Cost Estimation of the Entire Construction Process based on BAS Optimized CNN

Na Li[1*], Huan Wang[2]
[1]School of Data Science and Engineering, Xi'an Innovation College of Yan'an University, Xi'an 710100, China
[2]School of Continuing Education, Xi'an Innovation College of Yan'an University, Xi'an 710100, China
E-mail: ln_sunshine@126.com
[*]Corresponding author

Keywords: beetle antennae search algorithm (BAS), CNN, whole process of construction project, cost estimation, indicator system construction, convolutional layer

Received: February 14, 2025

In order to cope with the non-linear characteristics of construction project cost related data, capture the regularity between various influencing factors in the entire process of construction project cost, ensure the stability and accuracy of cost estimation, and improve the intelligent cost management and control in the construction field, this study proposes a construction project full process cost estimation method based on beetle antennae search algorithm optimized CNN. Firstly, analyze the factors affecting the overall cost of construction projects and construct an indicator system for these factors. Then, collect relevant data from the indicator system and normalize it, using the processed indicator data as input information for the convolutional neural network. Finally, the BAS algorithm is introduced to determine the optimal number of convolutional layers through iterative optimization, further enhancing the model's ability to capture nonlinear relationships between influencing factors and achieving more accurate cost estimation throughout the entire engineering process. The experiment collected data on M construction projects and real-time market price information of building materials from the Eurostat Construction Cost Index dataset. The expected cost, cost savings rate, and cost risk index were selected as comparison indicators, and the method was compared with reinforcement learning, decision tree modeling, and random forest method. Through experiments, it is known that this method can adaptively iteratively optimize the optimal number of convolutional layers, and the stability of cost estimation is good. In terms of estimated costs, the estimation results of the three comparison methods all exceeded the estimated upper limit set by the experiment, while the estimation result of this method did not exceed the limit; In terms of cost savings, the average cost savings rate of this method is 24.5%-25.1%, significantly higher than reinforcement learning (16.8%-18.5%), decision trees (14.7%-16.1%), and random forests (18.9%-20.4%); In terms of cost risk index, the average risk index of this method remains stable at 0.03-0.05, far lower than reinforcement learning (0.11-0.14), decision trees (0.17-0.21), and random forests (0.13-0.16). Experiments have shown that this method can improve cost savings in engineering projects and promote the development of intelligent cost management in the construction industry.

Povzetek: Članek predstavi metodo za oceno stroškov celotnega gradbenega procesa, ki združuje konvolucijsko nevronsko mrežo (CNN) z optimizacijo števila konvolucijskih slojev prek algoritma iskanja z antenami hrošča (BAS). S tem pristopom model dinamično prilagaja svojo arhitekturo glede na kompleksnost vhodnih podatkov.

1 Introduction

The current construction industry is facing a critical period of transformation and upgrading of engineering cost management methods. In engineering practice, cost estimation runs through the entire life cycle of a project, from early decision-making, design, construction to later operation and maintenance. The cost control of each link directly affects the overall economic benefits of the project [1]. However, traditional cost management has many pain points: subjective biases caused by relying on manual experience, collaborative difficulties caused by information silos, and the inability of static estimates to adapt to market fluctuations, which make project cost overruns common. With the continuous expansion and increasing complexity of construction projects, traditional cost estimation methods are no longer able to meet the precise and dynamic requirements of modern engineering management.

For construction projects, implementing a full cycle cost estimation mechanism has profound practical value. It can not only accurately estimate and regulate project expenses, ensuring stable progress within the predetermined budget, but also greatly benefit the construction enterprise's competitive advantage in the market [2]. For this reason, reference [3] proposed a hybrid modeling method combining Autoregressive Integrated

Moving Average (ARIMA) and Nonlinear Autoregressive Neural Network (NARNN) for time series analysis and cost estimation of building engineering material indicators. Among them, the ARIMA model effectively captures the linear trends and seasonal fluctuations in material prices through differential operations and linear parameter fitting; The NARNN model enhances the representation ability of complex nonlinear relationships through nonlinear activation functions and hidden layer structures. This method attempts to use the linear prediction results of ARIMA as input to NARNN for joint modeling to achieve full process cost estimation. However, ARIMA is essentially only able to handle linear components and lacks modeling capabilities for the abrupt fluctuations commonly present in building material prices, resulting in non-linear residuals being forcibly incorporated into the compensation process of NARNN, increasing the fitting burden of the latter and having application deficiencies in complex construction engineering scenarios. A conceptual cost estimation method based on project information conversion is proposed in reference [4] to address the difficulty of estimating modular building costs. This method consists of two stages. In the first stage, modular feasibility indicators are extracted by analyzing the design parameters of traditional cast-in-place projects; In the second stage, a comprehensive cost model is established by combining modular cost elements such as manufacturing, transportation, and assembly. However, the framework constructed in this method has poor applicability and is affected by specific project conditions and environmental factors, resulting in a higher cost risk index after applying this method. In reference [5], the general generating function method was used to estimate the engineering cost under uncertainty. This method directly constructs an activity cost function of arbitrary probability quality based on historical data to assist in project risk assessment. Its core consists of four parts: Firstly, establish a universal generation function for activity costs to quantify cost uncertainty; Secondly, abstract the project network as a parallel multi state system to capture complex dependencies between activities; Afterwards, the general generation function for each activity item is integrated to derive the general generation function for the overall project cost; Finally, the probability quality function of project cost is obtained through transformation to clarify the cost distribution. However, establishing a universal generation function requires complex mathematical and statistical methods, resulting in a complex cost estimation process and low effectiveness of this method. In reference [6], with the aim of improving the accuracy of cost estimation for building curtain wall engineering, a construction cost estimation model was designed on an AI platform by integrating case-based reasoning, random forest model, and artificial neural network technology. However, although the AI platform in this method can provide powerful data analysis, it relies too much on the accuracy and completeness of the data. If there are errors or omissions in the historical case data obtained, it will directly affect the accuracy of the model, leading to poor results in the final implementation of construction project cost estimation.

In actual construction project cost estimation, there are many and complex factors that affect the construction project cost, and there are nonlinear correlations between different factors. Convolutional neural networks have strong nonlinear mapping capabilities and can effectively extract regularities between different influencing factors [7]. However, CNN may get stuck in local optima during the training process, leading to a decrease in the model's generalization ability. Especially when dealing with complex and high-dimensional construction cost data, the problem of local optima is more prominent. To address this issue, this study considers using the Beetle Antennae Search algorithm [8] to optimize CNN. By iteratively optimizing, the optimal number of convolutional layers and other hyperparameters are determined to further enhance CNN's ability to capture nonlinear relationships between influencing factors.

Therefore, this study proposes a construction project cost estimation method based on the beetle antennae search algorithm optimized CNN, which can achieve more accurate construction project cost estimation, thereby assisting in the rational allocation of resources in construction projects and improving the scientific and effective management and control of construction project costs.

2 The design of the whole process cost estimation method for construction projects

2.1 Constructing a whole-process cost estimation index system for construction projects

Construction project construction cycle is long, which affects the whole process of construction engineering cost estimation of factors are more complex [9]. Constructing the whole process of construction cost estimation index system, mainly through the detailed classification and quantification of the cost of the project, helps project managers to more accurately grasp the cost composition and changes, so as to be able to timely discover the cost deviation, and improve the effect of cost control [10]. In order to simplify the complexity of estimating the full cycle cost of construction projects while ensuring their accuracy, this study selected the core influencing factors throughout the entire cycle of construction projects as the cornerstone of the cost estimation index system. For specific index details, please refer to Table 1.

Table 1: Index system of cost estimation in the whole process of construction project

Index	Units	Interpretation
Construction site	p.u.	The choice of construction site will directly affect factors such as raw material procurement costs, energy expenses, water source acquisition, transportation costs, and the availability of local construction forces. The combined effect of these factors will have a significant impact on the final project cost.
Expected construction period	Days	The length of the Expected Construction Period is directly related to the input of labor costs, the level of equipment rental expenses, and the overall time value of project funds. A longer construction period often means a higher total cost.
Cost index	p.u.	Cost index is an important indicator that reflects the overall trend of engineering cost changes. It can help estimators understand the current market cost level and predict project costs more accurately.
Area of structure	m^2	Area of structure is the foundation of construction cost estimation, which directly affects multiple costs such as material consumption, labor costs, equipment usage fees, etc. An increase in building area usually means an increase in total cost.
Building structure	p.u.	Different building structures have a significant impact on the amount of materials used, construction difficulty, and cost. For example, complex structures may require more materials and finer construction, thereby increasing costs.
Types of pile foundations	p.u.	Types of pile foundations directly affect the cost of foundation treatment. The construction difficulty, material consumption, and required construction period of different types of pile foundations vary, which in turn affects the overall cost.
Number of underground floors	Floor	The increase in the number of underground floors will increase the cost of earthwork excavation, support, waterproofing, and other engineering projects. Meanwhile, the development of underground spaces often requires more complex construction techniques and materials, further increasing costs.
Number of floors above ground	Floor	The increase in the number of floors above ground will directly lead to an increase in building area, which in turn increases material consumption, labor costs, and other costs. In addition, high-rise buildings also need to consider additional costs such as vertical transportation and safety protection.
Standard floor height	m	The change in standard floor height will affect the building area, which in turn will have an impact on the cost. Meanwhile, designs with different floor heights may require different materials and construction techniques, which can also have an impact on the cost.
Building height	m	The increase in building height will increase vertical transportation costs, safety protection costs, and structural strengthening costs under possible wind loads and earthquakes, thereby significantly affecting the cost.
Architectural form	p.u.	Complex architectural form design requires more building materials and more sophisticated construction techniques, thereby increasing costs. In addition, the unique exterior design may also involve special processing and installation costs.
Rate of change of average price of rebar	%	Steel reinforcement is one of the important building materials in construction engineering, and its price changes will directly affect the project cost. The increase in the average price of steel bars will lead to an increase in material costs, thereby pushing up the total cost.
Rate of change of average price of concrete	%	Concrete is another key material in construction engineering, and changes in its price can also have a significant impact on the cost. The fluctuation of the average price of concrete will directly affect the material cost, thereby affecting the total cost.
Types of wall decoration	p.u.	Different types of wall decorations involve different materials and techniques, which directly affect the cost. High end decoration materials and complex decoration techniques often significantly increase the cost.

As shown in Table 1, a total of 14 indicators of key influencing factors in the whole process of construction engineering were mainly extracted, of which the details of each indicator are as follows.

a. Construction location: affecting the construction period, construction quality, etc., the difference between different locations of construction projects is mainly reflected in the raw materials, energy, water, transportation, construction forces and other factors, affecting the final cost of the project.

b. Estimated duration: this cycle refers to the expected length of time required to complete all planned construction work while ensuring that the engineering project meets established quality standards. The length of the construction period is a key factor in labor cost consumption and has a direct impact on the overall cost of the final project.

c. Cost index: a measure of the relative change in the cost of construction works over time, reflecting the overall trend in the cost of works.

d. Floor area: one of the basic indicators for estimating the cost of construction works, including gross floor area, floor area and net usable area.

e. Building structure: refers to the building by the foundation, walls, columns, beams, floors, roofs and other components of the load-bearing system, different forms of building structure has a significant impact on the cost of the project.

f. Types of pile foundations: Refers to the types of pile foundations used in building foundation engineering. Different types of pile foundations have different impacts on project cost and construction difficulty.

g. Number of underground floors: the number of underground floors of a building located below ground level, where an increase in the number of basement floors will increase the cost of the project.

h. Number of floors above ground: this refers to the number of floors above the ground level of a building, which determines the total height and floor area of the building, and an increase in the number of floors above ground will result in an increase in the cost of the project.

i. Standard floor height: the average height of each floor of a building, with variations in the standard floor height affecting the floor area and the cost of the project.

j. Building height: it refers to the vertical distance between the highest point of a building from the ground surface to its top, an increase in building height increases the cost of the project.

k. Architectural form: refers to the exterior shape and design style of the building, complex building form requires more building materials and construction techniques, increasing the cost of the project.

l. Rate of change in the average price of rebar: the percentage change in the market price of rebar over time, reflecting supply and demand and price fluctuations in the rebar market.

m. Rate of change of average price of concrete: the percentage change in the market price of concrete over time, reflecting supply and demand and price fluctuations in the concrete market.

n. Type of wall decoration: refers to the materials and methods of finishing the walls of a building, with high-grade finishing materials and complex finishing methods increasing the cost of the project [11].

2.2 Establishing a cost estimation model for the whole process of construction engineering based on convolutional neural networks

Based on the selected core influencing factors throughout the entire lifecycle of construction projects, this study utilizes Convolutional Neural Networks (CNN) to learn a large amount of historical cost data, automatically extract key factors that affect the overall cost of the project, and establish a complex relationship model between influencing factors and cost, ultimately achieving full process cost estimation.

2.2.1 Determining the sequence of factors affecting the whole process of cost estimation in construction projects

In the practical operation of estimating the full cycle cost of construction projects, considering the specificity of construction site conditions may increase redundancy of convolutional neural networks in the estimation process, this study chooses to exclude this factor and instead focuses on other indicators that have a critical impact on the full cycle cost estimation of construction projects, and sets the total number of construction project sample data as c. At this point, the sequence of influencing factors for the cost estimation of the entire construction process is shown in Table 2.

Table 2: Input sequence of impact factors for cost estimation of the whole process of construction engineering

Index	Construction project 1	Construction project 2	…	Construction project c	Descriptor
Expected construction period a_1	$a_{1,1}$	$a_{1,2}$	…	$a_{1,c}$	Quantitative description
Cost index a_2	$a_{2,1}$	$a_{2,2}$	…	$a_{2,c}$	Quantitative description
Area of structure a_3	$a_{3,1}$	$a_{3,2}$	…	$a_{3,c}$	Quantitative description

Building structure a_4	$a_{4,1}$	$a_{4,2}$	…	$a_{4,c}$	Qualitative description
Types of pile foundations a_5	$a_{5,1}$	$a_{5,2}$	…	$a_{5,c}$	Qualitative description
Number of underground floors a_6	$a_{6,1}$	$a_{6,2}$	…	$a_{6,c}$	Quantitative description
Number of floors above ground a_7	$a_{7,1}$	$a_{7,2}$	…	$a_{7,c}$	Quantitative description
Standard floor height a_8	$a_{8,1}$	$a_{8,2}$	…	$a_{8,c}$	Quantitative description
Building height a_9	$a_{9,1}$	$a_{9,2}$	…	$a_{9,c}$	Quantitative description
Architectural form a_{10}	$a_{10,1}$	$a_{10,2}$	…	$a_{10,c}$	Qualitative description
Rate of change of average price of rebar a_{11}	$a_{11,1}$	$a_{11,2}$	…	$a_{11,c}$	Quantitative description
Change rate of average price of concrete a_{12}	$a_{12,1}$	$a_{12,2}$	…	$a_{12,c}$	Quantitative description
Types of wall decoration a_{13}	$a_{13,1}$	$a_{13,2}$	…	$a_{13,c}$	Qualitative description

In Table 2, a_4, a_5, a_{10} and a_{13} are qualitative indicators, while the rest are quantitative indicators. Among them: In a_4, the frame shear wall, shear wall, and frame correspond to -1, 0, and 1, respectively; In a_5, drilled cast-in-place piles, prefabricated reinforced concrete piles, and high-pressure rotary jet grouting piles correspond to -1, 0, and 1, respectively; In a_{10}, rectangles, cylinders or polygonal cylinders, and spherical objects correspond to -1, 0, and 1, respectively; In a_{13}, exterior wall coatings, glass curtain walls, and stone (brick) materials correspond to -1, 0, and 1, respectively.

For the above sequence of influence factors of the whole process cost estimation of construction project, before inputting into the convolutional neural network model, it is necessary to carry out the normalization process, so as to effectively unify the outline of the influence factors of the whole process cost estimation of construction project [12]. The normalization process is expressed as Equation (1):

$$A = \left(A_b - A_{\min}\right) / \left(A_{\max} - A_{\min}\right) \quad (1)$$

Among them, $A_{\max}$ and $A_{\min}$ respectively represent the maximum and minimum values of the influencing factor sequence of the input construction project cost estimation throughout the entire process. In this process, the minimum maximum standardization method was applied in this study. The cost estimation of construction projects requires handling complex nonlinear relationships. The minimum maximum normalization method can compress data to a fixed range, ensuring that the input values of the activation function are evenly distributed and avoiding gradient saturation or vanishing problems.

2.2.2 Estimation of the whole process cost of construction projects based on CNN

Based on the cost estimation impact factor sequence determined in section 2.2.1, comprehensively collect cost data throughout the entire construction process, and input the data into a convolutional neural network [13] model to automatically learn and identify complex correlations and potential patterns between various impact factors, ultimately obtaining accurate cost estimation results for the entire construction process.

Convolutional neural network is mainly composed of input layer, convolutional layer, pooling layer, fully connected layer and output layer [14]. The sequence of influencing factors for the whole process cost estimation of construction project obtained in subsection 2.2.1 is inputted into the convolutional neural network model to be processed, and the structure is shown in Fig. 1.

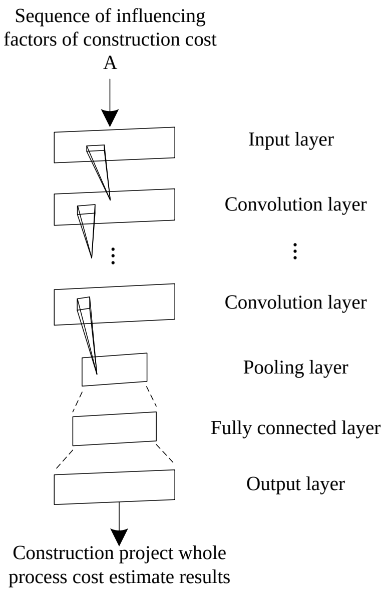

Fig 1: Convolution neural network model structure.

In Fig. 1, the input dimension of the input layer is 13×1 (corresponding to 13 normalized impact factor sequences), and the data format is single channel one-dimensional time-series data to adapt to the sequence characteristics of the impact factors in construction engineering; The convolution kernel size of the first convolutional layer is 3×1, the number of filters is 64, the step size is 1×1, and the activation function is ReLU (introducing nonlinearity and alleviating gradient vanishing problem); The convolution kernel size of the second convolutional layer is 2×1, the number of filters is 128, the step size is 1×1, and the activation function is ReLU; The pooling type of the pooling layer is one-dimensional max pooling, with a pooling window size of 2×1 and a step size of 2×1; There are a total of 256 neurons in the fully connected layer, with ReLU as the activation function; The number of neurons in the output layer is 1, and the activation function is linear activation.

As shown in Fig. 1, the roles of each layer in constructing the whole process cost estimation model of construction project based on convolutional neural network are as follows:

1) Input layer

It is mainly responsible for receiving the preprocessed data on the sequence of factors affecting the whole process of cost estimation of construction works, which is expressed in the form of a matrix as Equation (2):

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,c} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,c} \\ \vdots & \vdots & & \vdots \\ a_{13,1} & a_{13,2} & \cdots & a_{13,c} \end{bmatrix}_n \tag{2}$$

Among them, n represents the number of sequence data of influencing factors in the cost estimation of the entire construction process, and $n = 1, 2, \cdots, N$, N represents the total number of sequences.

2) Convolutional layers

Convolutional layer is the core component of CNN, which contains multiple convolutional kernels [15]. Through the convolution kernel on the input construction engineering cost estimation impact factor sequence data in the local region of the weighted sum and nonlinear activation operation, so as to realize the effective convolution operation, in order to adaptively learn the complex relationship between the input construction engineering cost estimation impact factor sequence data. The computational process is summarized as Equation (3) and (4):

$$y_{1C} = \text{sigmoid}\left(\omega_C \cdot A + \delta_C\right) \tag{3}$$

$$y_{dC} = \text{sigmoid}\left(\omega_C \cdot y_{(d-1)C} + \delta_C\right) \tag{4}$$

Among them, y_{1C} and y_{dC} represent the output values of the first and d -th convolutional layers, $\text{sigmoid}(\cdot)$ represents the sigmoid activation function, and ω_C and δ_C represent the convolution kernel weights and biases, respectively. In equation (4), it represents using the output value of the $d-1$ convolutional layer as the input value of the d -th convolutional layer.

3) Pooling layer

Pooling layer mainly through the down-sampling operation to reduce the output value obtained after the convolution operation of the convolution layer, while retaining the complex relationship between the sequence of data of the impact factor of the whole process cost estimation of construction engineering [16], through the maximum pooling method of the whole process cost estimation model of construction engineering convolution layer output value processing, expressed as Equation (5):

$$y_P = \text{MaxPooling}\left(y_C\right) \tag{5}$$

Among them, y_C represents the final output value of the convolutional layer of the cost estimation model for the entire construction process, $\text{MaxPooling}(\cdot)$ represents the maximum pooling calculation, and y_P represents the output value after maximum pooling.

4) Fully connected layers

The fully connected layer plays a crucial role in convolutional neural networks, as it is responsible for receiving the output from the pooling layer and flattening and mapping these multidimensional feature data to the final output space through a series of linear transformations and the application of activation functions. This process effectively integrates the complex features extracted by the preceding layer, providing the model with comprehensive mapping capabilities from input to output [17]. The calculation of a fully connected layer can be expressed as:

$$y_F = \mathrm{ReLU}\left(\omega_F \cdot y_P + \delta_F\right) \tag{6}$$

Among them, ω_F and δ_F respectively represent the weight and bias of the fully connected layer in the cost estimation model of the entire construction process, $\mathrm{ReLU}(\cdot)$ represents the ReLU activation function, and y_F represents the final output value after being flattened by the fully connected layer processing.

5) Output layer

Responsible for calculating and outputting the final cost estimate for the whole process of construction works, which is expressed as Equation (7):

$$Y = \omega_O \cdot y_F + \delta_O \tag{7}$$

Among them, ω_O and δ_O respectively represent the weights and biases of the output layer of the cost estimation model for the entire construction process, and Y represents the final cost estimation result for the entire construction process.

2.3 Estimation results of CNN optimization based on beetle antennae search algorithm

In the actual cost estimation process, it is easy to be affected by the complex and changeable construction engineering situation, and there are large uncertainties in many aspects of the whole process cost estimation of construction engineering, in order to more accurately identify and extract the key information in the sequence data of the impact factors of the whole process cost estimation of construction engineering, and optimize the convolutional neural network model by the Beetle Antennae Search Algorithm (BAS), to further improve the accuracy of the cost estimation [18].

Define the hypotheses: Compared with conventional CNN technology, BAS improves the performance of CNN.

Beetle Antennae Search Algorithm originates from the predatory activity of beetle with two whiskers, and it is a kind of intelligent optimization algorithm that can be used to search for the best quickly and accurately without explicit mathematical model [19]. The number of convolutional layers of the convolutional neural network was optimized by the Beetle Antennae Search Algorithm, let the prey in the Beetle Antennae Search Algorithm represent the optimal number of convolutional layers H of the convolutional neural network, the position is the trend of the number of convolutional layers, The convolution layer number value is the convolutional layer number pheromones, which guides the Beetle Antennae Search Algorithm to obtain the convolution layer structure of the optimal convolutional neural network. The structure of the optimized convolutional neural network model by the Beetle Antennae Search Algorithm is shown in Fig. 2.

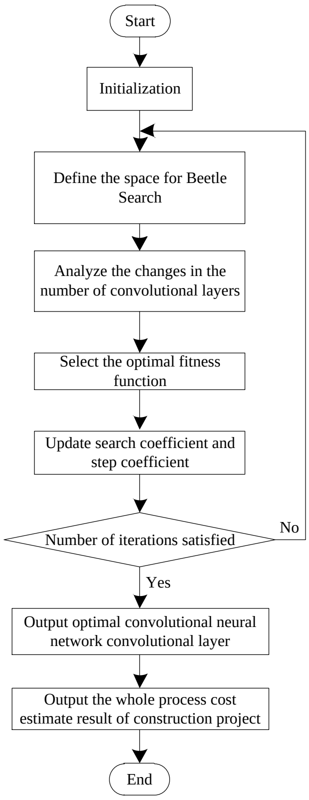

Figure 2: Optimization structure of CNN

The specific optimization process is as follows:

Step 1: Define the beetle search space, i.e., the optimization space of the number of convolutional layers in the constructed whole-process cost estimation model of the construction project, and carry out the normalization of the direction vectors, which is expressed as Equation (8):

$$\vec{\chi} = \frac{\mathrm{rand}\left(G,1\right)}{\left\|\mathrm{rand}\left(G,1\right)\right\|} \tag{8}$$

Among them, $\vec{\chi}$ represents the random direction vector in the beetle antennae search algorithm; $\mathrm{rand}(\cdot)$ represents a random number within [-1,1]; G represents the spatial dimension for optimizing the number of convolutional layers in the cost estimation model of the entire construction process.

Step 2: The actions of two antennae for beetle in BAS algorithm are regarded as the search behavior of the number of convolutional layers, and the length of the two antennae decreases gradually with the increase of the number of iterations, then the change of the number of convolutional layers is as follows:

$$\overrightarrow{S_L}(t) = \vec{S}(t) - \lambda(t)\vec{\chi} \tag{9}$$

$$\overrightarrow{S_R}(t) = \vec{S}(t) + \lambda(t)\vec{\chi} \tag{10}$$

Among Equation (9) and (10), $\overrightarrow{S_L}(t)$ represents the change in the number of convolutional layers in the estimation model of the t-th iteration of the Taurus whisker, $\vec{S}(t)$ represents the change in the number of convolutional layers in the t-th iteration, $\lambda(t)$ represents the length of the beetle antennae in the t-th iteration, and $\overrightarrow{S_R}(t)$ represents the change in the number of convolutional layers in the t-th iteration of the right beetle antennae estimation model.

Step 3: The optimal number of convolutional layers odor concentration of the convolutional neural network detected by beetle whiskers is expressed through the fitness function [20]. The optimal fitness function is selected to update the trend of the number of convolutional layers, which is computationally expressed as Equation (11):

$$\overrightarrow{S_L}(t+1)=\vec{S}(t)-\mu(t)\operatorname{sign}\left\{\operatorname{fit}\left[\overrightarrow{S_R}(t)\right]-\operatorname{fit}\left[\overrightarrow{S_L}(t)\right]\right\} \quad (11)$$

Among them, $\overrightarrow{S_L}(t+1)$ represents the trend vector of the change in the number of new convolutional layers; $\mu(t)$ represents the search step factor for the number of convolutional layers in the cost estimation model of the entire construction process, $\operatorname{sign}\{\cdot\}$ represents the sign function, and $\operatorname{fit}[\cdot]$ represents the fitness function for the optimal number of convolutional layers and odor concentration detected by the beetle antennae convolutional neural network.

Step 4: Since the search coefficient of the number of convolutional layers and the step coefficient in the whole process cost estimation model of the construction project change with the increase of the number of iterations, they are expressed as Equation (12) and (13):

$$\lambda(t)=0.95\lambda(t-1)+1 \quad (12)$$

$$\mu(t)=0.95\mu(t-1) \quad (13)$$

Among them, $\lambda(t-1)$ represents the length of the beetle's tendrils in the $t-1$-th iteration, and $\mu(t-1)$ represents the search step factor for the number of convolutional layers in the total cost estimation model of the construction project in the $t-1$-th iteration.

Step 5: The updated $\lambda(t)$ and $\mu(t)$ are inserted into the formula (9) - (11) to obtain the optimal number of convolutional neural networks in the whole process cost estimation model of construction engineering.

Through the above process, the convolutional neural network model can be optimized by the Beetle Antennae Search Algorithm, and the results of the whole process cost estimation of the construction project can be obtained finally.

3 Experiments and analysis of results

To verify the effectiveness of the method proposed in this article, the following validation experiments were conducted.

The experiment collected M construction project data and real-time market price information of building materials. These data items include the following:

(A) Construction project data: The experiment obtained full cycle data for M construction projects from the Eurostat Construction Cost Index, cooperative construction companies, and public project reports, covering the following core information:

Project attributes: construction site (commercial/industrial area, etc.), expected construction period (12-36 months), area of structure (5000-100000 square meters), building structure (frame/shear wall, etc.), number of underground floors (0-5 floors), number of floors above ground (5-50 floors), standard floor height (3.0-5.0 meters), building height (20-150 meters), etc.

Cost indicators: expected cost (5 million to 500 million yuan), actual cost deviation rate, cost risk index (0 to 0.3), cost savings rate (5% to 30%), etc.

Dynamic parameters: price change rate of steel bars and concrete (historical fluctuation range ± 20%), market cost index (1.0-1.5), etc.

(B) Building material price data: By integrating with the bulk commodity trading platform API, real-time market prices of key materials such as steel bars and concrete are obtained, and combined with historical price databases (covering nearly 5 years of data), a price volatility model is constructed to ensure data timeliness and continuity.

Input these data into an improved CNN model to achieve accurate cost estimation of the entire construction process. The architecture diagram of the estimation process is shown in Fig. 3.

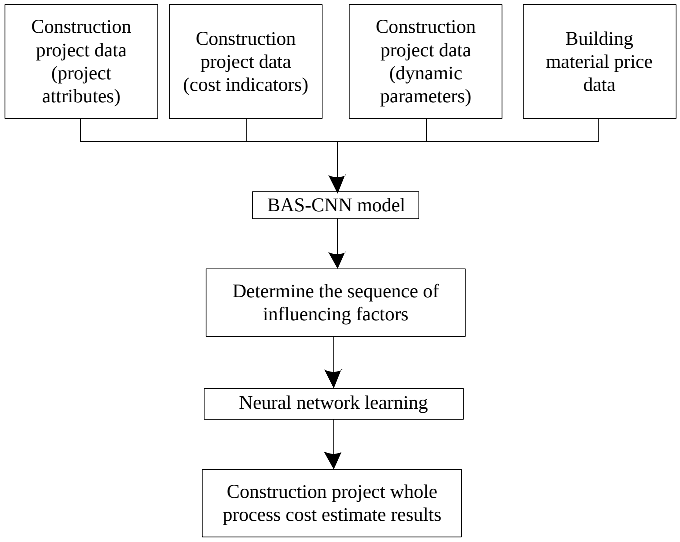

Figure 3: Topology construction whole process of cost estimating process.

For the topology of the whole process cost estimation process of construction project shown in Fig. 3, the data related to the construction project and the real-time market price of related materials are collected and counted, which provides a good data basis for the construction of the subsequent whole process cost estimation index system of the construction project and the initial data set inputted into the whole process cost estimation model of the construction project, and the actual data related to the construction project are shown in Table 3.

Table 3: Data related to construction projects

Parameters	Actual value
Construction site	A business district
Estimated construction period/month	24
Cost index	1.15
Floor area/m^2	50,000
Building structure	Reinforced concrete frame-shear wall structure
Number of underground floors/floors	3
Number of epipelagic floors/floors	20
Standard height/m	3.6
Building height/meter	90
Architectural form	Modern simple style,glass curtain wall and stone wall combination
Change rate of average price of steel bar/%	10
Change rate of average price of concrete/%	5
Types of wall decoration	Stone curtain wall+glass curtain wall
Pile foundation type	Bored pile

The CPU model is Intel Xeon Gold 6248R (2.5 GHz, 24 cores/48 threads), the GPU model is NVIDIA Tesla V100 (32GB video memory, 5120 CUDA core), the main storage is 2TB NVMe SSD, the operating system is Ubuntu 20.04 LTS, the deep learning framework is PyTorch 1.10+CUDA 11.3, and the BAS algorithm is implemented using Python 3.8.

The experimental parameters for the BAS algorithm are set as follows:

(1) The Influence of Step Size Factor on CNN Performance

When the step size factor is 0.01, the convergence of CNN is slow (reaching the optimum at about 200 epochs), but the accuracy of the test set is the highest (92.3%); When the step size factor is 0.05, the convergence speed increases (about 120 epochs), but the accuracy slightly decreases (91.5%), which may be due to skipping the optimal solution due to a large step size; When the step size factor is 0.1, the convergence is faster (about 80 epochs), but the accuracy is significantly reduced (89.8%), indicating that an excessively large step size can impair the optimization effect.

Therefore, the optimal range for the step size factor should be [0.02, 0.04], at which point the CNN converges around 150 epochs and the test accuracy stabilizes at 91.8% -92.1%.

(2) The influence of whisker length on CNN training

When the whisker length is 2, the CNN parameter update direction is relatively conservative, the training is stable but the convergence is slow (180 epochs), and the accuracy is 91.7%; When the whisker length is 6, balancing exploration and development, CNN converges around 130 epochs with an accuracy of 92.0%; When the whisker length is 10, the search range is too large, resulting in severe fluctuations in CNN parameters, a decrease in accuracy to 90.5%, and significant oscillations in the training curve.

Therefore, the optimal range for whisker length is [4, 8], where CNN converges quickly (within 140 epochs) and the accuracy remains stable at 91.5% -92.2%.

(3) The impact of search dimension scaling factor

When the search dimension scaling factor is 0.90, CNN converges quickly (100 epochs), but is prone to getting stuck in local optima, with a testing accuracy of only 90.2%; When the search dimension scaling factor is 0.95, the balance between convergence speed (150 epochs) and global search capability results in an accuracy rate of 91.8%; When the search dimension scaling factor is 0.99, the convergence is slow (220 epochs), but the accuracy is the highest (92.5%), making it suitable for high-precision tasks.

Therefore, the experiment chose a search dimension scaling factor of 0.97.

(4) The impact of population size on CNN optimization

When the population size is 20, the computational efficiency is high (single optimization takes 15 seconds), but the accuracy of CNN fluctuates greatly (± 0.8%); When the population size is 50, the optimization is more stable (with an accuracy fluctuation of ± 0.3%), but the time consumption increases (45s/time), and the accuracy only improves by 0.4%.

Therefore, the experimental population size is set between 30 and 40.

To verify the effectiveness of the Tianniu Xu search algorithm in optimizing the number of convolutional layers in CNN, the “dynamic” changes in the number of convolutional layers were statistically analyzed, and the results are shown in Fig. 4.

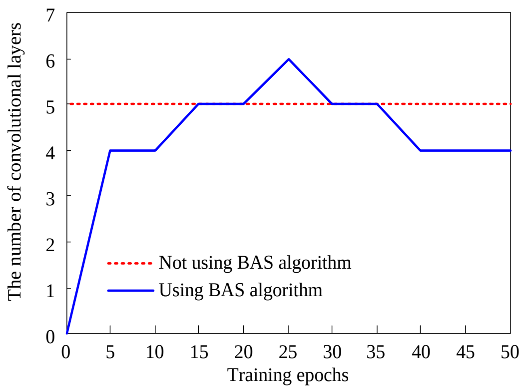

Figure 4: Optimization effect of dynamic changes in the number of convolutional layers.

As shown in Fig. 4, without the BAS algorithm, the number of convolutional layers in CNN is fixed at 5 and cannot be adjusted according to the training state. After using the BAS algorithm, the convolutional layers of CNN can dynamically change, thus dynamically responding to the model state.

In order to verify the effect of this paper's method to realize the whole process cost estimation of construction projects, 8 construction projects are randomly selected, and based on the constructed index system of the whole process cost estimation of construction projects, the cost estimation of different construction projects is carried out by this paper's method, and the results obtained are shown in Table 4.

Table 4: Cost estimation results of the whole process of construction engineering

Index	Construction number							
	1	2	3	4	5	6	7	8
Expected construction period a_1	0.39	0.39	0.31	0.16	0.45	0.41	0.43	0.15
Cost index a_2	0.06	0.26	0.46	0.4	0.25	0.23	0.11	0.17
Area of structure a_3	0.36	0.06	0.18	0.26	0.04	0.09	0.49	0.17
Building structure a_4	-1	0	0	1	-1	-1	0	1
Pile foundation type a_5	0	1	-1	1	0	1	-1	0
Subsurface number a_6	0.5	0.4	0.49	0.32	0.03	0.38	0.15	0.31
Epipelagic number a_7	0.06	0.02	0.44	0.33	0.39	0.47	0.31	0.22
Standard height a_8	0.3	0.5	0.45	0.13	0.05	0.04	0.46	0.03
Building height a_9	0.42	0.19	0.46	0.02	0.36	0.32	0.2	0.17
Architectural form a_{10}	0	-1	-1	1	0	1	-1	-1
Rate of change of average price of rebar a_{11}	0.41	0.27	0.03	0.02	0.29	0.29	0.34	0.28
Change rate of average price of concrete a_{12}	0.28	0.33	0.5	0.34	0.12	0.25	0.5	0.5
Types of wall decoration a_{13}	0	1	1	0	-1	0	1	1
Estimated total price/million yuan	14.93	12.33	18.9	16.35	12.8	8.58	18.38	18.74

Table 4 lists multiple key influencing factors of construction projects and their corresponding cost estimates, comprehensively covering the main aspects of construction project cost estimation. The data in the table has been normalized to ensure that data of different dimensions can be compared and analyzed at the same scale, improving the accuracy and stability of the model. By analyzing the estimation results in Table 4, it can be preliminarily judged that the cost estimation results of this model have high accuracy. For example, for projects with larger building areas and more floors (such as Project 3 and Project 7), the estimated total cost is relatively high, which conforms to the general law of construction project cost.

In order to further verify the actual effectiveness of the method proposed in this article, an estimation of the construction project cost under time variation was made, and the method was compared and analyzed with reinforcement learning method, decision tree model method, and random forest method. Under the premise of setting the maximum limit of 20 million yuan for the full cycle cost estimation of construction projects, the performance effects of various methods in the full cycle cost estimation of construction projects over time are shown in Fig. 5.

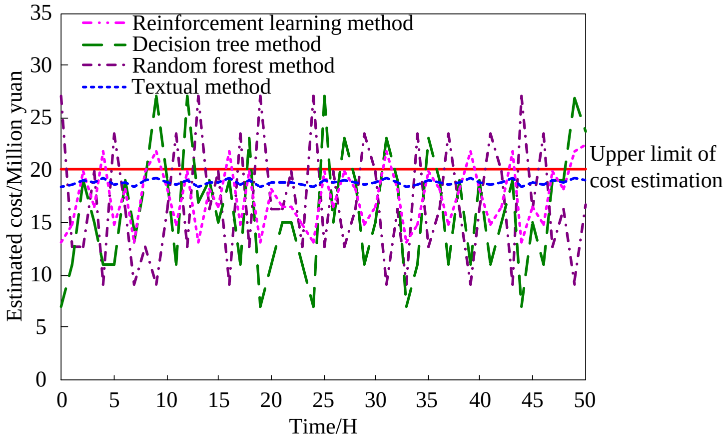

Figure 5: Estimating effect of whole process cost of construction project.

According to the analysis of Fig. 5, compared to the reinforcement learning method, decision tree model method, and random forest method, at different time periods, the volatility of cost estimation using method in this paper was smaller and did not exceed the upper limit of cost estimation (20 million yuan), effectively proving that method in this paper has good stability and reliability.

In order to further verify the effect of this method in realizing the whole process cost estimation of construction projects, 15 construction projects are selected, and compared with the actual investment, the cost saving rate and cost risk index of the cost estimation results achieved according to this method are counted, among which the higher the cost saving rate, the better the cost estimation effect, and the smaller the cost risk index. The actual results obtained according to the method in this paper are shown in Table 5.

Table 5: Evaluation of cost estimation effect achieved by this method

Construction number	Estimated investment/million yuan	Cost estimate/million	Cost saving rate /%	Cost risk index
1	25	17.22	31.12	0.03
2	22	18.48	16	0.06
3	10	8.59	14.1	0.07
4	20	17.66	11.7	0.02
5	25	18.84	24.64	0.11
6	20	15.14	24.3	0.01
7	25	19.09	23.64	0.01
8	20	15.87	20.65	0.04
9	15	12.3	18	0.05
10	15	11.87	20.87	0.07
11	16	12.98	18.88	0.04
12	22	18.56	15.64	0.03

13	20	16.15	19.25	0.03
14	22	18.91	14.05	0.07
15	16	12.2	23.75	0.05

As shown in Table 5, the method in this paper can achieve a better overall cost estimation of construction projects. Through the whole process cost estimation of construction projects realized by this method, the cost of 15 experimental construction projects has been effectively reduced, and the cost saving rate is between 14.05% and 31.12%, all of which have been well saved. The cost risk indexes are all small, all less than 0.15, which proves that this method can effectively save the cost of construction projects, effectively ensure the smooth and stable implementation of construction projects, and the risk of cost capital problems is small.

To verify the statistical significance of the cost savings rate, a single sample t-test was conducted on the cost savings rate data (14.05%-31.12%) of 15 construction projects to test the estimation effect of the method in this paper (assuming the benchmark value is the traditional method's average cost savings rate of 18.2%). The results showed that the average cost saving rate of method in this paper was 22.8%, with a standard deviation of 5.3%, a t-value of 4.62 (14 degrees of freedom), and a p-value of<0.001, indicating that the improvement had high statistical significance (α=0.05). Further calculation of the 95% confidence interval showed an error margin of ± 2.7% and a confidence interval of 20.120.1, demonstrating the robustness of the method in this paper.

On this basis, the method in this paper was compared and analyzed with the reinforcement learning method, decision tree method, and random forest method, using cost savings rate and cost risk index as indicators (the results were averaged). The results are shown in Table 6.

Table 6: Comparison of indicators

Number of tests conducted	Method in this paper		Reinforcement learning method		Decision tree method		Random forest method	
	Cost saving rate /%	Cost risk index	Cost saving rate /%	Cost risk index	Cost saving rate /%	Cost risk index	Cost saving rate /%	Cost risk index
10	24.5	0.05	18.2	0.12	15.3	0.18	20.1	0.14
20	23.8	0.04	17.6	0.13	16.1	0.19	19.5	0.15
30	25.1	0.03	16.8	0.14	14.7	0.21	18.9	0.11
40	24.2	0.04	18.5	0.11	15.9	0.17	20.4	0.13
50	24.7	0.03	17.9	0.12	15.2	0.20	19.8	0.14

Table 6 compares the average cost savings and risk indices of Method in this paper with Enforcement Learning Method, Decision Tree Method, and Random Forest Method under different testing times. The experimental results show that in terms of cost savings, the average cost savings rate of our method is 24.5%-25.1%, significantly higher than reinforcement learning (16.8%-18.5%), decision trees (14.7%-16.1%), and random forests (18.9%-20.4%). This is due to the dynamic optimization of CNN convolutional layers by the BAS algorithm, which can more accurately capture the nonlinear relationships between influencing factors, thereby improving estimation accuracy; In terms of cost risk index, the average risk index of our method remains stable at 0.03-0.05, far lower than reinforcement learning (0.11-0.14), decision trees (0.17-0.21), and random forests (0.13-0.16). This indicates that the optimized CNN model can still maintain stability in complex scenarios, effectively reducing estimation bias caused by data noise or dynamic market changes.

In summary, the method proposed in this article combines the BAS algorithm with CNN, which not only solves the problem of insufficient modeling ability of traditional models for nonlinear data, but also significantly improves the accuracy and stability of estimation through dynamic optimization of network structure. The experimental data further proves that it has significant advantages in cost control and risk avoidance, providing reliable technical support for the whole process cost management of construction projects.

4 Discussion

This study used Table 7 to analyze the limitations and contributions of three conventional methods: reinforcement learning method, decision tree method, and random forest method.

Table 7: Limitations and contribution analysis of conventional methods

Method	Brief idea	Limitation	Contribute
Reinforcement learning method	Optimize the cost estimation model by adjusting strategies through dynamic environmental feedback.	Low-cost savings rate: relying on a large amount of interactive data, long training time, and difficulty in capturing nonlinear relationships; High risk index: sensitive to hyperparameters, prone to falling into local optima, and high estimated volatility.	It has certain adaptability in dynamic scenes and can gradually optimize estimation strategies through reward mechanisms.
Decision tree method	Build a tree structure based on feature splitting and achieve cost prediction through rule partitioning.	The lowest cost saving rate: weak ability to model complex nonlinear relationships, prone to overfitting; The highest risk index: sensitive to data noise and poor generalization ability.	Strong interpretability, suitable for processing structured data, and quickly generating interpretable rules on small-scale datasets.
Random forest method	Integrate multiple decision trees and improve prediction stability through voting or averaging.	Limited cost savings rate: poor processing effect on high-dimensional sparse data (such as dynamic market indicators); High risk index: dependence on data integrity and accuracy, insufficient estimation stability.	By integrating to reduce variance and improve generalization ability, it is suitable for classification and regression tasks on medium-sized datasets.

On the other hand, the method in this paper is significantly superior to conventional methods in the following aspects:

(1) Dynamic optimization of nonlinear modeling capability: The BAS algorithm adaptively adjusts the CNN convolutional layer structure through iterative optimization, accurately capturing the complex nonlinear relationships between influencing factors in construction engineering, and improving cost savings.

(2) Anti noise and stability enhancement: The Tianniu Xu algorithm effectively suppresses data noise interference through direction vector normalization and step size adaptive adjustment. Combined with the local feature extraction ability of CNN, it maintains stable risk index in dynamic market environments.

(3) End to end efficient mapping: Directly inputting normalized multidimensional impact factors, avoiding the cumbersome process of traditional methods relying on manual feature engineering, and improving estimation efficiency.

(4) Dynamic response capability: By dynamically adjusting the number of convolutional layers, estimation bias can be quickly corrected to ensure that the results are always below the preset upper limit.

5 Conclusion

To optimize cost intelligent management and control in the construction field, this article proposes a whole process cost estimation method based on CNN optimized by Beetle Antenna Search Algorithm for construction projects. This method not only overcomes the limitations of conventional CNN in processing nonlinear data, but also significantly improves the accuracy and stability of estimation by dynamically optimizing the network structure. The experimental results show that this method performs well in improving cost savings and reducing risk index, providing strong support for cost control in construction projects. In the future, with the continuous optimization of algorithms and the accumulation of data, it is believed that this method will play a more important role in the field of construction cost estimation and contribute more to the intelligent management of the construction industry.

Funding

The study was supported by the Shaanxi Provincial Department of Education, Architectural acoustic design in the design and renovation of small and medium-sized cultural venues at the county level, 20JK0999.

References

[1] Xia, D. C., & Zhou, X. Y. (2024). Construction Cost control strategy of finished house based on BIM-5D. Journal of World Architecture, 8(4), 15-20. https://doi.org/10.26689/jwa.v8i4.8134

[2] Khalid, S. A. G., Abdullah, M. A., Naif, M. A., Abdulmajeed, A. A., & Abdulmohsen, S. A. (2025). ANN prediction model of final construction cost at an early stage. Journal of Asian Architecture and Building Engineering, 24(2), 775-799. https://doi.org/10.1080/13467581.2023.2294883

[3] Ikda, M., Hepsa, A., Süreyya mre Bykl, Z, D., Bekda, G., & Geem, Z. W. (2023). Estimating construction material indices with arima and optimized narnets. Computers, materials & continua, 74(1 Pt.1),113-129. https://doi.org/10.32604/cmc.2023.032502

[4] Choi, Y., Park, C. Y., Lee, C., Yun, S., & Han, S. H. (2022). Conceptual cost estimation framework for modular projects: a case study on petrochemical plant construction. Journal of Civil Engineering & Management, 28(2), 150-165. https://doi.org/10.3846/jcem.2022.16234

[5] Babaei, M., Rashidi-Baqhi, A., & Rashidi, M. (2022). Estimating project cost under uncertainty using universal generating function method. Journal of Construction Engineering and Management, 148(2), 4021194.1-4021194.10. https://doi.org/10.1061/(ASCE)CO.1943-7862.000223

[6] Duc, L. L., & Kim, A. T. (2024). Enhancing accuracy in cost estimation for faade works: integration of case-based reasoning, random forest, and artificial neural network techniques. Asian Journal of Civil Engineering: Building and Housing, 25(2), 1267-1280. https://doi.org/10.1007/s42107-023-00842-8

[7] Arkhipov, P. O., & Philippskih, S. L. (2022). Building an ensemble of convolutional neural networks for classifying panoramic images. Pattern Recognition and Image Analysis, 32(3), 511-514. https://doi.org/10.1134/S1054661822030051

[8] Khan, A. T., Cao, X., Brajevic, I., Stanimirovic, P. S., Katsikis, V. N., & Li, S. (2022). Non-linear activated beetle antennae search: a novel technique for non-convex tax-aware portfolio optimization problem. Expert Systems with Application, 197(Jul.), 116631.1-116631.10. https://doi.org/10.1016/j.eswa.2022.116631

[9] Lee, J. S. (2023). Quantifying costs of the productivity loss due to schedule changes in construction projects. Engineering construction & architectural management, 30(1), 56-73. https://doi.org/10.1108/ECAM-07-2021-0571

[10] Wu, H., Qian, Q. K., Straub, A., & Visscher, H. J. (2022). Factors influencing transaction costs of prefabricated housing projects in china: developers' perspective. Engineering construction & architectural management, 29(1), 476-501. https://doi.org/10.1108/ECAM-07-2020-0506

[11] Almashaqbeh, M., & El-Rayes, K. (2022). Minimizing transportation cost of prefabricated modules in modular construction projects. Engineering construction & architectural management, 29(10), 3847-3867. https://doi.org/10.1108/ECAM-11-2020-0969

[12] Shrestha, K. (2023). Cost comparison of highway rest area operations: in-house workforce versus outsourcing methods. Journal of Construction Engineering and Management, 149(8), 4023062.1-4023062.11. https://doi.org/10.1061/JCEMD4.COENG-13214

[13] Zhang H., Zhang Q., Yu J. Y. (2022). Analysis and improvement of properties of activation functions in convolutional neural networks. Computer Simulation, 39(4), 328-334. https://doi.org/10.3969/j.issn.1006-9348.2022.04.064

[14] André Luiz Carvalho Ottoni, Novo, M. S., & Costa, D. B. (2023). Hyperparameter tuning of convolutional neural networks for building construction image classification. The visual computer, 39(3), 847-861. https://doi.org/10.1007/s00371-021-02350-9

[15] Atik, S. O., Atik, M. E., & Ipbuker, C. (2022). Comparative research on different backbone architectures of deeplabv3+ for building segmentation. Journal of Applied Remote Sensing, 16(2), 024510-1-024510-18. https://doi.org/10.1117/1.jrs.16.024510

[16] Moghalles, K., Li, H., Al-Huda, Z., & Abdullah, E. (2022). Semantic segmentation of building extraction in very high-resolution imagery via optimal segmentation guided by deep seeds. Journal of Applied Remote Sensing, 16(2), 1-18. https://doi.org/10.1117/1.jrs.16.024513

[17] Zhou, Z. F., Li, H., Feng, Y. X., Lu, J. G., Qian, S. R., & Li, S. B. (2024). Research progress on designing lightweight deep convolutional neural networks. Computer Engineering and Applications, 60(22), 1-17. https://doi.org/10.3778/j.issn.1002-8331.2404-0372

[18] Ye, K. T., Shu, L. L., Li, W., & Hou, C. J. (2023). A beetle antenna search algorithm based on differential evolution strategy and its application. Computer Engineering & Science, 45(05), 920-930. http://joces.nudt.edu.cn/EN/Y2023/V45/I05/920

[19] Wang, C. Y., Liu, X., & Chaomurilige. (2024). A daptive and parallel beetle antennae optimization algorithm. Journal of Shandong University (Engineering Science), 54(05), 74-80. https://doi.org/10.6040/j.issn.1672-3961.0.2023.262

[20] Kumar, B. S., Santhi, S. G., & Narayana, S. (2022). Optimal energy-delay scheduling using improved beetle antennae search (bas) for energy-harvesting wsns. Wireless personal communications: An Internaional Journal, 126(3), 2533-2556.

https://doi.org/10.1007/s11277-022-09828-2