computing in Engineering*, 3(6), 58–77. <https://doi.org/10.1080/23311916.2022.2124236>
- [11] Chauhan, S., Kumar, R., Saxena, S., et al. (2024). Semsyn: Semantic-syntactic similarity based automatic machine translation evaluation metric. *IETE Journal of Research*, 70(4), 3823–3834. <https://doi.org/10.1080/03772063.2023.2288606>
- [12] Osth, A. F., & Zhang, L. (2024). Integrating word-form representations with global similarity computation in recognition memory. *Psychonomic Bulletin & Review*, 31(3), 1000–1031. <https://doi.org/10.3758/s13423-023-02293-4>
- [13] Zhang, Y., Zhao, H., Wei, J., et al. (2022). Context-based semantic communication via dynamic programming. *IEEE Transactions on Cognitive Communications and Networking*, 8(3), 1453–1467. <https://doi.org/10.1109/TCCN.2022.3176618>
- [14] Asudani, D. S., Nagwani, N. K., & Singh, P. (2023). Impact of word embedding models on text analytics in deep learning environment: A review. *Artificial Intelligence Review*, 56(9), 10345–10425. <https://doi.org/10.1007/s10462-023-10409-8>
- [15] Rodriguez, P. L., & Spirling, A. (2022). Word embeddings: What works, what doesn't, and how to tell the difference for applied research. *The Journal of Politics*, 84(1), 101–115. <https://doi.org/10.1086/715618>
- [16] Mars, M. (2022). From word embeddings to pre-trained language models: A state-of-the-art walkthrough. *Applied Sciences*, 12(17), 8805–8817. <https://doi.org/10.3390/app12178805>
- [17] Eminagaoglu, M. (2022). A new similarity measure for vector space models in text classification and information retrieval. *Journal of Information Science*, 48(4), 463–476. <https://doi.org/10.1177/01655515211069131>
- [18] Ichien, N., Lu, H., & Holyoak, K. J. (2022). Predicting patterns of similarity among abstract semantic relations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 48(1), 108. <https://doi.org/10.1037/xlm0000948>
- [19] Gao, Q., Huang, X., Dong, K., et al. (2022). Semantic-enhanced topic evolution analysis: A combination of the dynamic topic model and word2vec. *Scientometrics*, 127(3), 1543–1563. <https://doi.org/10.1007/s11192-022-04337-8>
- [20] Zhang, Y., Zhang, C., & Hu, F. (2025). Optimization of science and technology project management system based on hybrid semantic similarity evaluation framework. *Journal of Computational Methods in Sciences and Engineering*, 25(4), 3384–3396. <https://doi.org/10.1177/14727978251319396>

# Image Classification in Coal Production Using Deep Neural Networks: A Comprehensive Benchmarking Study

Wenmi Chai<sup>1\*</sup>, Zhiyao Yang<sup>2</sup>, Rui Zhao<sup>3,4</sup>, Qian Xiang<sup>5</sup>, Xinxin Niu<sup>1</sup>, Ling Liang<sup>1</sup>

1.New Energy Technology Research Institute Co., Ltd., CHN ENERGY Investment Group Co., Ltd., Beijing, 102209, China

2.National Institute of Clean-and-Low-Carbon Energy, Beijing, 102209, China

3.Chengdu Jinjiang Center for Disease Control and Prevention, Chengdu, 610000, China

4.School of Economics and Management, Sichuan Normal University, Chengdu, 610101, China

5.Laboratory of Intelligent Control, PLA Rocket Force University of Engineering, Xi'an, 710025, China

\*Corresponding author: Wenmi Chai, wenmi.chai@ceic.com

**Copyright:** 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

**Abstract:** During the intelligent transformation of the coal industry, image classification technology plays a crucial role in process monitoring, quality inspection, and safety early warning. Taking the DsCGF-1 dataset in the coal production environment as the research object, this study conducts a multi-dimensional performance and efficiency evaluation on 12 mainstream deep learning models, aiming to establish industrial-level model selection criteria for intelligent coal separation. The results indicate that RepVGG-B3 exhibits the optimal comprehensive performance, with a test accuracy of 97.92%, a coal recall rate of 99.8%, and the best AUC value across all categories. Furthermore, RepViT-M3 achieves a test accuracy of 97.85% with a parameter count of merely 9.66M, demonstrating excellent lightweight characteristics, which makes it suitable for resource-constrained scenarios such as underground edge computing. This study establishes a model selection benchmark for coal separation, providing technical support for the development of intelligent sorting systems in industrial scenarios.

**Keywords:** Coal Separation; Image Classification; Deep Learning; Convolutional Neural Network; Model Benchmarking

**Published:** Dec 27, 2025

**DOI:** <https://doi.org/10.62177/jaet.v2i4.958>

## 1.Introduction

### 1.1 Research Background and Importance

As the cornerstone of China's energy security, coal plays an irreplaceable strategic role in ensuring the country's energy independence and self-reliance against the backdrop of the restructuring of the global industrial chain and energy supply chain[1]. In 2024, coal consumption accounted for 53.2% of China's total primary energy consumption[2]. In the advancement of the "carbon peaking and carbon neutrality" goals, the clean and efficient utilization of coal as well as its intelligent sorting directly influence the process of low-carbon transformation in the coal industry [1,2]. Traditional coal separation technologies struggle to meet the requirements of real-time performance and accuracy in modern industry [3,4]. With the development of computer technology and digital image acquisition equipment, machine vision technology, leveraging the advantage of non-contact detection, has become a core support for breaking through the efficiency bottleneck of traditional mine separation[4].

Early machine vision methods relied on the combination of manual feature extraction and machine learning-based classification. Numerous researchers have constructed ore recognition models using support vector machines (SVMs) [5–7], which exhibit favorable recognition performance on images collected in laboratory environments. However, these studies are detached from actual production scenarios and impose stringent requirements on both the acquisition environment and image resolution. The emergence of deep learning technology has offered a brand-new solution for coal image classification. A number of scholars [4,8–13] have adopted convolutional neural networks to replace the manual feature extraction process, which reduces the reliance on high-resolution images and enhances the classification accuracy of the models. Notably, Lv et al. [3] recently released the large-scale coal image dataset DsCGF, which includes coal images captured under both production and non-production conditions. This dataset provides a reliable benchmark for the systematic evaluation of the performance of deep learning models in real industrial scenarios.

## 1.2 Overview of Deep Learning Models

The aforementioned deep learning-based coal image classification research has laid the foundation for the technical advancement of intelligent sorting equipment in production environments. In recent years, deep learning model architectures have undergone continuous iterative optimization, forming a multi-technical route pattern that caters to diverse computational resource constraints and accuracy requirements, thereby providing abundant technical options for coal image classification tasks:

- ResNet [14]: Alleviates the gradient vanishing problem in deep networks through a residual connection mechanism, and simplifies training via residual mapping.
- Xception [15]: Adopts depthwise separable convolution to decompose the standard convolution process, significantly reducing computational complexity while ensuring performance.
- ResNeXt [16]: Introduces the concept of cardinality and parallel grouped convolution, enhancing feature diversity without increasing the network scale.
- MNASNet [17]: Pioneers platform-aware neural architecture search, with the latency of mobile devices as a constraint, achieving an accuracy of 75.2% with a latency of 78ms on Pixel phones.
- RepVGG [18]: Employs a reparameterization strategy of multi-branch training and single-path inference, featuring excellent model training accuracy and deployment inference speed.
- GhostNetV2 [19]: Proposes a hardware-friendly attention mechanism to enhance the extended features generated by low-cost operations, which can aggregate both local and long-range information simultaneously and optimize the performance of the network architecture.
- MobileOne [20]: Addresses the weak correlation between the traditional optimization objectives (such as FLOPs or parameter quantity) and the actual inference latency of devices by alleviating the architectural and optimization bottlenecks of existing efficient networks, realizing dual improvements in inference speed and accuracy on mobile devices.
- FastViT [21]: Optimizes memory access efficiency using the RepMixer module based on structural reparameterization, achieving a significant improvement in mobile inference speed while maintaining high accuracy, and demonstrating excellent latency-accuracy trade-off and cross-task generalization capabilities.
- ConvNeXt V2 [22]: Enhances the representation learning capability of pure convolutional networks by introducing a fully convolutional masked autoencoder framework and a global response normalization layer, achieving breakthrough performance in multiple visual tasks and providing a full range of efficient models with parameters ranging from 3.7M to 650M.
- RepViT [23]: Integrates the efficient design of Vision Transformer (ViT) into lightweight CNNs, achieving a latency of 1.0 millisecond with an accuracy exceeding 80% on the iPhone 12, and its inference speed increases by 10 times when combined with the Segment Anything Model (SAM).
- MobileNetV4 [24]: Adopts a universal Unified Inception Block (UIB) and a mobile-optimized attention mechanism to achieve comprehensive high efficiency across mobile hardware. Its large model reaches 87% accuracy on ImageNet with only 3.8ms inference latency on EdgeTPU.
- InceptionNeXt [25]: Decomposes large-kernel depthwise convolution along the channel dimension into four parallel branch-

es: small square kernels, two orthogonal strip kernels, and identity mapping. It maintains the advantage of large receptive field while significantly improving training throughput, and achieves a 0.2% accuracy improvement on ImageNet, serving as a lightweight architecture that balances performance and energy efficiency.

### 1.3 Research Motivation and Structure Arrangement

Each of the aforementioned models has its own advantages in terms of accuracy, efficiency, and deployment adaptability. However, most existing studies focus on the laboratory verification of a single model, lacking systematic evaluation of mainstream models in real production environments, which makes it difficult to form industrial-level model selection criteria. Based on this, this paper conducts a multi-dimensional evaluation on 12 advanced deep learning models, aiming to provide technical support for the research and development of intelligent coal separation equipment.

The structure of this paper is arranged as follows: Section 2 introduces the data and methods used in this study; Section 3 presents the multi-dimensional comparison results and corresponding analysis; Section 4 discusses the research limitations and future research directions; finally, the research achievements of the whole paper are summarized.

## 2. Material and Methods

### 2.1 Dataset

This study adopts the production-conditions subset of the Guobei Coal Preparation Plant in Anhui Province from the DsCGF-1 dataset released by Lv et al. [3] as the experimental data source. Collected from a real industrial environment, the dataset was acquired using an acA4096-40gc industrial camera manufactured by Basler, which was installed on the manual sorting conveyor belt. During the acquisition process, a stable illumination of 1200 ( $\pm 100$ ) Lux was maintained to ensure that the image quality meets industrial inspection standards. The dataset includes four categories: coal, gangue, unknown and foreign objects. All annotation work underwent strict quality control and was completed by an annotation team that passed professional assessments, ensuring the consistency and accuracy of the annotations.

The dataset was split into training, validation, and test sets in a 6:2:2 ratio based on the chronological order of image acquisition, effectively mitigating data leakage. As shown in Table 1, all images were uniformly preprocessed to a resolution of 224 $\times$ 224 pixels and normalized using the statistical parameters of the ImageNet dataset, providing a standardized data foundation for model training and evaluation.

Table 1: Subset division of the DsCGF-1 dataset and the corresponding sample sizes.

Dataset	Classes of Coal Images				Sum of Images
	Coal	Gangue	Unknown	Foreign Objects	
Training Set	2,926	57,799	2,201	1,890	64,816
Validation Set	975	19,266	734	630	21,605
Test Set	976	19,266	734	630	21,606
Total	4,877	96,331	3,669	3,150	108,027

This data construction method based on real industrial scenarios not only maintains the authenticity of the production environment but also ensures the reliability of model evaluation through scientific partitioning methods, laying a solid foundation for the performance evaluation of the deep learning models in this study.

### 2.2 Experimental Setup

The experimental environment of this study is built on the Windows 11 operating system. The hardware platform configuration includes an NVIDIA GeForce RTX 3090 GPU, an Intel Core i5-12400F processor (2.5 GHz, 6 cores and 12 threads), and 32 GB of physical memory. For the software environment, Python 3.12.11 is adopted as the programming language. The deep learning framework is based on PyTorch 2.8.0+cu129, and it is combined with the Torchvision 0.23.0+cu129 and Timm 1.0.19 libraries to implement model construction and training.

### 2.3 Data Preprocessing

The image preprocessing workflow adopts a standard computer vision task pipeline, with specific steps as follows:

1. Size Adjustment: All images are uniformly resized to a specified square size;

2. Format Conversion: Convert PIL images into PyTorch tensors, and normalize the pixel values to the range of [0, 1];
3. Standardization: Perform standardization using the statistical parameters of the ImageNet dataset, with a mean of [0.485, 0.456, 0.406] and a standard deviation of [0.229, 0.224, 0.225].

Regarding the configuration of the data loader, the key parameters are specified as follows: For the training set, the batch size is set to 128 with random data shuffling enabled, while the validation and test sets adopt a batch size of 256 without altering the data order, which ensures accurate evaluation of model performance. The optimizer used for all datasets is AdaBoB[26], with a total of 90 training epochs. The initial learning rate is set to  $10^{-3}$ , and it is decayed by a factor of 0.1 every 30 training epochs. Additionally, a weight decay of  $10^{-4}$  is applied to prevent overfitting.

## 2.4 Evaluation Metrics

This study employs standard evaluation metrics to quantitatively assess the model performance. For the four-class coal image classification task (coal, gangue, unknown and foreign objects), the evaluation is performed based on the confusion matrix, whose structure is presented in Table 2. This matrix intuitively illustrates the model's classification performance across different categories by conducting statistical analysis of the correspondence between true labels and predicted labels.

Table2: Four-Class Confusion Matrix

Class	Predicted coal	Predicted gangue	Predicted unknown	Predicted foreign object
Actual coal	$TP_{11}$	$FP_{12}$	$FP_{13}$	$FP_{14}$
Actual gangue	$FP_{21}$	$TP_{22}$	$FP_{23}$	$FP_{24}$
Actual unknown	$FP_{31}$	$FP_{32}$	$TP_{33}$	$FP_{34}$
Actual foreign object	$FP_{41}$	$FP_{42}$	$FP_{43}$	$TP_{44}$

Note:  $TP_{ii}$  denotes the number of correct predictions for class  $i$ , and  $FP_{ij}$  denotes the number of samples from class  $i$  incorrectly predicted as class  $j$  ( $i \neq j$ ).

The classification performance metrics include Accuracy, Precision, Recall, and F1-Score. Accuracy reflects the proportion of all samples correctly classified; Precision denotes the proportion of correctly predicted samples among those detected as a specific class; Recall evaluates the proportion of actual positive samples of a target class that are correctly identified as such class; F1-Score provides a comprehensive measure of Precision and Recall. Based on the confusion matrix, the calculation formulas for each evaluation metric are as follows, where  $N$  denotes the total number of samples,  $w_i$  represents the weight of class  $i$ .

Accuracy: Overall classification accuracy.

$$Accuracy = \frac{\sum_{i=1}^4 TP_{ii}}{N} \quad (1)$$

Precision: Prediction accuracy calculated for each class  $i$ , followed by a weighted average.

$$Precision = \sum_{i=1}^4 w_i \cdot Precision_i = \sum_{i=1}^4 w_i \cdot \frac{TP_{ii}}{TP_{ii} + \sum_{j=1, j \neq i}^4 FP_{ji}} \quad (2)$$

Recall: Recognition completeness calculated for each class  $i$ , followed by a weighted average.

$$Recall = \sum_{i=1}^4 w_i \cdot Recall_i = \sum_{i=1}^4 w_i \cdot \frac{TP_{ii}}{TP_{ii} + \sum_{j=1, j \neq i}^4 FP_{ij}} \quad (3)$$

F1-Score: Harmonic mean of Precision and Recall, weighted by class.

$$F1-Score = \sum_{i=1}^4 w_i \cdot 2 \times \frac{Precision_i \times Recall_i}{Precision_i + Recall_i} \quad (4)$$

Considering the class imbalance characteristics of the coal image dataset, all metrics are calculated using the weighted average method, where the weight of each class is the proportion of samples in that class.

## 3. Experimental Results and Analysis

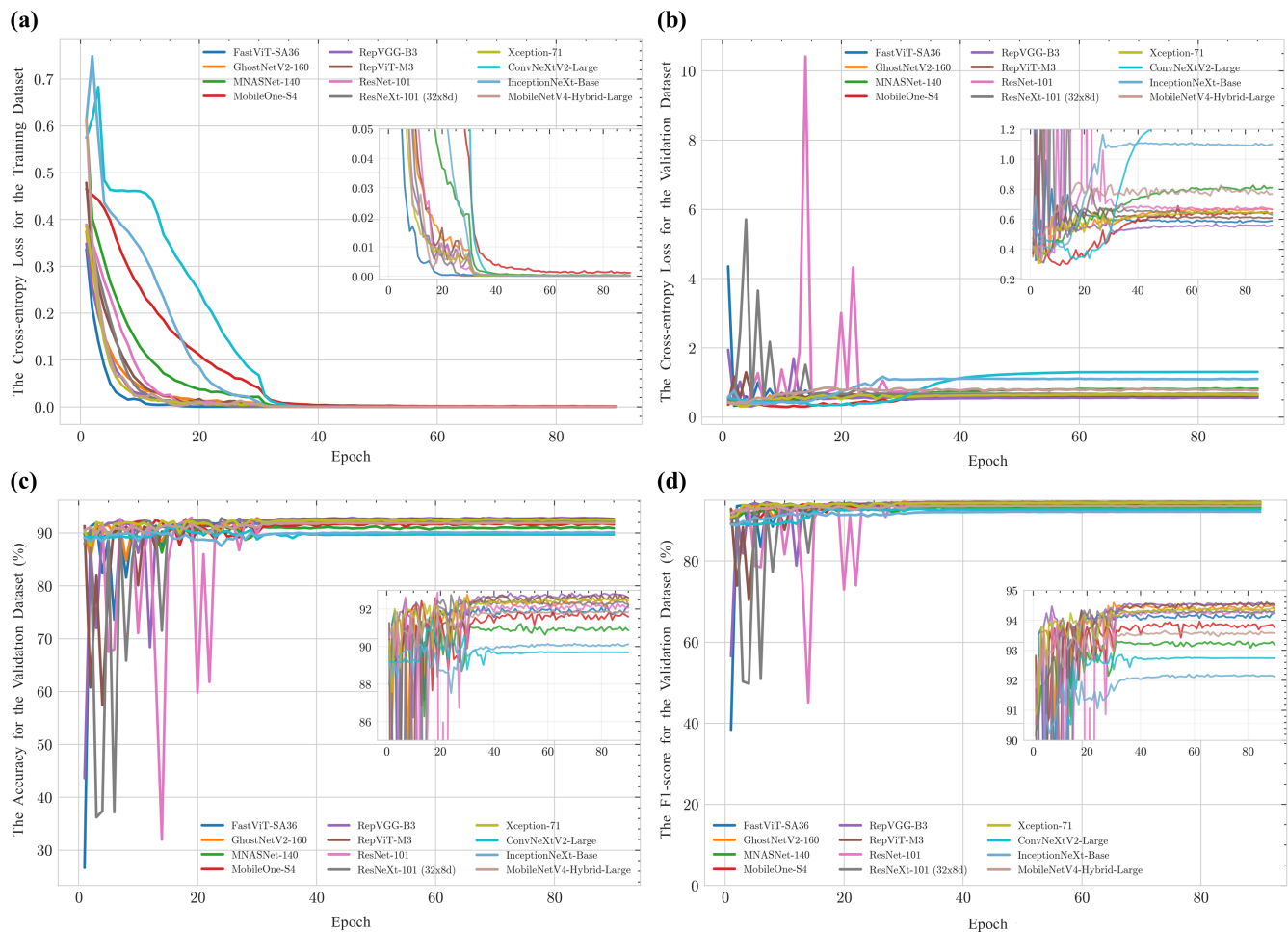
### 3.1 Model Training

By investigating the performance metrics of 12 deep learning models over 90 training epochs, this study systematically analyzed the dynamic evolution patterns of the models on both the training set and validation set. Figure 1 illustrates the

training processes of these 12 models on the training set and validation set. The loss values of all models on the training set exhibited a typical exponentially decaying convergence pattern, decreasing rapidly and then stabilizing (Figure 1a). Initially, the loss values ranged from 0.3 to 0.8; after 40 epochs, all converged to below 0.05, indicating that all models possessed the ability to effectively capture the essential features of coal images. However, significant differences were observed among different models in terms of convergence speed and training stability. Among them, FastViT-SA36 exhibited the fastest convergence speed: its loss dropped below 0.05 after 20 epochs, with minimal overall fluctuations. ConvNeXtV2-Large and InceptionNeXt-Base showed slower convergence and exhibited abnormal peaks in the early training stages, but stabilized after 40 epochs. From the locally enlarged view, it can be observed that MobileOne-S4 still presented slight fluctuations after 40 epochs, indicating its relatively weak training stability.

Figure 1b shows the loss of 12 models on the validation set, where a smaller loss indicates better generalization ability of the model. The loss of all models on the training set is less than 0.01, while their loss on the validation set stabilizes in the range of 0.3-0.8. Combined with the training set loss curves of each model, ConvNeXtV2-Large and InceptionNeXt-Base perform poorly: their convergence speed is lower than that of other models, and their validation loss on the validation set is higher than that of other models. Although ResNeXt-101 (32×8d) and ResNet-101 perform well on the training set, their validation loss fluctuates significantly in the first 25 epochs. After 40 epochs, the validation loss curves of all models gradually stabilize. The loss values of InceptionNeXt-Base, MNASNet-140, and MobileOne-S4 are greater than 0.7, which indicates that lightweight models do have an impact on performance while reducing model parameters.

Figure1: Training results of 12 deep learning models for coal image classification task on DsCGF-1 dataset. (a) Cross-entropy loss convergence curve on training dataset; (b) Cross-entropy loss generalization performance curve on validation dataset; (c) Accuracy curve on validation dataset; (d) F1-score curve on validation dataset.



Consistent conclusions can be drawn from the variation trends of accuracy and F1-score on the validation set (Figure 1c, d). ResNeXt-101 (32×8d) and ResNet-101 experienced severe fluctuations during the 0–20 epoch period: their minimum accuracy dropped below 40%, and the F1-score fell below 60%. This completely corresponds to the loss peaks of the two models on the validation set (Figure 1b), verifying the stability issues of these models. FastViT-SA36 also showed similar fluctuations but with a smaller amplitude. Eventually, the accuracy of each model also exhibited obvious differentiation: the accuracy of ConvNeXtV2-Large and InceptionNeXt-Base was around 90%, that of MNASNet-140 stabilized at approximately 91%, and the accuracy of most other models reached about 92%. A similar differentiated result was also observed in the F1-score, but the differentiation appeared earlier in the training epochs.

In terms of model performance effectiveness, Accuracy, as the core evaluation metric for classification tasks, reveals distinct hierarchical differences among the 12 models (Table 3). Models with high accuracy include ResNet-101 (92.88%), RepVGG-B3 (92.84%), RepViT-M3 (92.79%), GhostNetV2-160 (92.77%), and ResNeXt-101 (32×8d) (92.75%). These models exhibit superior performance in complex feature extraction and classification decision-making. Models with moderate accuracy cover MNASNet-140 (91.22%), ConvNeXtV2-Large (91.45%), MobileNetV4-Hybrid-Large (91.95%), and FastViT-SA36 (92.17%), whose performance is slightly inferior to that of the high-accuracy group. In contrast, InceptionNeXt-Base achieves the lowest validation accuracy (90.15%), showing poor overall classification performance. Models with high Precision generally have a Precision rate of over 99%, among which GhostNetV2-160 (99.96%), RepVGG-B3 (99.95%), and Xception-71 (99.94%) perform optimally, indicating an extremely low misjudgment rate in their prediction of positive samples. GhostNetV2-160 and RepViT-M3 achieve the highest F1-scores (both 94.59%), followed by RepVGG-B3 (94.54%) and Xception-71 (94.31%). This demonstrates that these models have greater advantages in balancing the classification performance of positive and negative samples, making them particularly suitable for task scenarios with imbalanced category distribution.

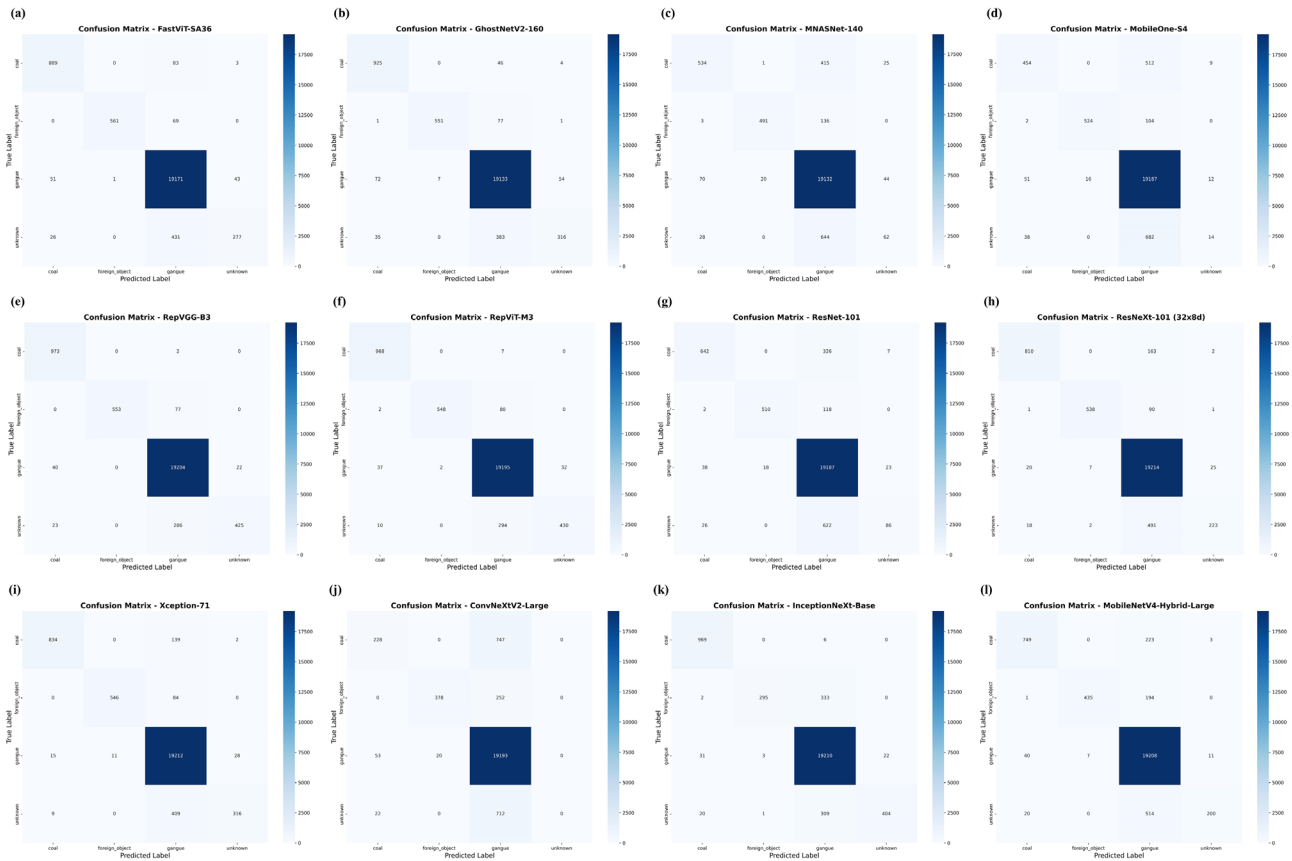
Table3: Results of 12 models on the DsCGF-1 validation set.

Model	Loss ( $\times 10^{-3}$ )	Accuracy (%)	Precision (%)	F1-score (%)
InceptionNeXt-Base	110.05	90.15	99.86	92.14
MNASNet-140	60.29	91.22	99.92	93.17
ConvNeXtV2-Large	32.98	91.45	96.48	92.69
MobileNetV4-Hybrid-Large	80.07	91.95	99.31	93.59
FastViT-SA36	58.88	92.17	99.94	94.24
MobileOne-S4	30.28	92.42	99.46	93.93
Xception-71	57.47	92.6	99.94	94.31
ResNeXt-101 (32x8d)	69.76	92.75	99.4	94.3
GhostNetV2-160	56.81	92.77	99.96	94.59
RepViT-M3	61.29	92.79	99.47	94.59
RepVGG-B3	56.93	92.84	99.95	94.54
ResNet-101	79.32	92.88	98.87	94.23

### 3.3 Test Results

To comprehensively evaluate the classification capability of each model in the coal image task, this study conducts an analysis of the performance of 12 models based on the results of the DsCGF-1 test set. Figure 2 presents the confusion matrix results of four image categories, and the models show significant differences in performance.

Figure 2: The confusion matrix of the 12 models on the DsCGF-1 test set: (a) FastViT-SA36, (b) GhostNetV2-160, (c) MNASNet-140, (d) MobileOne-S4, (e) RepVGG-B3, (f) RepViT-M3, (g) ResNet-101, (h) ResNeXt-101 (32x8d), (i) Xception-71, (j) ConvNeXtV2-Large, (k) InceptionNeXt-Base and (l) MobileNetV4-Hybrid-Large. X-axis: predicted labels; Y-axis: truth labels.



FastViT-SA36 achieved relatively high precision in classifying coal, gangue, and foreign objects, yet 432 unknown samples were misidentified as gangue (Figure 2a). GhostNetV2-160 exhibited similar performance to FastViT-SA36, with 383 unknown samples misclassified as gangue (Figure 2b). MNASNet-140 performed slightly worse: only 534 coal samples were correctly identified, while nearly 42% of coal samples, 21% of foreign objects, and 88% of unknown samples were misidentified as gangue (Figure 2c). MobileOne-S4 showed comparable performance to MNASNet-140, merely 454 coal samples were correctly recognized, and approximately 53% of coal samples, 16.5% of foreign objects, and 93% of unknown samples were misclassified as gangue (Figure 2d). RepVGG-B3 delivered the optimal performance, as nearly all coal samples were correctly identified. However, a small number of gangue and unknown samples were misclassified as coal. Additionally, the recall rate of this model for unknown samples—an area where other models performed poorly—reached approximately 58% (Figure 2e). RepViT-M3 exhibited similar performance to RepVGG-B3 but with slightly lower accuracy and precision across all categories (Figure 2f). ResNet-101 performed poorly, with a large number of coal, foreign object, and unknown samples misidentified as gangue (Figure 2g). ResNeXt-101 (32×8d) showed similar performance to ResNet-101 but was slightly more effective (Figure 2h). Xception-71 outperformed ResNet-101 yet was inferior to RepVGG-B3, as a considerable number of coal and unknown samples were still misclassified as gangue (Figure 2i). ConvNeXtV2-Large had very limited effectiveness, a large number of samples from other categories were misidentified as gangue. In particular, 77% of coal samples and 97% of unknown samples were misclassified as gangue, and none of the unknown samples were correctly recognized (Figure 2j). InceptionNeXt-Base fell into the second-tier performance group; its recall rate for coal classification reached 99.4%, but its performance in recognizing foreign objects and unknown samples was inferior to that of RepVGG-B3 (Figure 2k). MobileNetV4-Hybrid-Large showed similar performance to ConvNeXtV2-Large, with a large number of samples from other categories misidentified as gangue—more than 70% of foreign objects were misclassified as gangue

(Figure 2l).

By observing the confusion matrices, it can be found that more than 75% of the models perform poorly in classifying unknown samples and gangue. Some models, such as ConvNeXtV2-Large, even misidentify a large number of coal samples as gangue, which will undoubtedly cause economic losses in the production process. For coal production, image classification serves intelligent sorting, which requires as accurate identification of coal as possible to improve production and operation benefits. Therefore, the recall rate of the coal category is particularly important. Among the 12 models, RepVGG-B3 performs the best, with a recall rate of 99.8% and a precision rate of 93.9% for the coal category, which is of great significance to the coal production process.

Figure 3: The ROC curves of 12 models on the DsCGF-1 test set, with the abscissa representing the false positive rate and the ordinate representing the true positive rate. Among them, (a) coal, (b) gangue, (c) unknown and (d) foreign object.

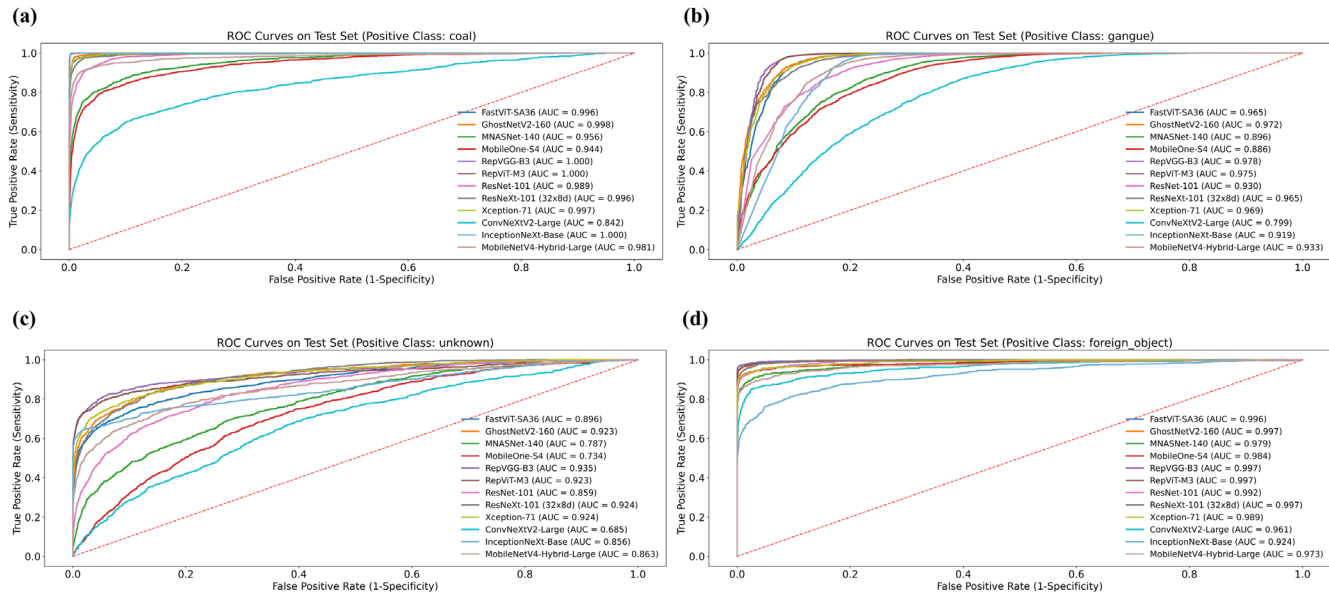


Figure 3 presents the ROC curve results of 12 models, where the AUC value (Area Under the ROC Curve) is used to evaluate the performance of classifiers. The closer the AUC value is to 1, the better the classifier performance; conversely, the closer the AUC value is to 0, the poorer the classifier performance. The AUC values of RepVGG-B3 for coal, gangue, unknown samples, and foreign objects are 1, 0.978, 0.935, and 0.997 respectively, which are the highest among the 12 models in all categories. This is consistent with the results of the confusion matrix. The AUC values of ConvNeXtV2-Large for the four categories are 0.842, 0.799, 0.685, and 0.961 respectively, ranking at a low level among the 12 models. Therefore, in terms of recognition performance alone, RepVGG-B3 is the optimal model on the DsCGF-1 test set, while ConvNeXtV2-Large has poor comprehensive performance and is not suitable as an intelligent coal separation model in this scenario. However, due to the limited equipment conditions in the coal production scenario, factors such as model training time, parameter quantity, and hardware requirements also need to be considered.

### 3.3 Model Complexity

Parameter scale and training time directly determine the deployment cost and training cost of a model. The results show (Table 4) that lightweight models such as MNASNet-140(5.84M), RepViT-M3(9.66M), GhostNetV2-16(11.12M), MobileOne-S4 (12.91M), and MobileNetV4-Hybrid-Large (36.49M) all have their parameter quantities controlled within 40M, which are quite consistent with the requirements of deployment scenarios with limited computing power, such as edge devices at coal mine wellheads. As models adapted to mobile lightweight scenarios, MNASNet-140 (5.84M), GhostNetV2-160, MobileOne-S4, RepViT-M3, and MobileNetV4-Hybrid-Large have small parameter quantities, which are much lower than those of large models like ConvNeXtV2-Large (196.43M) and RepVGG-B3 (120.53M).

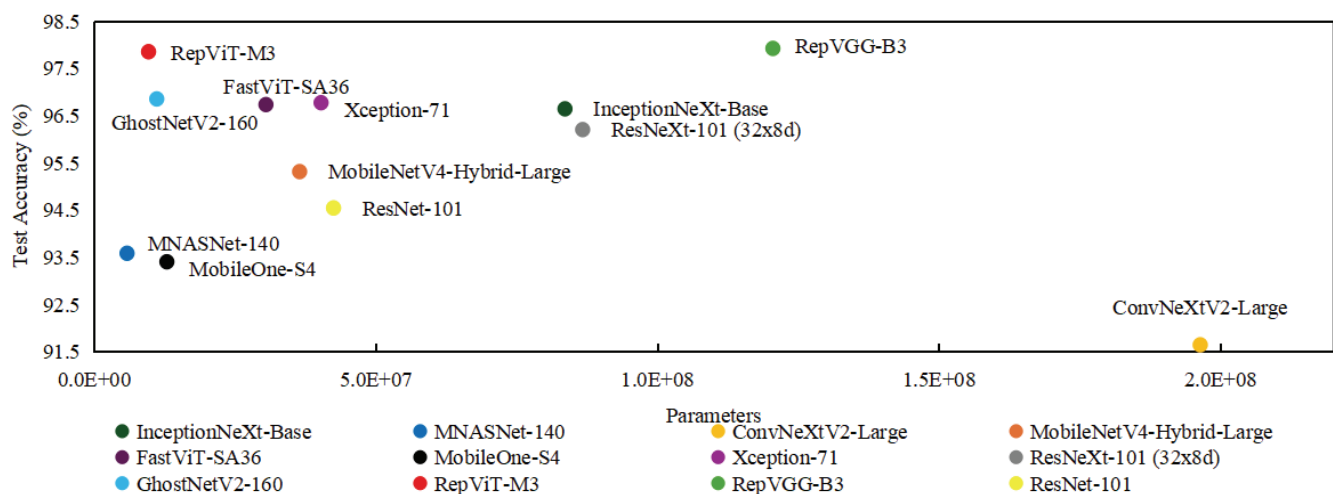
Table 4: Parameter quantity and training time of 12 models after 90 epochs on the DsCGF-1 set.

Model	Parameters(M)	Training Time (s)
InceptionNeXt-Base	83.61	20812.82
MNASNet-140	5.84	11774.49
ConvNeXtV2-Large	196.43	26651.59
MobileNetV4-Hybrid-Large	36.49	19378.11
FastViT-SA36	30.51	23373.23
MobileOne-S4	12.91	25883.84
Xception-71	40.3	29388.68
ResNeXt-101 (32x8d)	86.75	20791.29
GhostNetV2-160	11.12	28372.77
RepViT-M3	9.66	18442.13
RepVGG-B3	120.53	17857.85
ResNet-101	42.51	15773.18

In terms of training efficiency and time consumption, MNASNet-140 has obvious advantages among lightweight models, with a training time of only 11,774.49 seconds. As the fastest-trained model among the 12 models, it can reduce the model iteration cost. RepViT-M3 has a training time of 18,442.13 seconds, which is at a medium-low level. The training time of MobileNetV4-Hybrid-Large is 19,378.11 seconds, while that of MobileOne-S4 reaches 25,883.84 seconds. Although GhostNetV2-160 has only 11.12M parameters, its training time is 28,372.77 seconds. It is inferred that its attention mechanism increases the computational complexity of feature processing. Xception-71 has a medium-level parameter quantity, but its training time is the longest among all models.

Considering the parameter quantity and accuracy comprehensively (Figure 4), RepViT-M3 performs the best among lightweight models. This advantage benefits from the integration of ViT's efficient design into lightweight CNNs, enabling its pure convolutional structure to achieve Transformer-like global feature capture. The recall rate of coal classification reaches 99.3%, which is slightly lower than that of RepVGG-B3. In addition, compared with RepViT-M3, the parameter quantity of RepVGG-B3 is two orders of magnitude larger, but its training time is shorter. Moreover, RepVGG-B3 has distinct training and inference structures: during training, it adopts a multi-branch structure consisting of  $3 \times 3$  convolution,  $1 \times 1$  convolution and identity mapping to enhance feature extraction capability and alleviate gradient vanishing; during inference, through structural reparameterization, the multi-branch parameters are equivalently fused into a single  $3 \times 3$  convolution layer, forming an extremely simple architecture similar to VGG, which further improves the inference speed.

Figure 4: Scatter plot of test accuracy and parameter quantity for 12 models, where the X-axis represents parameter quantity and the Y-axis represents test accuracy.



## 4. Conclusions and Recommendations

### 4.1 Conclusions

This study systematically evaluated the performance and efficiency of 12 deep learning models in the coal image classification task, and revealed the key patterns of the models in terms of training dynamics, classification accuracy, and resource consumption. The main findings are as follows:

1. In terms of the training and validation process, the training set loss of all models shows an exponential decay trend, converging to below 0.05 after 40 training epochs, which proves their ability to capture the core features of coal images. In the later training stage, there are obvious differences in convergence speed and stability: FastViT-SA36 converges the fastest (loss below 0.05 after 20 epochs), while ConvNeXtV2-Large and Inception Next-Base converge slowly with significant early fluctuations. The validation set loss stabilizes in the range of 0.3-0.8, and that of InceptionNext-Base, MNASNet-140 and MobileOne-S4 exceeds 0.7, indicating that lightweight models do affect performance while reducing parameters.
2. The core contradiction of model efficiency lies in the balance among parameter scale, training time, and performance. Lightweight models with parameters less than 40M, such as MNASNet-140 and RepViT-M3, are suitable for resource-constrained scenarios like edge devices at coal mine. However, training efficiency is not entirely determined by parameter quantity. MNASNet-140, with 5.84M parameters, achieves the shortest training time of 11,774.49 seconds, while GhostNetV2-160 (11.12M parameters) takes 28,372.77 seconds for training due to its attention mechanism. Among large models, RepVGG-B3 shows an advantage: although its parameter quantity (120.53M) is one order of magnitude higher than that of RepViT-M3, its training time (17,857.85s) is shorter, achieving a good training efficiency.
3. Considering classification performance, efficiency, and scenario adaptability, RepVGG-B3 and RepViT-M3 have emerged as the optimal representatives of large and lightweight models, respectively. RepVGG-B3 ranks first in test set accuracy (97.92%), coal recall rate (99.8%), and AUC values of various categories (coal AUC=1). The confusion matrix shows that it can almost correctly identify all coals, with an unknown object recall rate of 58%, making it the best model for intelligent sorting in coal production. RepViT-M3, with 9.66M parameters, achieves a high accuracy of 97.85% and a coal recall rate of 99.3%. Through its integrated architecture of “lightweight CNN + ViT design”, it achieves the optimal balance between performance and efficiency in resource-constrained scenarios. In addition, models have common shortcomings in category recognition: over 75% of models perform poorly in classifying unknown objects and gangue, and some models (such as ConvNeXtV2-Large) even misclassify 77% of coal as gangue, highlighting the importance of task adaptability in model selection.

### 4.2 Practical Significance and Industrial Application Suggestions

Based on the above conclusions, this study provides the following practical suggestions for the technological implementation of intelligent coal separation systems:

Differentiated model deployment schemes should be adopted for different application scenarios. Large-scale coal mine separation centers should give priority to RepVGG-B3, and leverage its advantage in high-precision coal identification to improve washing and separation efficiency, thereby reducing economic losses. For small and medium-sized coal mines or edge device, RepViT-M3 is recommended.

To promote the technological implementation and continuous optimization, it is also necessary to establish a guarantee mechanism from three dimensions: data, hardware, and mechanism. At the data level, it is suggested that coal mining enterprises cooperate with scientific research institutions to build an industry-level annotated dataset, which includes image samples from different regions, coal types, and working conditions, so as to solve the scenario limitations of the existing test set. At the hardware level, adaptive hardware solutions can be designed for the optimized models—for example, configuring a dedicated inference chip for RepVGG-B3 to improve speed, and developing edge computing modules for RepViT-M3 to reduce deployment costs. At the mechanism level, a closed-loop mechanism of “model performance monitoring - data update - fine-tuning optimization” can be established: misjudged samples from the separation site are collected regularly, and model parameters are optimized through incremental training to ensure that model performance is continuously adapted to changes in working conditions.

In addition, for enterprises with strong technical capabilities, model improvement can be carried out based on the conclusions of this study. For instance, drawing on the structural design idea of RepVGG-B3, customized models with “lightweight and reparameterization” can be developed to further compress parameters while retaining the advantage of high performance. Future work can expand sample diversity, explore multi-modal data fusion, and develop customized models balancing lightweight and high performance, so as to promote the intelligent upgrading of coal separation technology.

## Funding

This research was supported by the Science and Technology Project of CHN ENERGY Investment Group entitled “Research and Application of an Integrated Intelligent Simulation Platform for Full Industrial Chain Production and Operations Based on Digital Twin Technology.” (Grant No. GJNY-23-176).

## Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Reference

- [1] Kang H, Xie H, Ren S, et al. (2022) Development Strategy of China’s Coal Industry under the Reconstruction of Global Industrial Chain and Energy Supply Chain. *Chinese Journal of Engineering Science* 24: 26.
- [2] National Bureau of Statistics of China. Statistical Communiqué on National Economic and Social Development of the People's Republic of China in 2024 [EB/OL]. (2025-02-28). [https://www.stats.gov.cn/sj/zxfb/202502/t20250228\\_19588-17.html](https://www.stats.gov.cn/sj/zxfb/202502/t20250228_19588-17.html).
- [3] Lv Z, Fan Y, Sha T, et al. (2025) A large-scale open image dataset for deep learning-enabled intelligent sorting and analyzing of raw coal. *Sci Data* 12: 403.
- [4] Liu Y, Zhang Z, Liu X, et al. (2021) Deep Learning Based Mineral Image Classification Combined With Visual Attention Mechanism. *IEEE Access* 9: 98091–98109.
- [5] Patel AK, Chatterjee S, Gorai AK (2019) Development of a machine vision system using the support vector machine regression (SVR) algorithm for the online prediction of iron ore grades. *Earth Sci Inform* 12: 197–210.
- [6] Wang W, Lv Z, Lu H (2021) Research on methods to differentiate coal and gangue using image processing and a support vector machine. *International Journal of Coal Preparation and Utilization* 41: 603–616.
- [7] Zhang L, Sui Y, Wang H, et al. (2022) Image feature extraction and recognition model construction of coal and gangue based on image processing technology. *Sci Rep* 12: 20983.
- [8] Pu Y, Apel DB, Szmigiel A, et al. (2019) Image Recognition of Coal and Coal Gangue Using a Convolutional Neural Network and Transfer Learning. *Energies* 12: 1735.
- [9] Si L, Xiong X, Wang Z, et al. (2020) A Deep Convolutional Neural Network Model for Intelligent Discrimination between Coal and Rocks in Coal Mining Face. *Mathematical Problems in Engineering* 2020: 1–12.
- [10] Liu Q, Li J, Li Y, et al. (2021) Recognition Methods for Coal and Coal Gangue Based on Deep Learning. *IEEE Access* 9: 77599–77610.
- [11] Liu H, Xu K (2023) Recognition of gangues from color images using convolutional neural networks with attention mechanism. *Measurement* 206: 112273.
- [12] Cao Z, Fang L, Li R, et al. (2023) Research on image classification of coal and gangue based on a lightweight convolution neural network. *Energy Science & Engineering* 11: 3042–3054.
- [13] Pengcheng Y, Heng Z, Xuyue K, et al. (2024) Lightweight detection method of coal gangue based on multispectral and improved YOLOv5s. *International Journal of Coal Preparation and Utilization* 44: 399–414.
- [14] He K, Zhang X, Ren S, et al. (2016) Deep Residual Learning for Image Recognition, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, IEEE, 770–778.
- [15] Chollet F (2017) Xception: Deep Learning with Depthwise Separable Convolutions, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, IEEE, 1800–1807.
- [16] Xie S, Girshick R, Dollár P, et al. (2017) Aggregated Residual Transformations for Deep Neural Networks.

- [17] Tan M, Chen B, Pang R, et al. (2019) MnasNet: Platform-Aware Neural Architecture Search for Mobile.
- [18] Ding X, Zhang X, Ma N, et al. (2021) RepVGG: Making VGG-style ConvNets Great Again, 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, IEEE, 13728–13737.
- [19] Tang Y, Han K, Guo J, et al. (2022) GhostNetV2: Enhance Cheap Operation with Long-Range Attention.
- [20] Vasu PKA, Gabriel J, Zhu J, et al. (2023) MobileOne: An Improved One millisecond Mobile Backbone.
- [21] Vasu PKA, Gabriel J, Zhu J, et al. (2023) FastViT: A Fast Hybrid Vision Transformer using Structural Reparameterization.
- [22] Woo S, Debnath S, Hu R, et al. (2023) ConvNeXt V2: Co-designing and Scaling ConvNets with Masked Autoencoders.
- [23] Wang A, Chen H, Lin Z, et al. (2024) RepViT: Revisiting Mobile CNN From ViT Perspective.
- [24] Qin D, Leichner C, Delakis M, et al. (2024) MobileNetV4 -- Universal Models for the Mobile Ecosystem.
- [25] Yu W, Zhou P, Yan S, et al. (2025) InceptionNeXt: When Inception Meets ConvNeXt.
- [26] Xiang Q, Wang X, Lei L, et al. (2025) Dynamic bound adaptive gradient methods with belief in observed gradients. Pattern Recognition 168: 111819.

# Coupled Topology–Material Design Enables Manufacturable Lattice Metamaterials with Tailored Mechanics

Xinru Li, Yifei Zhou, Danyang Qiu, Ningping Zhan, Jianxiang Qiao, Yaoxin Huang\*

Faculty of Mechanical Engineering and Mechanics, Ningbo University, Ningbo 315211, China

\*Corresponding author: Yaoxin Huang, 226003965@nbu.edu.cn

**Copyright:** 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

**Abstract:** With the growing demand for lightweight and multifunctional integration in fields such as aerospace and impact protection, lattice metamaterials have attracted widespread attention due to their designable mechanical properties and excellent load-bearing and energy-dissipation characteristics. However, traditional homogeneous designs are limited by single topology-material combinations, making it difficult to overcome the performance trade-offs among stiffness, strength, and energy absorption, which constrains their further engineering application. This study focuses on the cutting-edge direction of “topology-material coupling design,” aiming to systematically elaborate the theoretical and methodological framework for cross-domain performance regulation of lattice metamaterials through spatially heterogeneous design. Unlike previous research that often emphasized single performance aspects or isolated processes, this work adopts a closed-loop “design-manufacturing-performance” perspective. It integrates collaborative strategies such as functional gradients, multi-material composites, and hybrid topologies, combined with additive manufacturing techniques like laser powder bed fusion and stereolithography, to reveal the regulating mechanisms of these designs on stiffness, strength, and energy absorption under both quasi-static and dynamic loading. The research demonstrates that topology-material coupling design not only significantly expands the tunable performance space of lattice metamaterials but also promotes a paradigm shift from “homogeneous configuration” to “functional customization.” Furthermore, this paper outlines future research directions, including intelligent inverse design and process-performance correlation modeling, providing a systematic theoretical reference and design framework to advance the practical application of lattice metamaterials in high-end equipment and protective structures.

**Keywords:** Lattice Metamaterials; Heterogeneous Topology Design; Spatial Heterostructure; Mechanical Performance

**Published:** Dec 30, 2025

**DOI:** <https://doi.org/10.62177/jaet.v2i4.976>

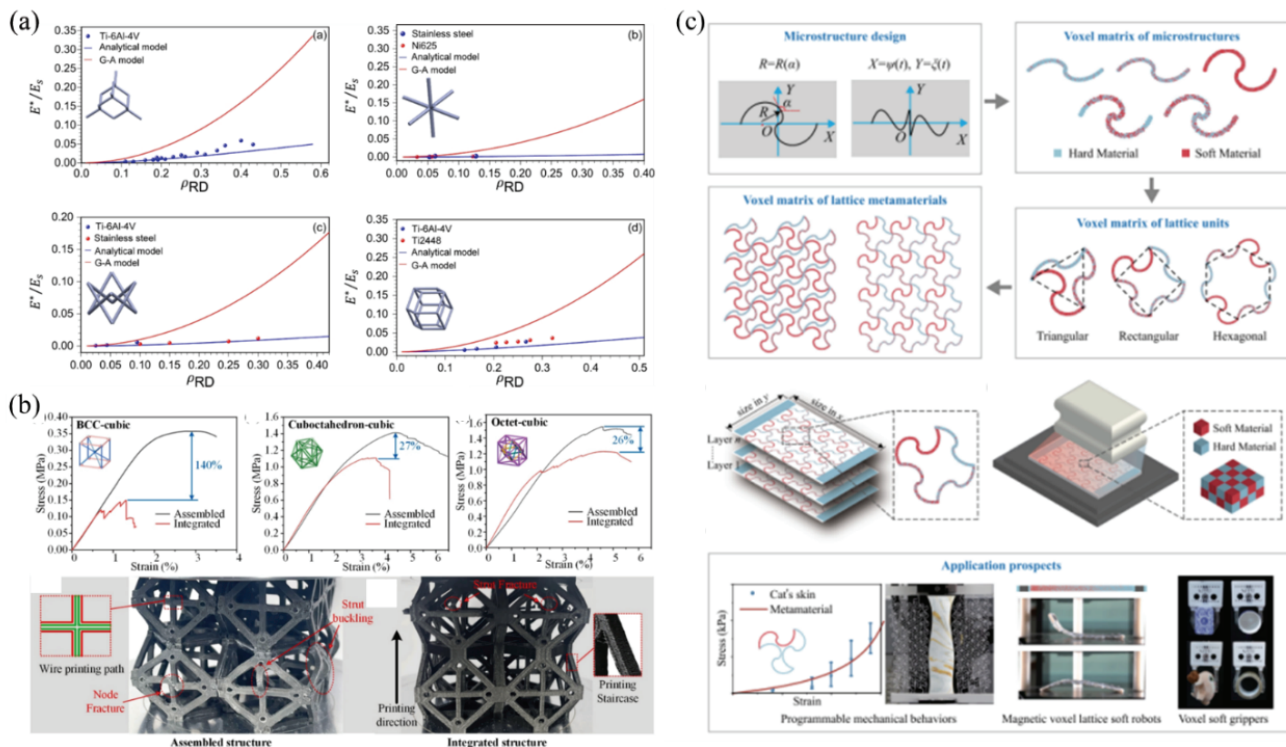
## 1.Introduction

Metamaterials are a class of engineered functional materials characterized by periodically arranged microstructural units. Their macroscopic mechanical properties depend more on the microstructural topology than on the inherent properties of the constituent material<sup>[1,2]</sup>. Such materials often combine high specific stiffness, high specific strength, and excellent energy absorption capacity. These properties make them promising for applications in lightweight aerospace structures, automotive crash-resistant systems, biomedical implants, and other fields<sup>[3,4]</sup>. However, as engineering demands on material performance grow increasingly stringent, the limitations of conventional uniform lattice designs, which typically employ a single topology and a single constituent material are becoming more apparent. Such designs are often constrained by their fixed deformation

modes, offering limited tunability in performance. Their inherent anisotropy makes them poorly suited for complex multiaxial loading scenarios, while the inherent trade-offs between strength and toughness, or between stiffness and large deformation capacity, remain difficult to reconcile<sup>[5]</sup>. Tension-dominated uniform lattices typically provide high stiffness, yet their energy absorption capacity is often limited. In contrast, bending-dominated structures exhibit favorable deformation capability, though they generally suffer from relatively low strength<sup>[6]</sup>.

To overcome these constraints, researchers are increasingly turning to the concept of “topology-material coupled design.” This approach advocates for the co-optimization of microstructural topology and spatial material distribution, thereby enabling programmable mechanical properties and multifunctional integration in lattice-based metamaterials<sup>[7]</sup>. Specifically, as shown in Fig. 1(b), such a coupled design allows for precise control over internal load transfer paths, enables controlled progressive deformation, and alleviates stress concentration. This approach thereby enhances structural strength, toughness, and energy absorption efficiency in a synergistic manner while maintaining lightweight characteristics<sup>[8]</sup>. It is worth noting that such complex designs with spatially heterogeneous characteristics heavily rely on the support of additive manufacturing technologies. Processes like laser powder bed fusion and stereolithography, illustrated in Fig. 1(c), overcome the geometric limitations of conventional subtractive manufacturing, enabling high-precision fabrication of predefined topologies and material distributions<sup>[9]</sup>. A systematic study of the interplay among topology-material coupling principles, additive manufacturing processes, and mechanical performance characterization is therefore essential to advance lattice-based metamaterials from fundamental research to engineering applications<sup>[10]</sup>. This paper focuses on how coupled design strategies regulate the key mechanical properties of lattice-based metamaterials, with emphasis on such core metrics as quasi-static stiffness and strength, dynamic energy absorption efficiency, and stress-wave manipulation. The review is structured around a closed “design-manufacturing-performance” loop. It systematically examines topology optimization and co-design strategies for metamaterials, provides an in-depth analysis of additive manufacturing technologies and their corresponding design principles, thoroughly characterizes the resulting mechanical performance with underlying mechanisms explained, and concludes with a summary of current research challenges and prospects for future development.

*Figure 1: Design framework and performance comparison of 3D voxel-printed lattice metamaterials. (a) Geometric design process, comprising voxel matrix generation for microstructures, unit cells, and lattice metamaterials. (b) Schematic of material distribution in a dual-material lattice, showing a coordinated layout of rigid and compliant constituents. (c) Additive manufacturing process routes, such as laser powder bed fusion and stereolithography.*

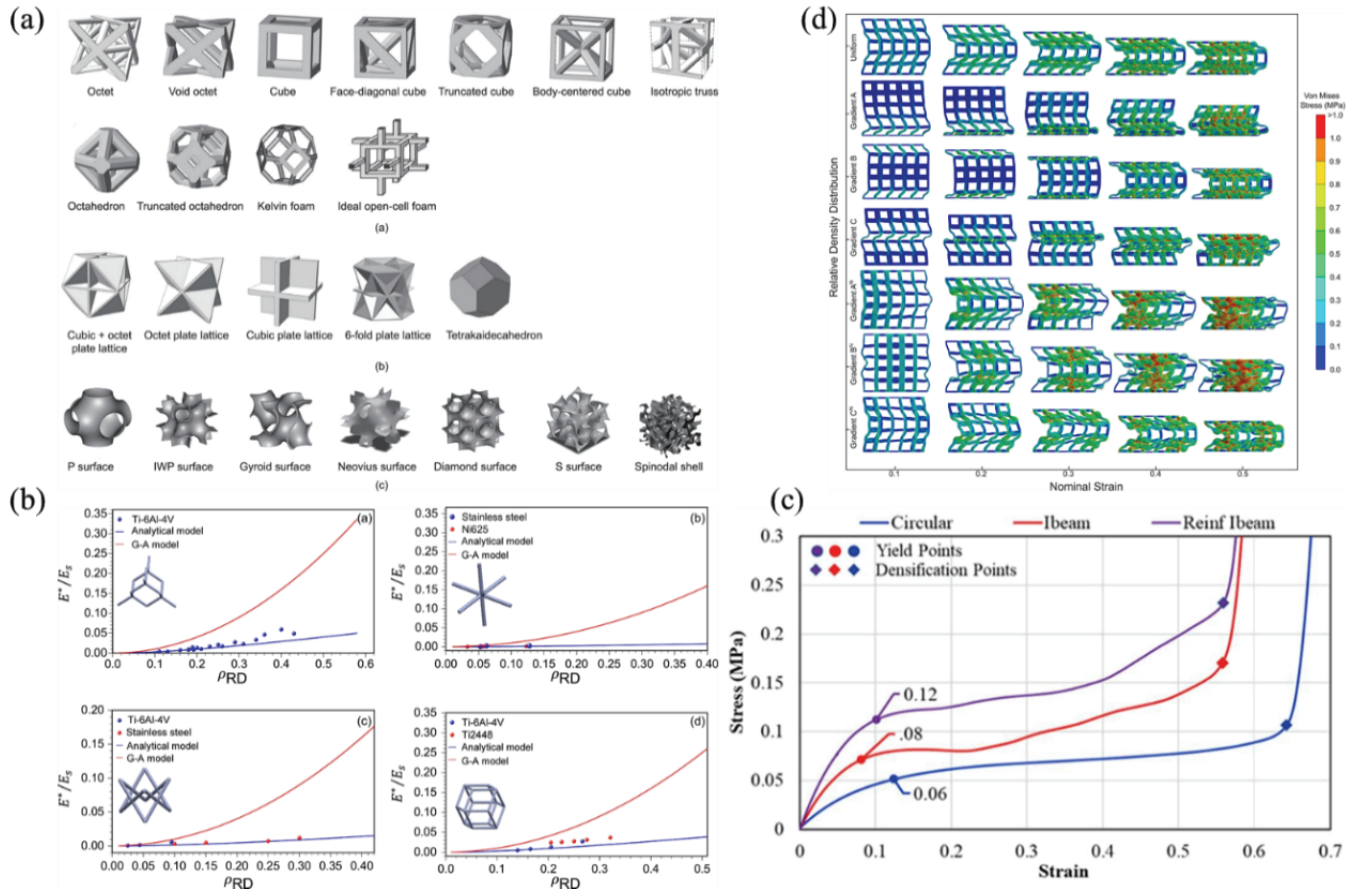


## 2. Integrated Design of Structural Mechanics and Topology Optimization for Lattice Metamaterials

### 2.1 Topology and Mechanical Behavior

The topology of a lattice metamaterial serves as the primary determinant of its macro-scale mechanical behavior<sup>[11]</sup>. As shown in Fig. 2(a), common topological types mainly include strut-based structures (such as body-centered cubic (BCC), face-centered cubic (FCC), and octet truss), plate-based structures (e.g., Kagome honeycombs), and structures based on triply periodic minimal surfaces (like Gyroid and Schwarz P)<sup>[12]</sup>. Based on the classical Maxwell criterion, these topologies can be broadly categorized into two types: stretch-dominated and bending-dominated structures. As illustrated in Fig. 2(b), struts in stretch-dominated structures (e.g., octet truss) primarily bear axial tensile or compressive loads, typically resulting in high stiffness and strength. In contrast, members in bending-dominated structures (e.g., BCC, Gyroid) undergo mainly bending deformation, which endows the structure with excellent energy absorption capacity and large deformation capability<sup>[6]</sup>. As shown in Fig. 2(c), under quasi-static compression, a BCC lattice typically undergoes progressive layer-by-layer collapse, while an Octet lattice exhibits a distinct linear elastic stage followed by a sharp yield plateau. These fundamental topological configurations provide a diverse “performance library” for subsequent coupled design strategies. By co-optimizing them with spatial material distribution, more complex and customized mechanical responses can be achieved<sup>[13]</sup>.

Figure 2: Topological configurations and mechanical response curves of lattice metamaterials. (a): Rod-based topologies (Octet, BCC, FCC, etc.), plate-based topologies (Cubic octet, Tetraikadehedron plate lattice, etc.), Topological Minimal Surface (TPMS) topologies (Gyroid, P surface, Diamond surface, etc.); (c) Quasi-static compressive stress-strain curves for different topological lattices (contrasting elastic stages and yield plateau differences between tension-dominant and bending-dominant structures); (d) Gradient-guided progressive propagation of plastic collapse.



## 2.2 Spatial Heterogeneity Design Strategies

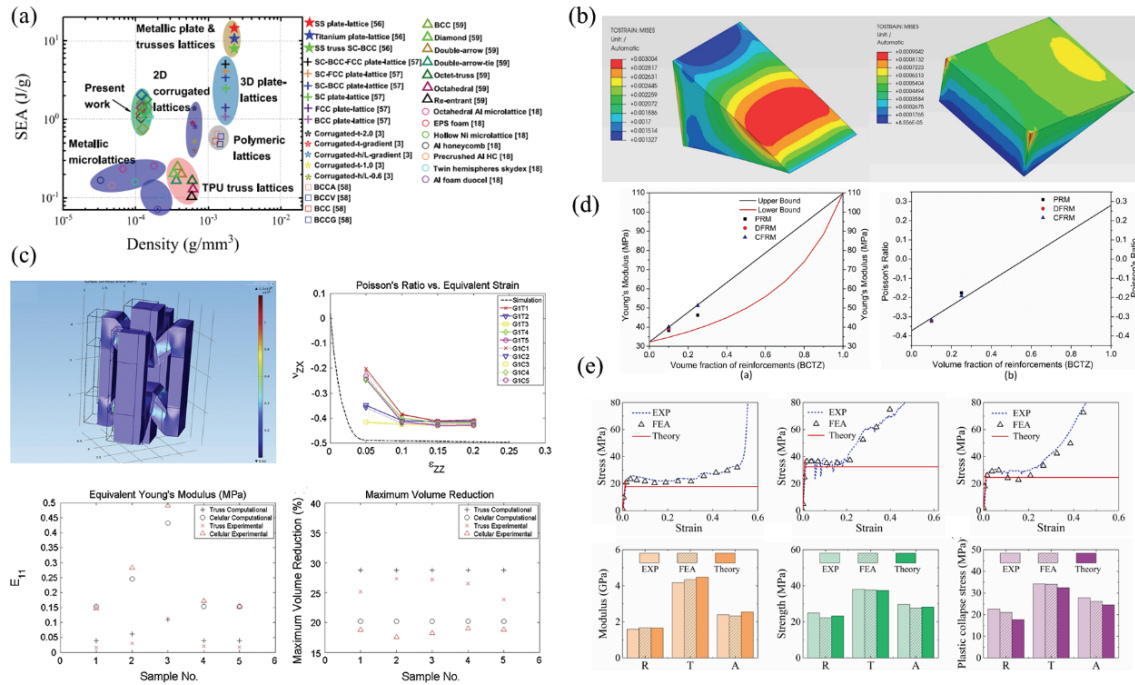
Spatial heterogeneity design serves as the central approach to coupling topology with material properties. By introducing deliberate variations in geometry and material composition <sup>[14,15]</sup>, this strategy enables precise performance control beyond what uniform structures can offer. The main design schemes can be categorized into three types: functionally graded design, multi-material composite design, and hybrid topology design <sup>[16]</sup>.

Functionally graded design optimizes the mechanical response of structures by allowing material properties (such as elastic modulus and yield strength) or topological parameters (e.g., strut diameter, wall thickness, relative density) to vary continuously along spatial coordinates <sup>[17,18]</sup>. For example, Niknam et al. <sup>[19]</sup> designed a lattice structure with relative density graded perpendicular to the loading direction. Compared to a uniform lattice, the graded design showed a 60% increase in stiffness and a 110% increase in energy absorption capacity. As demonstrated in Fig. 2(d), this enhancement is attributed to the progressive propagation of plastic collapse, guided by the density gradient from low- to high-density regions, which prevented global buckling and instability of the structure <sup>[19]</sup>. Fig. 3 (a) Zhang et al. <sup>[20]</sup> developed a graded structure by continuously scaling the unit size of a sinusoidal corrugated lattice. The resulting continuous interfaces allowed the structure to exhibit highly stable plateau stress during compression, achieving a specific energy absorption ranging from 0.77 to 1.85 J/g <sup>[20]</sup>. Moreover, functionally graded design can effectively mitigate stress concentration under complex loading conditions. For instance, as shown in Fig. 3(b), introducing a through-thickness density gradient in a bending-loaded beam results in a more uniform stress distribution <sup>[3]</sup>.

Multi-material composite design assigns materials with distinct mechanical properties to specific functional regions within the lattice unit, such as load-bearing struts versus deformable nodes, or structural cores versus surface claddings thereby synergistically combining their respective strengths <sup>[7]</sup>. In Fig. 3(c), Wang et al. <sup>[7]</sup> proposed a bi-material auxetic metamaterial design. By utilizing a stiff material for the primary load-bearing struts and a compliant material for the deformable connecting nodes, they successfully resolved the inherent trade-off between structural stiffness and auxetic deformation capability found in single-material designs. This configuration enables the structure to maintain a stable negative Poisson's ratio within a strain range of 5% to 20% <sup>[7]</sup>. Fig. 3 (d) Inspired by the reinforcement principle of composite materials, Han et al. <sup>[5]</sup> embedded a high-stiffness BCTZ lattice as the reinforcing phase into a 3D re-entrant auxetic lattice matrix. The resulting phase-reinforced metamaterial (PRM) structure showed an 81.3% improvement in energy absorption compared to the pure matrix, and its macroscopic mechanical properties can be effectively predicted using composite mixture rules <sup>[5]</sup>. The effectiveness of this strategy hinges on precisely matching material properties with structural functions. For example, high-toughness materials are placed at stress-concentrated joints to delay crack initiation, while high-strength materials are used for the primary load-bearing struts to enhance overall load capacity <sup>[21]</sup>.

Hybrid topology design can be regarded as a discrete "property-tailored material" placement strategy, in which two or more distinct topological unit cells are integrated into separate regions of the overall structure <sup>[22]</sup>. Xu et al. <sup>[22]</sup> developed a graded AuxHex honeycomb by using an auxetic hexagonal honeycomb as the main structure and replacing selected cell walls with substructures such as triangles and double arrows. As shown in Fig. 3(e), this design achieved a 45% to 180% increase in specific modulus compared to conventional honeycombs, primarily due to the synergistic deformation and energy dissipation between the primary and secondary structures <sup>[22]</sup>. Fig. 4 (a) Mizzi et al. <sup>[23]</sup> introduced a hierarchical triangular truss into rotating square units, achieving over 80% weight reduction while maintaining favorable auxetic behavior. The deformation mechanism shifts from joint stretching in solid structures to a combined effect of ligament bending and unit rotation in the truss system <sup>[23]</sup>. Hybrid topology design enables the tailoring of local properties to match specific loading demands in different regions of a structure. For example, stretch-dominated topologies can be employed in primary load-bearing areas to ensure stiffness and strength, while bending-dominated topologies can be used in energy-absorbing buffer zones to maximize energy dissipation <sup>[24]</sup>.

Figure 3: Mechanical performance characterization of spatially heterogeneous design strategies. (a) Specific energy absorption (SEA) of the corrugated 3D lattice compared to 2D/3D polymer/metal lattices reported in the literature. The structure in this study achieved an SEA range of 0.77–1.85 J/g at a lower density. (b) Equivalent Tresca stress distribution in a rotating triangular system, comparing stress concentration regions between solid and truss-based structures. (c) Poisson's ratio versus equivalent strain curves for multi-material composite lattices, illustrating the stability of Poisson's ratio under different material combinations. (d) Prediction of elastic modulus and Poisson's ratio for multi-phase lattices using composite mixture rules, showing agreement between theoretical and simulated results. (e) Enhancement in specific modulus achieved by the graded AuxHex honeycomb structure.



## 2.3 Comparison and Synthesis of Design Strategies

The three spatial heterogeneity design strategies discussed above couple topology with material properties in distinct ways<sup>[25]</sup>. They differ in their design variables, implementation methods, and target performance metrics, as summarized in the table below.

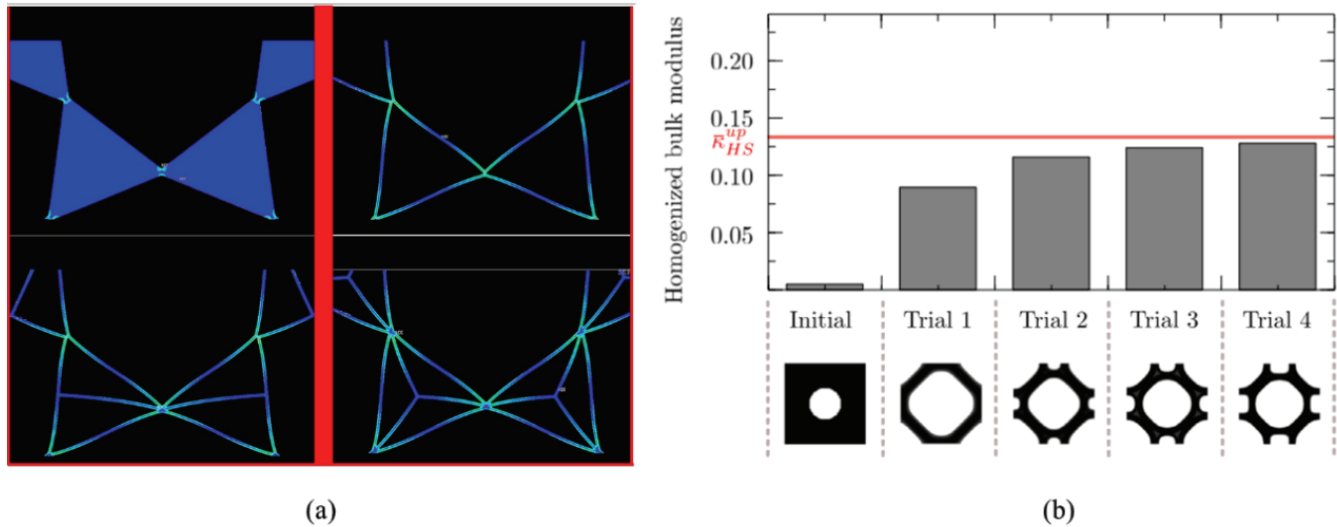
Table 1 Comparative Overview of Design Variables, Implementation Methods, and Target Performance Across Three Spatial Heterogeneity Design Strategies

Design Strategy	Core design variables	Means of implementation	Target performance	References
Functional Gradient Design	Relative density, rod diameter/wall thickness	Topological parameters vary continuously	Uniform stress distribution, controlled collapse sequence	[26,27]
Multi-material composite design	Material type, material distribution area	Dual-material/multi-material additive manufacturing	Stiffness-Toughness Synergy, Damage Tolerance Enhancement	[28,29]
Heterogeneous Topology Design	Topological type, partition boundary	Integration of Different Topological Unit Partitions	Multifunctional integration, localised performance customisation	[30,31]

A comparative analysis with Fig. 4 (b) reveals that the core advantages of coupled design over traditional homogeneous design can be summarized as three main aspects. First, it significantly expands design freedom in performance tuning, enabling the achievement of multifunctional integration that is difficult to attain with uniform structures<sup>[32]</sup>. Second, under equivalent mass constraints, it substantially optimizes the trade-offs between stiffness, strength, and energy absorption, thereby enhancing overall performance efficiency<sup>[33]</sup>. Third, it improves the structure's adaptability to complex loading conditions

and service environments, allowing the design to precisely meet specific engineering requirements<sup>[34]</sup>.

*Figure 4: Design strategy optimization iteration and performance comparison. (a) Lightweight design combining rotating square units with hierarchical triangular trusses while retaining auxetic behavior. (b) The optimization iteration process of homogenized bulk modulus across different trials (progressively approaching the theoretical limit from Initial to Trial 4), along with a comparison of performance boundaries between coupled design and traditional uniform lattices (highlighting the advantages of coupled design in stiffness–toughness and strength–energy absorption trade-offs).*



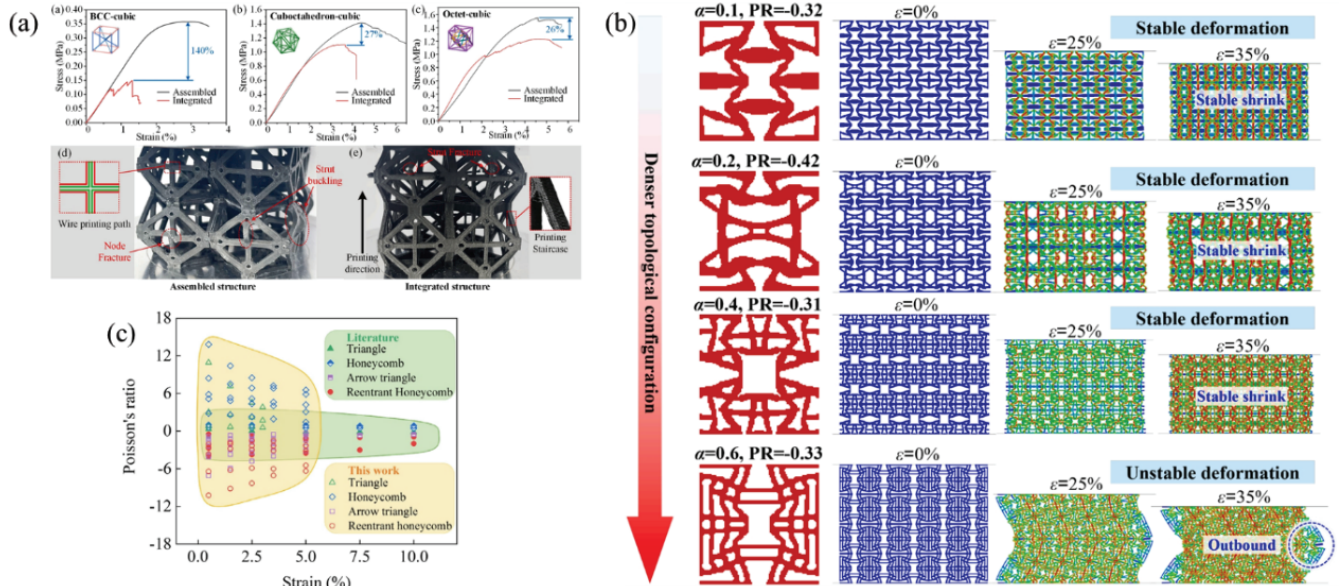
### 3. Additive Manufacturing-Enabled Fabrication Technologies for Lattice Metamaterials

#### 3.1 Essential Additive Manufacturing Technologies

Additive manufacturing enables the fabrication of complex geometries without the need for molds, making it essential for realizing lattice metamaterials based on topology-material coupled design<sup>[35]</sup>. To meet the requirements of different material systems and performance objectives, mainstream additive manufacturing technologies currently include: Laser powder bed fusion (LPBF), widely used for metal components; Stereolithography (SLA), suitable for high-performance polymers and high-precision complex structures; Fused deposition modeling (FDM) and material jetting (MJF), among others<sup>[36,37]</sup>. Based on the principle of photopolymerization, SLA technology offers extremely high fabrication accuracy and surface quality, making it particularly well-suited for producing lattice metamaterials with fine topological features. It is therefore recognized as one of the most effective methods for fabricating high-resolution, complex microstructures<sup>[38,39]</sup>.

Selective laser melting (SLM) technology uses a high-energy laser beam to selectively melt metal powder beds, enabling the high-precision fabrication of lattice structures made from high-strength metals such as titanium alloys (e.g., Ti6Al4V) and stainless steel (e.g., 316L). With a minimum feature size down to the 50  $\mu\text{m}$  scale, SLM is particularly well-suited for producing performance-critical lightweight components in aerospace and other fields<sup>[8]</sup>. For example, as shown in Fig. 5 (a), Gong et al.<sup>[8]</sup> fabricated a modular-assembled composite lattice structure using selective laser melting (SLM), achieving a strut diameter accuracy of approximately 0.7 mm. Compared to integrally printed uniform lattices, this structure exhibited a peak strength increase of 26% to 140%<sup>[8]</sup>. Stereolithography (SLA) technology operates on the principle of selective UV curing of photosensitive resin. It achieves higher fabrication accuracy, enabling feature sizes as fine as 20  $\mu\text{m}$ , which makes it particularly suitable for manufacturing polymer lattice structures with intricate features<sup>[40]</sup>. As shown in Fig. 5 (b), Han et al.<sup>[41]</sup> used stereolithography (SLA) to print nylon samples, successfully validating the effectiveness of a feature-controlled topology-optimized design in achieving a stable negative Poisson's ratio<sup>[41]</sup>. While fused deposition modeling (FDM) offers relatively lower fabrication accuracy, typically with feature sizes  $\geq 100 \mu\text{m}$ , it remains widely used for rapid validation of lattice design concepts due to its low equipment cost and broad material compatibility. For example, as illustrated in Fig. 5 (c), Ling et al.<sup>[42]</sup> fabricated nylon samples via FDM, achieving a broadly tunable Poisson's ratio ranging from  $-10.24$  to  $13.79$ <sup>[42]</sup>.

Figure 5: Fabrication outcomes and mechanical responses of different additive manufacturing processes. (a) Modular-assembled composite lattice structure fabricated via SLM, with a strut diameter accuracy of  $\approx 0.7$  mm and a peak strength increase of 26% to 140% compared to monolithic printing. (b) Stress–strain curve of a polymer lattice produced by SLA, demonstrating stable auxetic (negative Poisson's ratio) behavior. (c) Tunable range of Poisson's ratio ( $-10.24$  to  $13.79$ ) achieved in a nylon lattice printed using FDM.



### 3.2 Application in Metamaterial Design

The value of additive manufacturing extends beyond enabling the precise fabrication of complex topologies; it provides a critical pathway for realizing metamaterial designs. By supporting multi-material integration, graded structure fabrication, and heterogeneous unit assembly, this technology effectively bridges the gap between design concepts and physical prototypes in topology-material coupling, continuously expanding the performance limits and application potential of such materials [43].

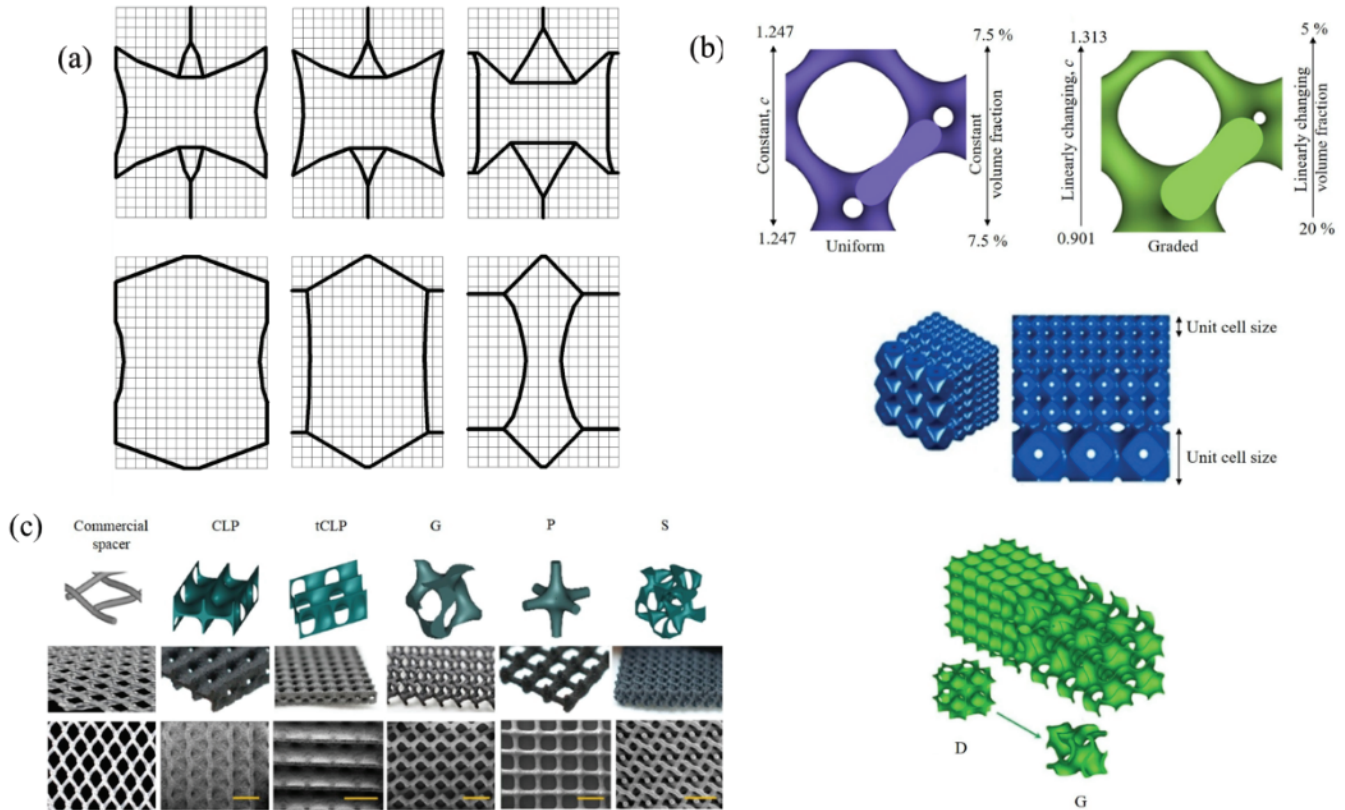
For multi-material and functional integration, additive manufacturing enables precise spatial arrangement and functional allocation of materials [44]. As shown in Fig. 6 (a), Wang et al. [7] successfully fabricated a bi-material auxetic metamaterial using PolyJet technology to simultaneously deposit rigid and compliant photopolymer resins. The resulting strong interfacial bonding ensures the structure meets the combined demands of load-bearing capacity and large deformation [45]. Technologies like HP Multi Jet Fusion go a step further, enabling the composite printing of polymers and metal powders [5]. This opens new possibilities for developing multi-material lattice structures that combine high stiffness with high toughness.

For producing graded structures, additive manufacturing facilitates continuous spatial variation in material properties or geometric features through real-time adjustment of process parameters [46]. In Fig. 6 (b), AlKetan et al. [47] employed selective laser melting (SLM) to fabricate a triply periodic minimal surface (TPMS) graded lattice with a gradual variation in relative density from 0.1 to 0.4 by adjusting the laser energy input. This design guided orderly deformation propagation and promoted uniform stress distribution. Similarly, the application of functionally graded design in sinusoidal corrugated lattices [19] and vertically densitygraded lattices [19] also enhanced both the stability and efficiency of energy absorption.

Furthermore, additive manufacturing is driving the advancement of metamaterials toward greater customization, intelligence, and scalability. In the field of intelligent response, 4D printing technology integrates shape-memory materials with lattice topologies, enabling structures to undergo predefined deformations under external stimuli. This offers a new pathway for applications such as adaptive cushioning and reconfigurable carriers [48]. Addressing the challenge of fabricating large-scale components, the “meta-assembly” strategy proposed by Chen et al. [6] presents a promising approach. As illustrated in Fig. 6(c), this strategy involves first fabricating standardized voxel units with distinct functions via additive manufacturing or CNC processes, which are then assembled into meter-scale composite lattices [49]. This method effectively overcomes the limitations of equipment build volume while preserving design freedom, providing a viable solution for large-scale architectural

elements and protective engineering structures<sup>[6]</sup>. Meanwhile, in the biomedical field, titanium alloy porous bone scaffolds fabricated via SLM, as referenced in Fig. 7(a), demonstrate pore structures and mechanical properties that can be well-matched to human bone tissue. This highlights the broad potential of this technology for manufacturing personalized implants<sup>[47]</sup>.

Figure 6: Fabrication of multi-material and graded structures with a meta-assembly strategy. (a) Microstructure of a bi-material auxetic metamaterial, featuring rigid struts for load-bearing connected by compliant nodes. (b) Relative density distribution in a functionally graded lattice, varying gradually from 0.1 to 0.4 to guide orderly deformation. (c) Schematic of the meta-assembly strategy, showing the fabrication of standardized voxel units and their assembly into a meter-scale composite lattice.

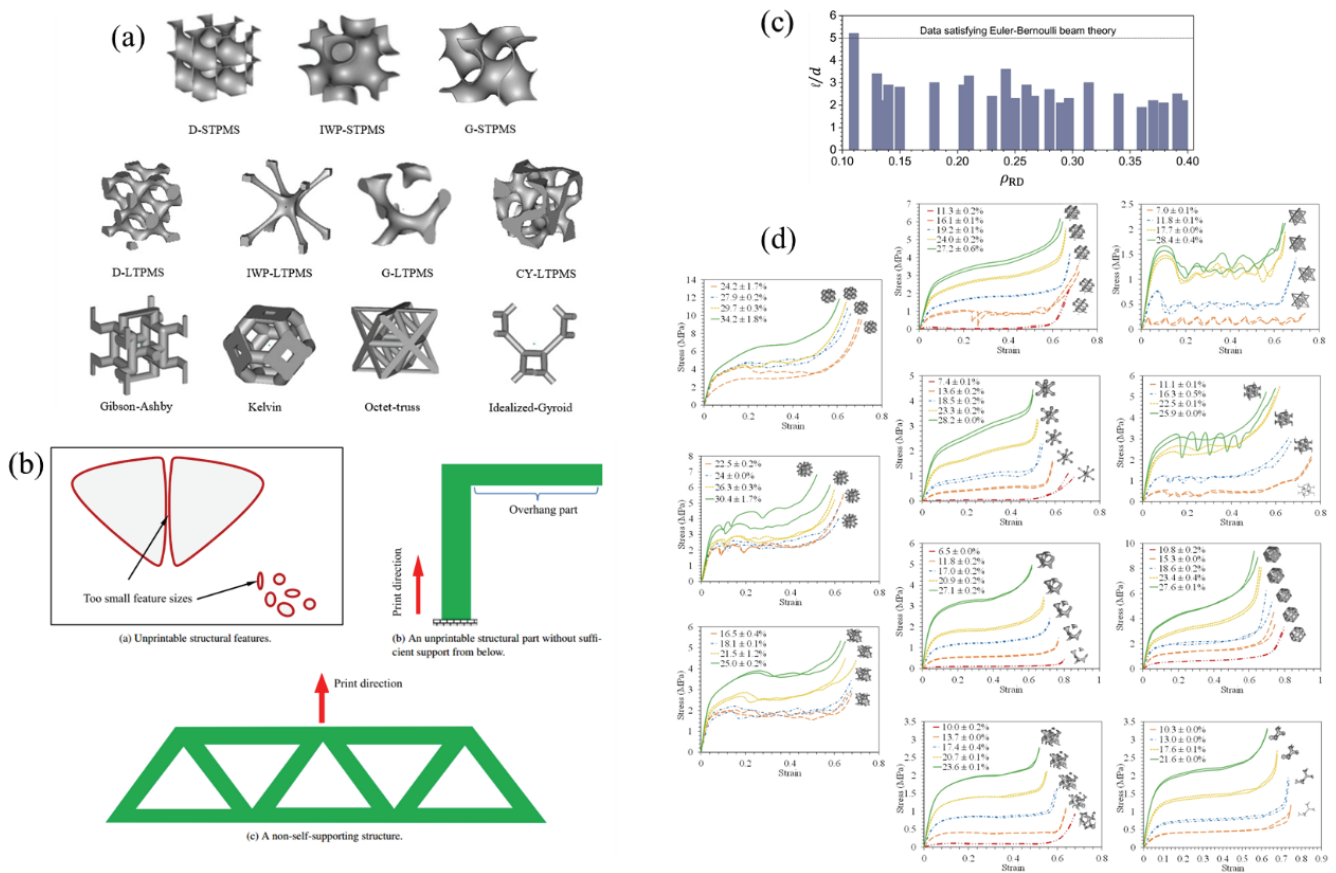


### 3.3 Design for Additive Manufacturing of Lattice Metamaterials

To realize coupled designs, it is essential to adopt a “design for additive manufacturing” (DFAM) approach from the outset, fully accounting for the constraints of the specific fabrication process and material characteristics<sup>[50]</sup>. This involves the appropriate selection of manufacturing technology and matching of feature sizes. The forming capability of each additive process directly dictates the minimum achievable strut diameter or wall thickness of a lattice, requiring the choice of technology to align with both material type and performance goals. For example, as shown in Fig. 7(b), selective laser melting (SLM) is commonly used for metal lattices, typically achieving a minimum strut diameter around 300  $\mu\text{m}$ . In contrast, stereolithography (SLA) for polymer lattices can produce features finer than 20  $\mu\text{m}$ . Designs must ensure geometric dimensions do not exceed these process limits<sup>[36]</sup>. Similarly, fused deposition modeling (FDM) is suitable for large-scale polymer lattices, while electron beam melting (EBM) can enhance the density of metal parts. These technical distinctions should be considered comprehensively throughout the design process<sup>[12]</sup>. Structurally, aligning the design with self-supporting requirements and process constraints is particularly critical. Large overhang angles in lattices often necessitate support structures or must be avoided through topology optimization<sup>[51]</sup>. For instance, Fig. 7(c) demonstrates a self-supporting design method based on moving morphable components, which automatically satisfies overhang-angle constraints and is compatible with processes like SLM and SLA<sup>[52]</sup>. For complex curved lattice structures such as those based on TPMS, curved layer manufacturing (CLM) can be employed to fabricate directly onto 3D surfaces, reducing reliance on supports and improving accuracy<sup>[53]</sup>. Moreover, designs must proactively compensate for manufacturing defects. Additive processes inevitably introduce imper-

fections such as porosity, lack-of-fusion, and poor interlayer bonding, with the type and severity of these defects varying across different technologies<sup>[54]</sup>. For instance, as illustrated in Fig. 7(d), metal lattices fabricated by SLM typically exhibit a porosity of 1–5%, which can reduce the actual strength by 10–20% compared to theoretical predictions<sup>[6]</sup>. Polymer lattices made via FDM are often limited by weak interlayer bonding<sup>[12]</sup>. Targeted measures can be incorporated during the design phase to preemptively mitigate such performance losses, for example, increasing the cross-sectional dimensions of critical load-bearing struts, optimizing the topology to alleviate stress concentration, or employing graded designs to distribute the influence of defects<sup>[36]</sup>. In one example, shown in Fig. 8(a), increasing the wall thickness of a TPMS lattice during design helps minimize the adverse effect of residual unmelted powder<sup>[36]</sup>. For FDM, adjusting the printing orientation and layer thickness can enhance interlayer bonding quality<sup>[55]</sup>. Manufacturing efficiency and reproducibility must also be considered. For large-scale lattices, a modular assembly strategy, where standardized voxel units are printed first and then assembled can overcome equipment build-volume limits and improve overall efficiency<sup>[6]</sup>. Furthermore, the design should clearly define the relationship between critical geometric tolerances and key process parameters (such as laser power and scan speed in SLM) to ensure consistency and stability in batch production<sup>[56]</sup>.

*Figure 7: Design constraints and solutions in additive manufacturing. (a) Matching the properties of titanium alloy TPMS lattice bone scaffolds with human bone tissue. (b) Comparison of minimum feature sizes across processes: metal SLM lattices typically reach about 300  $\mu\text{m}$ , polymer SLA can achieve  $\leq 20 \mu\text{m}$ , and FDM is generally  $\geq 100 \mu\text{m}$ . (c) A self-supporting design solution for overhang-angle constraints using topology optimization based on moving morphable components. (d) Stress–strain responses of different lattice topologies from two repeat specimens measured under varying relative densities.*



## 4. Mechanical Characterization and Mechanisms of Lattice Metamaterials

### 4.1 Quasi-Static Compressive Response

Quasi-static compression testing serves as a fundamental method for characterizing the basic mechanical properties of lattice metamaterials. By adjusting the stiffness distribution, strength, and deformation-failure modes of the structure, topology–material coupled designs achieve significant performance improvements<sup>[57]</sup>.

In terms of stiffness and strength regulation, functionally graded design optimizes the spatial distribution of performance. For example, the sinusoidally corrugated graded lattice developed by Zhang et al.<sup>[20]</sup> (Fig. 8(b)) exhibits a 25.85 % increase in compressive modulus compared to a uniform lattice, achieved through continuous interface design<sup>[20]</sup>. Multi-material composite design enables the simultaneous enhancement of stiffness and toughness. As shown in Fig. 8(c), the bi-material lattice reported by Wang et al.<sup>[7]</sup> maintains a stable negative Poisson's ratio while allowing its equivalent Young's modulus to be independently controlled by adjusting the stiffness of the compliant phase<sup>[7]</sup>. Hybrid topology design facilitates the customization of local performance. For instance, the graded AuxHex honeycomb by Xu et al.<sup>[22]</sup> (Fig. 8(d)) shows a 54 % increase in specific strength in its T-AuxHex configuration compared to conventional honeycombs<sup>[22]</sup>.

Controlling deformation and failure modes represents another key advantage of coupled designs<sup>[58]</sup>. Under compression, uniform BCC lattices tend to form localized shear bands, leading to abrupt failure. In contrast, graded designs can guide the progressive propagation of plastic collapse along the gradient direction. For example, in the Gradient3 structure (Fig. 8(e)), deformation is successfully confined to specific interfacial subspaces, effectively avoiding global buckling<sup>[8]</sup>. Multi-material designs steer deformation through material distribution: in the DMAMs lattice (Fig. 9(a)), deformation is mainly limited to the compliant connecting regions, while the rigid struts remain stable, preventing premature buckling<sup>[7]</sup>. Hybrid topology designs achieve progressive failure through synergistic deformation between primary and secondary structures. As illustrated in Fig. 9(b), the substructure walls in the graded honeycomb undergo local fracture and twisting, thereby significantly improving energy absorption efficiency<sup>[22]</sup>.

Fig. 7 Design constraints and solutions in additive manufacturing. (a) Matching the properties of titanium alloy TPMS lattice bone scaffolds with human bone tissue. (b) Comparison of minimum feature sizes across processes: metal SLM lattices typically reach about 300  $\mu\text{m}$ , polymer SLA can achieve  $\leq 20 \mu\text{m}$ , and FDM is generally  $\geq 100 \mu\text{m}$ . (c) A self-supporting design solution for overhang-angle constraints using topology optimization based on moving morphable components. (d) Stress-strain responses of different lattice topologies from two repeat specimens measured under varying relative densities.

Figure 8: Stiffness control and deformation patterns under quasi-static compression: (a)-(b) Compare stress-strain curves and deformation processes of lattices at different scaling factors, illustrating scaling effects on mechanical curve trends and structural deformation patterns (e.g., negative Poisson's ratio behaviour); (c) Displacement contour plots of metamaterial at  $n = +1.0$ ; (d) Mechanical responses of regular hexagonal honeycomb and graded hexagonal honeycomb under uniaxial compression, derived from experiments, numerical simulations, and theoretical models; (e) Deformation sequence diagrams for normal lattice, gradient 1 lattice, gradient 2 lattice, and gradient 3 lattice.

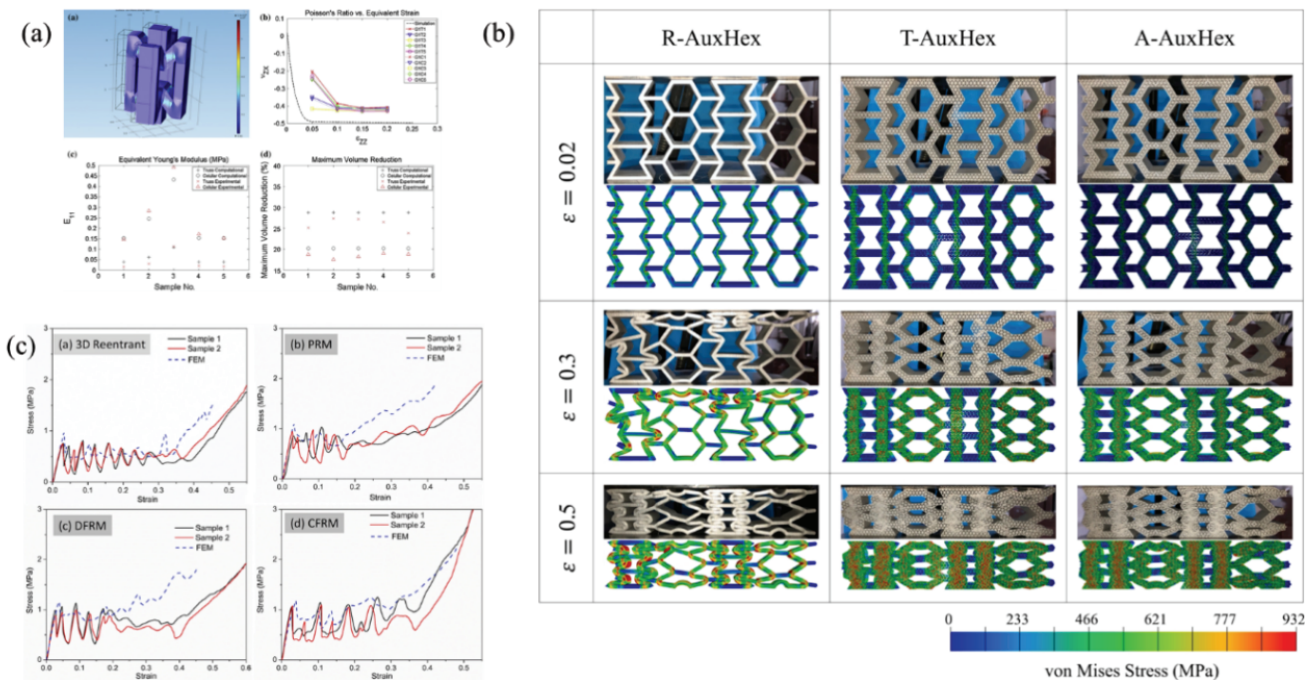
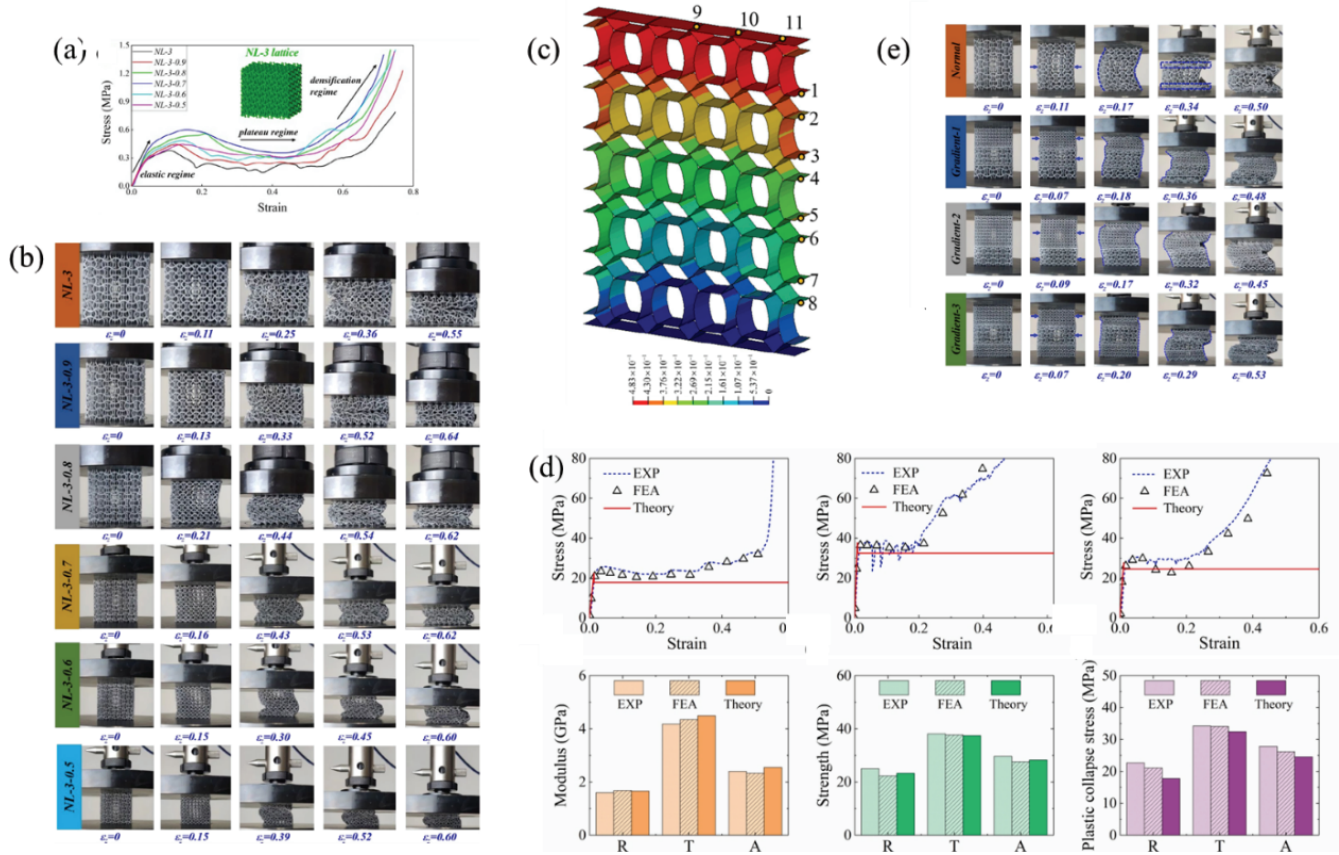


Figure 9: Mechanisms of deformation and failure mode control. (a) Deformation localization in a DMAMs lattice, where compliant connecting regions dominate the deformation while rigid struts remain stable. (b) Progressive failure mode in a graded honeycomb, featuring local fracture and twisting of substructure walls to enhance energy absorption. (c) Comparison of compressive stress–strain curves between a pure 3D re-entrant lattice and three multi-phase lattice structures (PRM, DFRM, CFRM), showing both experimental and simulated results.



## 4.2 Dynamic Response and Energy Absorption Behavior

The dynamic response and energy absorption behavior of lattice metamaterials are key metrics for their application in impact protection<sup>[59]</sup>. Coupled designs provide the possibility of tailoring the energy absorption curve.

Functionally graded designs effectively adjust the plateau stress level through density gradients<sup>[58]</sup>. For instance, the graded lattice developed by Niknam et al.<sup>[19]</sup> exhibits a 60% higher plateau stress in the medium-to-high strain range and a 110% increase in energy absorption capacity compared to its uniform counterpart<sup>[19]</sup>. Multi-material composite designs optimize energy dissipation mechanisms. As seen in the PRM multi-phase lattice, the synergistic interaction between plastic deformation of the reinforcing phase and auxetic deformation of the matrix improves specific energy absorption by 81.3% relative to the pure matrix<sup>[5]</sup>. Hybrid topology designs enable stable energy absorption over a wide strain range. For example, the Dprime TPMS lattice achieves twice the specific energy absorption of a conventional BCC lattice at a relative density of only 12%<sup>[24]</sup>. Metamaterials based on coupled design also demonstrate unique advantages in stress-wave manipulation. The flower-shaped lattice designed by Sarracino et al.<sup>[60]</sup> localizes impact energy within the struck petal cluster through specific topological features, reducing the transmitted peak pressure to the back face by two orders of magnitude compared to a solid structure<sup>[60]</sup>. Under dynamic impact, the sinusoidally corrugated biphasic lattice exhibits significant strain-rate strengthening, with dynamic strength increasing by 183% over quasi-static conditions. Moreover, the arrangement of the reinforcing phase can control the propagation path of shear bands<sup>[60]</sup>. These characteristics give coupled-design lattices considerable application potential in fields such as vibration damping, shock isolation, and blast protection.

## 4.3 Comparative Analysis and Validation

Extensive experimental and numerical results confirm that lattice metamaterials based on topology–material coupled design

consistently outperform conventional uniform lattices in overall mechanical performance<sup>[61]</sup>.

In terms of quasistatic compressive performance, the modularassembled lattice by Gong et al.<sup>[8]</sup> achieves a 26–140 % increase in peak strength and a 30–510 % improvement in energy absorption capacity compared to its monolithicprinted uniform counterpart<sup>[8]</sup>. Under dynamic impact, the sinusoidally corrugated biphasic lattice reported by Wang et al.<sup>[60]</sup> exhibits a 74 % higher strength than a uniform lattice at a strain rate of  $120 \text{ s}^{-1}$ <sup>[60]</sup>. Regarding stiffness, the multimaterial lattice obtained through topology optimization by Yang et al.<sup>[9]</sup> shows a 47–185 % enhancement in specific stiffness relative to a uniform design<sup>[9]</sup>.

The underlying mechanisms behind these performance gains can be revealed through finiteelement simulations. Coupled designs promote a more uniform internal stress distribution<sup>[62]</sup>. For instance, in multimaterial lattices, the rigid reinforcing phase carries a larger share of the load, while the compliant matrix effectively alleviates stress concentration<sup>[5]</sup>. Graded designs guide deformation in a more coordinated manner, preventing early failure caused by localized overdeformation<sup>[20]</sup>. Hybrid topology designs, through the synergistic interaction of different unit cells, enable multipath and multimode energy dissipation<sup>[22]</sup>. This deepened understanding of the structure–performance relationship provides a solid theoretical foundation for further optimization of coupled designs in the future.

## 5. Conclusions and Future Perspectives

This review summarizes the recent advances in the fabrication and mechanical performance of lattice metamaterials based on topology–material coupled design. The main findings are outlined below:

- (1) Spatial heterogeneity design serves as the central approach to coupling topology with material properties<sup>[63]</sup>. Strategies such as functional grading, multi-material composition, and hybrid topology have substantially expanded both the performance tunability and design freedom of lattice metamaterials by enabling continuous parameter variation, targeted material-function matching, and zoned integration of distinct topological units.
- (2) Additive manufacturing is the essential enabling technology for fabricating such complex designs<sup>[64]</sup>. Processes such as SLM, SLA, and FDM are capable of producing lattice structures with varying material compatibility and accuracy needs. However, their successful application must follow design-for-manufacturing principles to achieve the desired structural integrity and performance.
- (3) By precisely tailoring stiffness distribution, guiding controlled deformation patterns, and optimizing energy dissipation mechanisms, coupled designs significantly enhance the overall mechanical properties of lattice metamaterials under both quasi-static and dynamic loading<sup>[49]</sup>, consistently outperforming traditional uniform designs.

In summary, topology–material coupled design effectively overcomes the performance limitations of conventional uniform lattices, driving lattice metamaterials from “singleproperty optimization” toward “multifunctional integration and customization.” Establishing a seamless “design–manufacturing–performance” closed loop is the essential framework for advancing such materials toward practical engineering applications.

Although significant progress has been made, several key challenges remain. Future research could focus on the following directions:

**Integration of machine learning and topology optimization.** Combining algorithms such as generative adversarial networks and reinforcement learning with topology optimization theory to build efficient, automated frameworks that map target performance requirements to optimal topology–material distributions<sup>[65]</sup>. For example, as illustrated in Fig. 10(a), Garland et al.<sup>[52]</sup> integrated a convolutional neural network with a genetic algorithm and required only about 3,500 numerical simulations to identify highperformance lattice designs along the Pareto front<sup>[52]</sup>.

**Defectaware modeling and performance prediction.** Delving deeper into the formation mechanisms and distribution patterns of defects introduced during additive manufacturing (e.g., porosity, lackoffusion, residual stress). Developing macroscopic mechanicalproperty prediction models that account for these defects will enable quantitative and precise optimization of coupled designs. For instance, Zhong et al.<sup>[66]</sup> proposed a modified Gibson–Ashby model that improves the accuracy of strength predictions for additively manufactured lattices<sup>[6]</sup>(see Fig. 10(b)).

Figure 10: Intelligent design and model refinement. (a) Pareto front optimization results driven by machine learning, showing the performance boundary evolution from Gen 1 to Gen 11 compared with intuitive design. (b) Comparison between the modified Gibson-Ashby (G-A) model and experimental results, demonstrating improved accuracy in predicting the strength of additively manufactured lattices.

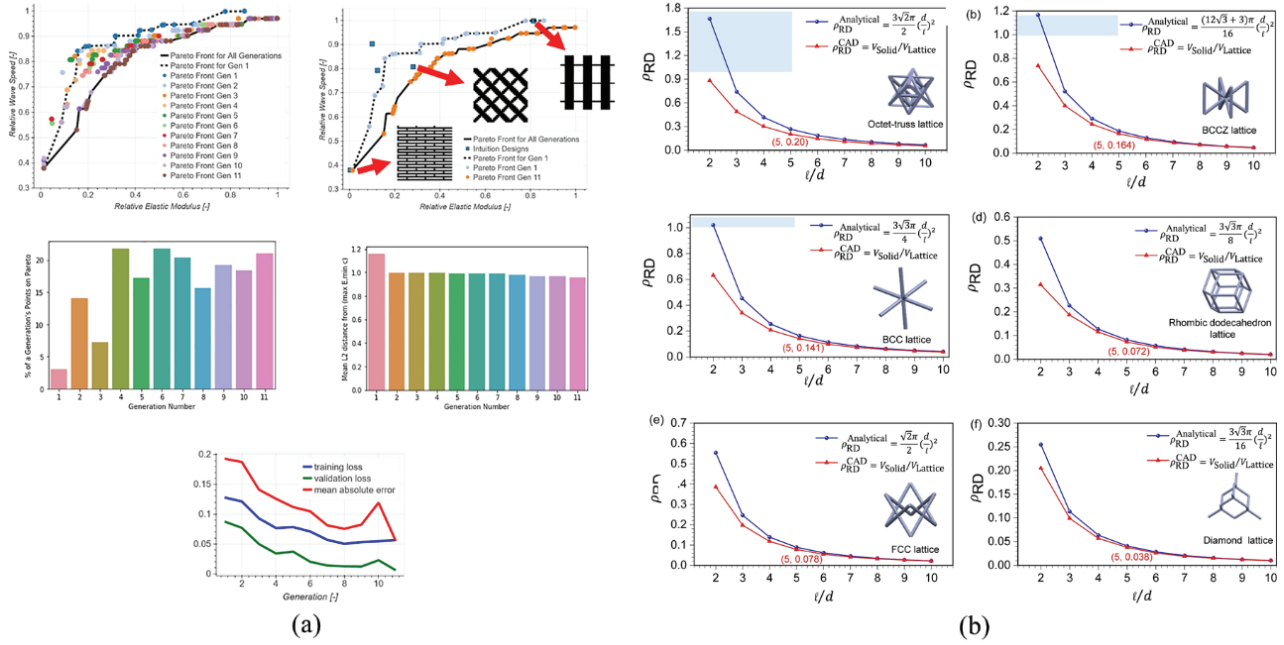
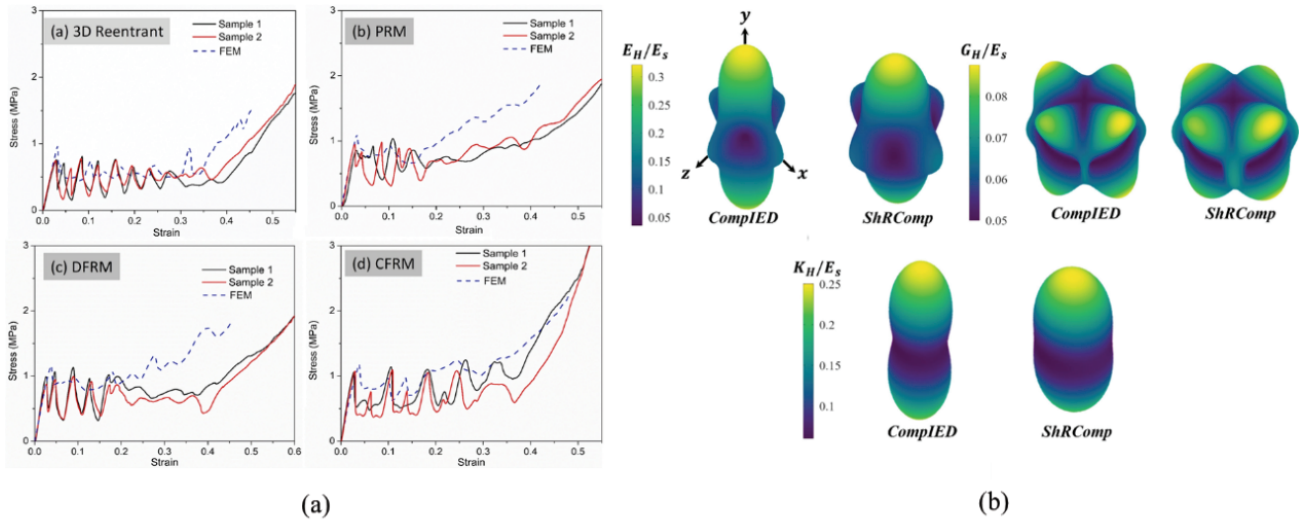


Figure 11: Response under complex loading conditions and multi-scale design. (a) Compressive stress-strain curves for 3D concave lattice, PRM, DFRM and CFRM structures; (b) Distribution of elastic anisotropy in CompIED and ShRComp topological structures at a relative density of 0.4, showing: uniaxial modulus, shear modulus and bulk modulus.



Current research predominantly focuses on uniaxial quasi-static or dynamic impact loading. Future work should intensify the study of lattice metamaterial performance under complex service environments, such as multi-axial loading, cyclic fatigue, thermo-mechanical coupling, and hygro-mechanical coupling. For instance, a deeper investigation into the interfacial failure dynamics of multi-material lattices under dynamic impact, as suggested by Fig. 11(a), would be valuable<sup>[5]</sup> Integrating microstructural material characteristics (e.g., grain orientation, composite interfacial properties) into the macro-scale topology and material distribution design framework can achieve multiscale co-design and performance regulation from the micro to the macro level. This approach holds promise for discovering new pathways for performance enhancement, as illustrated in Fig. 11(b)<sup>[21]</sup>. As design methods become more intelligent, manufacturing processes more precise, and

performance characterization systems more robust, lattice metamaterials based on topology–material coupled design are poised for broader and deeper engineering applications in fields such as aerospace, high-end equipment, biomedical devices, and protective systems<sup>[67]</sup>. They are likely to become a pivotal direction guiding the development of next-generation advanced structural materials.

## Funding

No

## Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Reference

- [1] Yan, H., Zhang, Y., Teng, X., Jiang, W., Xie, Y., & Wu, W. (2024). A study on surface reconfigurability of future architecture based on lattice metamaterials. *Composites Structures*, 337. <https://doi.org/10.1016/j.compstruct.2024.118075>
- [2] Zhang, Y., Guo, X., Wu, Y., Zhang, Y., & Lü, C. (2025). Active control of cables with piezoelectric actuation considering geometric and material nonlinearities. *Engineering Structures*, 340, 120773. <https://doi.org/10.1016/j.engstruct.2025.120773>
- [3] Cibrario, L., Gastaldi, C., Delprete, C., & Cozza, I. (2025). Mechanical design of lattice metamaterials: A multiscale homogenization-based operational procedure. *Materials & Design*, 251. <https://doi.org/10.1016/j.matdes.2025.113614>
- [4] Zhang, Y., Huang, Y., Zhao, S., Jiao, Z., Lü, C., & Yang, J. (2025). Nonlinear dynamic response and stability of piezoelectric shells with piezoelectric nonlinearities. *International Journal of Mechanical Sciences*, 304, 110731. <https://doi.org/10.1016/j.ijmecsci.2025.110731>
- [5] Sahariah, B., Namdeo, A., & Khanikar, P. (2022). Composite-inspired multilattice metamaterial structure: An auxetic lattice design with improved strength and energy absorption. *Materials Today Communications*, 30, 103159. <https://doi.org/10.1016/j.mtcomm.2022.103159>
- [6] Chen, Y., Wang, Z., & Ma, L. (2023). Super-assembly strategy based on discretization design for composite lattice metamaterials. *Construction and Building Materials*, 409, 133955. <https://doi.org/10.1016/j.conbuildmat.2023.133955>
- [7] Wang, K., Chang, Y., Chen, Y., Zhang, C., & Wang, B. (2015). Designable dual-material auxetic metamaterials using three-dimensional printing. *Materials & Design*, 67, 159–164. <https://doi.org/10.1016/j.matdes.2014.11.033>
- [8] Gong, C., Ritchie, R., Wei, X., Liu, Q., & Xiong, J. (2025). Mechanical properties of modular assembled composite lattice architecture. *Journal of the Mechanics and Physics of Solids*, 195. <https://doi.org/10.1016/j.jmps.2024.105967>
- [9] Dong, L., Wang, J., & Wang, D. (2023). Modeling and design of three-dimensional voxel printed lattice metamaterials. *Additive Manufacturing*, 69. <https://doi.org/10.1016/j.addma.2023.103532>
- [10] Abali, B., & Barchiesi, E. (2021). Additive manufacturing introduced substructure and computational determination by means of the asymptotic homogenization. *Continuum Mechanics and Thermodynamics*, 33, 993–1009. <https://doi.org/10.1007/s00161-020-00941-w>
- [11] Kobir, M., Liu, X., & Yang, Y. (2023). Additive manufacturing of body-centered cubic metamaterials with novel I-shaped beam lattice towards enhanced mechanical properties. *Manufacturing Letters*, 35, 509–515. <https://doi.org/10.1016/j.mfglet.2023.08.004>
- [12] Zhu, J., Zhou, H., Wang, C., Zhou, L., Yuan, S., & Zhang, W. (2021). A review of topology optimization for additive manufacturing: Status and challenges. *Chinese Journal of Aeronautics*, 34, 91–110. <https://doi.org/10.1016/j.cja.2020.09.020>
- [13] Ochoa, O., Cuan-Urquiza, E., Alvarez-Trejo, A., Roman-Flores, A., & Silva, R. (2024). Additively manufactured auxetic arc-based architected metamaterial: Mechanical properties and their directional dependency. *Mechanics of Advanced Materials and Structures*, 31, 6450–6463. <https://doi.org/10.1080/15376494.2023.2231453>
- [14] Zhang, Y., Zhu, H., Zhao, S., Ni, Z., Lü, C., & Yang, J. (2025). Nonlinear dynamic response of functionally graded plates with piezoelectric nonlinearity. *European Journal of Mechanics - A/Solids*, 114, 105776. <https://doi.org/10.1016/j.euromechsol.2025.105776>

- [15] Zhang, Y., Guo, X., Wu, Y., Zhang, Y., Zhang, H., & Lü, C. (2024). Nonlinear thermo-electro-mechanical responses and active control of functionally graded piezoelectric plates subjected to strong electric fields. *Thin-Walled Structures*, 205, 112375. <https://doi.org/10.1016/j.tws.2024.112375>
- [16] Mukhopadhyay, T., Naskar, S., & Adhikari, S. (2020). Anisotropy tailoring in geometrically isotropic multi-material lattices. *Extreme Mechanics Letters*, 40, 100934. <https://doi.org/10.1016/j.eml.2020.100934>
- [17] Zhang, P., Qi, D., Xue, R., Liu, K., Wu, W., & Li, Y. (2021). Mechanical design and energy absorption performances of rational gradient lattice metamaterials. *Composites Structures*, 277, 114606. <https://doi.org/10.1016/j.comp-struct.2021.114606>
- [18] Zhang, Y., Guo, X., Wu, Y., Zhang, Y., Zhang, H., & Lü, C. (2024). Vibration control of membrane structures by piezoelectric actuators considering piezoelectric nonlinearity under strong electric fields. *Engineering Structures*, 315, 118413. <https://doi.org/10.1016/j.engstruct.2024.118413>
- [19] Niknam, H., & Akbarzadeh, A. (2020). Graded lattice structures: Simultaneous enhancement in stiffness and energy absorption. *Materials & Design*, 196. <https://doi.org/10.1016/j.matdes.2020.109129>
- [20] Zhang, H., Lin, G., & Sun, W. (2023). Structural design and tunable mechanical properties of novel corrugated 3D lattice metamaterials by geometric tailoring. *Thin-Walled Structures*, 184. <https://doi.org/10.1016/j.tws.2022.110495>
- [21] Song, J., Wang, Y., Zhou, W., Fan, R., Yu, B., Lu, Y., et al. (2019). Topology optimization-guided lattice composites and their mechanical characterizations. *Composites Part B: Engineering*, 160, 402–411. <https://doi.org/10.1016/j.compositesb.2018.12.027>
- [22] Xu, M., Zhao, Z., Wang, P., Duan, S., Lei, H., & Fang, D. (2022). Mechanical performance of bio-inspired hierarchical honeycomb metamaterials. *International Journal of Solids and Structures*, 254. <https://doi.org/10.1016/j.ijsol-str.2022.111866>
- [23] Mizzi, L., & Spaggiari, A. (2020). Lightweight mechanical metamaterials designed using hierarchical truss elements. *Smart Materials and Structures*, 29. <https://doi.org/10.1088/1361-665X/aba53c>
- [24] Xiao, L., Shi, G., Feng, G., Li, S., Liu, S., & Song, W. (2024). Large deformation response of a novel triply periodic minimal surface skeletal-based lattice metamaterial with high stiffness and energy absorption. *International Journal of Solids and Structures*, 296. <https://doi.org/10.1016/j.ijsolstr.2024.112830>
- [25] Lu, C., Hsieh, M., Huang, Z., Zhang, C., Lin, Y., Shen, Q., et al. (2022). Architectural design and additive manufacturing of mechanical metamaterials: A review. *Engineering*, 17, 44–63. <https://doi.org/10.1016/j.eng.2021.12.023>
- [26] Zhang, P., Qi, D., Xue, R., Liu, K., Wu, W., & Li, Y. (2021). Mechanical design and energy absorption performances of rational gradient lattice metamaterials. *Composites Structures*, 277, 114606. <https://doi.org/10.1016/j.comp-struct.2021.114606>
- [27] Niknam, H., & Akbarzadeh, A. H. (2020). Graded lattice structures: Simultaneous enhancement in stiffness and energy absorption. *Materials & Design*, 196, 109129. <https://doi.org/10.1016/j.matdes.2020.109129>
- [28] Wang, K., Chang, Y.-H., Chen, Y., Zhang, C., & Wang, B. (2015). Designable dual-material auxetic metamaterials using three-dimensional printing. *Materials & Design*, 67, 159–164. <https://doi.org/10.1016/j.matdes.2014.11.033>
- [29] Sahariah, B. J., Namdeo, A., & Khanikar, P. (2022). Composite-inspired multilattice metamaterial structure: An auxetic lattice design with improved strength and energy absorption. *Materials Today Communications*, 30, 103159. <https://doi.org/10.1016/j.mtcomm.2022.103159>
- [30] Xu, M., Zhao, Z., Wang, P., Duan, S., Lei, H., & Fang, D. (2022). Mechanical performance of bio-inspired hierarchical honeycomb metamaterials. *International Journal of Solids and Structures*, 254–255, 111866. <https://doi.org/10.1016/j.ijsolstr.2022.111866>
- [31] Mizzi, L., & Spaggiari, A. (2020). Lightweight mechanical metamaterials designed using hierarchical truss elements. *Smart Materials and Structures*, 29, 105036. <https://doi.org/10.1088/1361-665X/aba53c>
- [32] Cheng, L., Liu, J., Liang, X., & To, A. (2018). Coupling lattice structure topology optimization with design-dependent feature evolution for additive manufactured heat conduction design. *Computer Methods in Applied Mechanics and*

- Engineering, 332, 408–439. <https://doi.org/10.1016/j.cma.2017.12.024>
- [33] Alvarez-Trejo, A., Cuan-Urquizo, E., Roman-Flores, A., Trapaga-Martinez, L., & Alvarado-Orozco, J. (2021). Bezier-based metamaterials: Synthesis, mechanics and additive manufacturing. *Materials & Design*, 199. <https://doi.org/10.1016/j.matdes.2020.109412>
- [34] Saurabh, S., Gupta, A., & Chowdhury, R. (2023). Impact of parametric variation to achieve extreme mechanical metamaterials through topology optimization. *Composites Structures*, 326. <https://doi.org/10.1016/j.compstruct.2023.117611>
- [35] Pan, C., Han, Y., & Lu, J. (2020). Design and optimization of lattice structures: A review. *Applied Sciences-Basel*, 10. <https://doi.org/10.3390/app10186374>
- [36] Abou-Ali, A., Lee, D., & Abu Al-Rub, R. (2022). On the effect of lattice topology on mechanical properties of SLS additively manufactured sheet-, ligament-, and strut-based polymeric metamaterials. *Polymers*, 14. <https://doi.org/10.3390/polym14214583>
- [37] Huang, Y., Zhang, Y., Li, J., Hoang, T.-D., & Jiang, Q. (2025). Metamaterials and smart structure integration enabled by 3D and 4D printing technologies: A review. *Materials Emerging Technologies for Sustainability*, 1, 2530003. <https://doi.org/10.1142/S3060932125300033>
- [38] Zegard, T., & Paulino, G. (2016). Bridging topology optimization and additive manufacturing. *Structural and Multidisciplinary Optimization*, 53, 175–192. <https://doi.org/10.1007/s00158-015-1274-4>
- [39] Jiang, J., Huang, Y., Yin, A., He, Y., Xu, X., Yan, X., et al. (2025). Evaluation of microstructure and mechanical properties of annealed SLMed Ti–6Al–4V alloy using laser ultrasonics. *Journal of Materials Research and Technology*, 37, 380–388. <https://doi.org/10.1016/j.jmrt.2025.05.241>
- [40] Wang, H., You, J., Tian, Y., Chen, Z., & Yin, S. (2025). Dynamic mechanical behavior of sinusoidal corrugated dual-phase lattice metamaterials by additive manufacturing. *Experimental Mechanics*, 65, 541–551. <https://doi.org/10.1007/s11340-025-01160-7>
- [41] Han, Z., Zhou, Y., Xu, Z., Wei, K., Zhao, J., He, Z., et al. (2025). High stable auxetic metamaterials developed through feature-control topology optimization and additive manufacturing. *Thin-Walled Structures*, 213. <https://doi.org/10.1016/j.tws.2025.113305>
- [42] Ling, B., Wei, K., Wang, Z., Yang, X., Qu, Z., & Fang, D. (2020). Experimentally program large magnitude of Poisson's ratio in additively manufactured mechanical metamaterials. *International Journal of Mechanical Sciences*, 173. <https://doi.org/10.1016/j.ijmecsci.2020.105466>
- [43] Li, Y., Jiang, D., Zhao, R., Wang, X., Wang, L., & Zhang, L. (2024). High mechanical performance of lattice structures fabricated by additive manufacturing. *Metals*, 14. <https://doi.org/10.3390/met14101165>
- [44] Dong, G., Tang, Y., Li, D., & Zhao, Y. (2020). Design and optimization of solid lattice hybrid structures fabricated by additive manufacturing. *Additive Manufacturing*, 33. <https://doi.org/10.1016/j.addma.2020.101116>
- [45] Qin, H., Yang, D., & Ren, C. (2018). Design method of lightweight metamaterials with arbitrary Poisson's ratio. *Materials*, 11. <https://doi.org/10.3390/ma11091574>
- [46] Lu, Z., Liu, L., Chen, Z., Wang, C., Zhu, X., Lu, X., et al. (2025). Progress in electromagnetic wave absorption of multifunctional structured metamaterials. *Polymers*, 17. <https://doi.org/10.3390/polym17182559>
- [47] Sharma, D., & Hiremath, S. (2022). Additively manufactured mechanical metamaterials based on triply periodic minimal surfaces: Performance, challenges, and application. *Mechanics of Advanced Materials and Structures*, 29, 5077–5107. <https://doi.org/10.1080/15376494.2021.1948151>
- [48] Wang, Y., Gao, J., Luo, Z., Brown, T., & Zhang, N. (2017). Level-set topology optimization for multimaterial and multifunctional mechanical metamaterials. *Engineering Optimization*, 49, 22–42. <https://doi.org/10.1080/0305215X.2016.1164853>
- [49] Liu, P., Kang, Z., & Luo, Y. (2020). Two-scale concurrent topology optimization of lattice structures with connectable microstructures. *Additive Manufacturing*, 36. <https://doi.org/10.1016/j.addma.2020.101427>
- [50] Guo, X., Zhou, J., Zhang, W., Du, Z., Liu, C., & Liu, Y. (2017). Self-supporting structure design in additive manufac-

- turing through explicit topology optimization. *Computer Methods in Applied Mechanics and Engineering*, 323, 27–63. <https://doi.org/10.1016/j.cma.2017.05.003>
- [51] Wang, X., Zhang, L., Song, B., Zhang, Z., Zhang, J., Fan, J., et al. (2022). Tunable mechanical performance of additively manufactured plate lattice metamaterials with half-open-cell topology. *Composites Structures*, 300. <https://doi.org/10.1016/j.compstruct.2022.116172>
- [52] Garland, A., White, B., Jensen, S., & Boyce, B. (2021). Pragmatic generative optimization of novel structural lattice metamaterials with machine learning. *Materials & Design*, 203. <https://doi.org/10.1016/j.matdes.2021.109632>
- [53] Mccaw, J., & Cuan-Urquizo, E. (2018). Curved-layered additive manufacturing of non-planar, parametric lattice structures. *Materials & Design*, 160, 949–963. <https://doi.org/10.1016/j.matdes.2018.10.024>
- [54] Zhong, H., Song, T., Li, C., Das, R., Gu, J., & Qian, M. (2023). The Gibson-Ashby model for additively manufactured metal lattice materials: Its theoretical basis, limitations and new insights from remedies. *Current Opinion in Solid State and Materials Science*, 27. <https://doi.org/10.1016/j.cossms.2023.101081>
- [55] Wang, Y., Li, S., Yu, Y., Xin, Y., Zhang, X., Zhang, Q., et al. (2020). Lattice structure design optimization coupling anisotropy and constraints of additive manufacturing. *Materials & Design*, 196. <https://doi.org/10.1016/j.matdes.2020.109089>
- [56] Teng, F., Sun, Y., Guo, S., Gao, B., & Yu, G. (2022). Topological and mechanical properties of different lattice structures based on additive manufacturing. *Micromachines*, 13. <https://doi.org/10.3390/mi13071017>
- [57] Zhao, Z., Yuan, C., Lei, M., Yang, L., Zhang, Q., Chen, H., et al. (2019). Three-dimensionally printed mechanical metamaterials with thermally tunable auxetic behavior. *Physical Review Applied*, 11. <https://doi.org/10.1103/PhysRevApplied.11.044074>
- [58] Zhao, M., Cui, J., Chen, L., Jin, K., & Zeng, Z. (2025). Enhanced mechanical properties and energy absorption of lattice metamaterials inspired by crystal imperfections. *Composites Structures*, 356. <https://doi.org/10.1016/j.compstruct.2025.118894>
- [59] Zhao, W., Yue, C., Liu, L., Leng, J., & Liu, Y. (2023). Mechanical behavior analyses of 4D printed metamaterials structures with excellent energy absorption ability. *Composites Structures*, 304. <https://doi.org/10.1016/j.compstruct.2022.116360>
- [60] Sarracino, A., Hammetter, C., Whetten, S., Bishop, S., McCoy, C., Clem, P., et al. (2024). Elastic wave suppression through additively manufactured petal lattice metamaterials. *Journal of Applied Physics*, 135. <https://doi.org/10.1063/5.0196004>
- [61] Li, Z., Luo, Z., Zhang, L., & Wang, C. (2021). Topological design of pentamode lattice metamaterials using a ground structure method. *Materials & Design*, 202. <https://doi.org/10.1016/j.matdes.2021.109523>
- [62] Wang, Y., Luo, Z., Zhang, N., & Kang, Z. (2014). Topological shape optimization of microstructural metamaterials using a level set method. *Computational Materials Science*, 87, 178–186. <https://doi.org/10.1016/j.commatsci.2014.02.006>
- [63] Almesmari, A., Barsoum, I., & Abu Al-Rub, R. (2024). Topology optimised novel lattice structures for enhanced energy absorption and impact resistance. *Virtual and Physical Prototyping*, 19. <https://doi.org/10.1080/17452759.2024.2361463>
- [64] Srivatsa, S., Kumar, R., Selva, D., & Silberstein, M. (2022). Examining the impact of asymmetry in lattice-based mechanical metamaterials. *Mechanics of Materials*, 172. <https://doi.org/10.1016/j.mechmat.2022.104386>
- [65] Gong, C., Ritchie, R., Wei, X., Liu, Q., & Xiong, J. (2025). Mechanical properties of modular assembled composite lattice architecture. *Journal of the Mechanics and Physics of Solids*, 195. <https://doi.org/10.1016/j.jmps.2024.105967>
- [66] Ma, W., Yang, H., Zhao, Y., Li, X., Ding, J., Qu, S., et al. (2025). Multi-physical lattice metamaterials enabled by additive manufacturing: Design principles, interaction mechanisms, and multifunctional applications. *Advanced Science*, 12. <https://doi.org/10.1002/advs.202405835>
- [67] Mizzi, L., Simonetti, A., & Spaggiari, A. (2024). Mechanical properties and failure modes of additively-manufactured chiral metamaterials based on Euclidean tessellations: An experimental and finite element study. *Rapid Prototyping Journal*, 30, 59–71. <https://doi.org/10.1108/RPJ-06-2023-0190>

**Dear Researchers and Scholars :**

Greetings from Asia Pacific Science Press, a beacon of academic and scientific publishing, located in the vibrant city of Hong Kong.

We extend our heartfelt gratitude for your relentless pursuit of knowledge, and your significant contributions to the advancement of science and society. It is researchers and scholars like you who propel humanity forward, and we at the Asia Pacific Science Press are devoted to ensuring that your groundbreaking works receive the global recognition they rightfully deserve.

In light of our commitment to disseminating pioneering research across various disciplines, such as medicine, architecture, education, and electronics, we are reaching out with two pivotal opportunities to augment our collaboration with the global academic community:

**Call for Paper Submissions:**

We cordially invite you to submit your original research articles to our fast-growing, peer-reviewed, and open-access journals. Our platform guarantees an extensive, global reach, enabling your work to garner maximum visibility and citation in the academic sphere. Rest assured, your work will be meticulously assessed by experts in the field, ensuring it receives the acknowledgment and exposure it merits.

**Join Our Esteemed Team:**

We are fervently searching for passionate researchers and scholars interested in joining our burgeoning team at Asia Pacific Science Press. We offer numerous roles, such as peer reviewers, editors, and advisory board members, where your expertise will significantly shape the content and quality of our publications. In return, you will gain invaluable experience, network with preeminent scholars, and play a pivotal role in molding the future of global academic publishing.

**Why Choose Asia Pacific Science Press?**

**Global Reach:** Your work will be accessible to a worldwide audience, free from any access barriers.

**Collaboration with Renowned Universities:** We have established extensive publishing systems in cooperation with world-renowned universities, such as Wuhan University, Hong Kong University, and the University of Malaya.

**Diverse Disciplines:** Your research will be housed among numerous journals across a multitude of academic projects and disciplines.

As we stride forward in the academic landscape, we envision a future where our collective efforts shape a more enlightened, innovative, and interconnected global society. We sincerely hope that you consider this invitation to join us on this auspicious journey towards knowledge, discovery, and global impact.

Should you wish to submit your work or express interest in joining our team, please do not hesitate to contact us. You can submit your manuscript or personal profile to [info@apspublisher.com](mailto:info@apspublisher.com) or visit our website at [www.apspublisher.com](http://www.apspublisher.com) for more information.

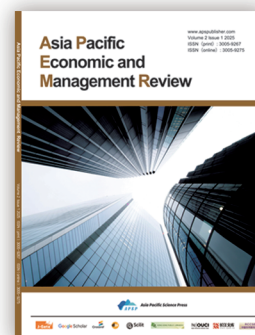
Thank you for considering this opportunity, and we eagerly anticipate the possibility of welcoming you to the Asia Pacific Science Press family. Together, let's forge a future of unparalleled scientific advancement and discovery.

Warm regards  
Asia Pacific Science Press



## OUR JOURNALS

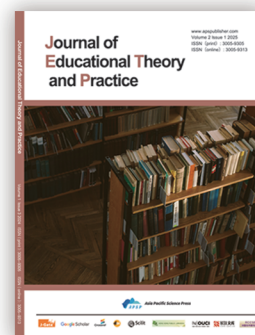
**Asia Pacific Economic and Management Review** is an international, peer-reviewed and open access journal which focuses on theoretical and applied studies of corporate and financial behavior. Aiming to promote the research in fields of business economics and management, it covers mainly but not limits to the following areas: accounting and financial management, economics, human resource management and organizational behavior, information management, international business, strategy and innovation, management science and operations management, marketing and retailing, finance.



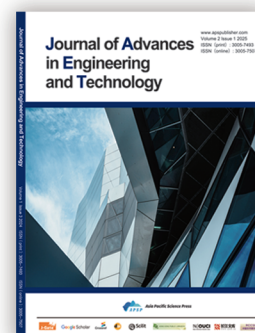
**Critical Humanistic Social Theory** is an journal that publishes papers specifically using quantitative or qualitative research methods for social science research. The journal encourages scholars to conduct social science theory research from the perspective of social critical theory and emphasizes research concerned with issues or methods that cut across traditional disciplinary lines.



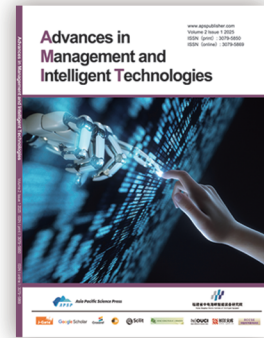
**Journal of Educational Theory and Practice** is an international, peer-reviewed and open access journal which is to promote the evaluative, integrative, theoretical and methodological research on contemporary education; shape a novel, broader view of issues in contemporary education; enhance the caliber of humanities research through active use of best domestic and foreign practices; and integrate the achievements of various sciences and knowledge areas with unconventional approaches.



**Journal of Advances in Engineering and Technology** is an international, peer-reviewed and open access journal which publishes original articles, reviews, short communications, case studies and letters in the field of electronic research and application.



**Advances in Management and Intelligent Technologies** is an international, peer-reviewed, open-access academic journal, hosted by the Fujian Strait Institute of Intelligent Equipment and managed and published by Asia-Pacific Science Press. It focuses on the latest research in the fields of management and intelligent technologies, and aims to advance both theoretical and applied research in management, technological innovation, and intelligent development.



**Asia Pacific Journal of Clinical Medical Research** is an international, peer-reviewed, open access journal dedicated to advancing clinical medical research across multiple disciplines. The journal serves as a platform for publishing high-quality original research, reviews, and clinical studies that enhance the understanding of medical practices, treatment innovations, and healthcare outcomes, thereby supporting patient care and medical advancements in the Asia Pacific region and beyond.



**Asia Pacific Journal of Educational Research** is an international, peer-reviewed, open-access academic journal focusing on educational theory and practice. It publishes high-quality research on educational reform, teaching methods, educational equity, and policy studies. The journal addresses practical needs and institutional changes in the education systems of the Asia-Pacific region, advocating a balance between theoretical inquiry and practical experience. It encourages original studies from multicultural, comparative, and interdisciplinary perspectives, aiming to support educational innovation and policy development across the region.



**Asia Pacific Economic and Social Development** is an international, peer-reviewed, open-access academic journal openly distributed to the global academic community. The journal is committed to publishing original research with theoretical depth and practical value in the fields of economic and social development. It focuses on issues such as economic behavior, social structure transformation, policy innovation, and regional coordinated development in the Asia-Pacific region. The journal encourages interdisciplinary perspectives and promotes the integration of economics, sociology, management, and related disciplines.

