www.cad-journal.net
Computer-Aided Design AND APPLICATIONS

Computer-Aided Architectural Image Recognition and Optimization Based on Deep Convolution Neural Network

Haohao Guo[1], Yang Lin[2], Qing Yang[3], Minhui Sun[4] and Peng Wang[5]

[1,3,4,5]China Construction Sixth Engineering Bureau Corp., Ltd 310000, China,
[1]guojiaoshuo163@163.com, [3]13776139663@163.com, [4]15558191962@163.com,
[5]18272742267@163.com
[2]China Construction 6th Engineering Bureau 8th Corp. Limited 310000,China,
huanghaocscec@163.com

Corresponding author: Minhui Sun, 15558191962@163.com

Abstract. This paper aims to construct and optimize a DCNN (Deep convolution neural network) model for architectural image recognition and optimization. This paper deeply analyzes the characteristics and recognition requirements of architectural images and then constructs a model suitable for architectural image recognition and optimization based on DCNN. The model employs multi-level convolution operations to extract feature information from the image, subsequently utilizing the fully connected layer for feature classification and optimization. Through a lot of training and optimization, the model can accurately identify the key elements in architectural images and optimize them effectively. Experiments show that the DCNN model performs well in recognition accuracy, RMSE (Root Mean Square Error), and response time, which effectively improves the processing efficiency and quality of architectural image information. This shows that the model has high application value. Future research will focus on improving the model's generalization ability, exploring more advanced NN (Neural network) structure and optimization algorithms, and applying the model to more practical scenarios to promote the sustainable development of this field.

Keywords: Architectural Image Recognition; Optimization Model; Deep Convolution Neural Network; Deep Learning
DOI: https://doi.org/10.14733/cadaps.2025.S1.90-103

1 INTRODUCTION

Amidst the rapid advancements in science and technology, computer vision technology has found its applications across diverse industries. Notably, architectural image recognition and optimization occupy a pivotal position in architectural design and urban planning fields. With the acceleration of global urbanization, more and more residents are living in noisy urban environments. Bar et al. [1] proposed an innovative method that combines acoustic analysis with parametric landscape design, incorporating architectural image recognition technology to expand the analysis. Next, using architectural image recognition technology, we can quickly and accurately identify key architectural and landscape features in the urban environment, which will serve as a reference for subsequent landscape design. This method first collects noise data online and on-site to comprehensively understand the distribution and characteristics of noise. Although modern architectural design has begun to widely use parametric tools to improve indoor acoustic environments, research on using digital technology to alleviate noise through landscape and urban design is still relatively limited. Especially, further exploration is needed to optimize the urban noise environment using ground morphology design. The design of these landscape structures will be based on parametric methods to ensure the efficiency and flexibility of the design. To verify the effectiveness of the method, they conducted practical applications at a case study site near Munich Airport. The acoustic performance of these design schemes was compared and evaluated using acoustic simulation software. Through architectural image recognition technology, it quickly obtained detailed architectural and landscape information of the research area, and based on this, developed three different park terrain landscape designs. These architectural images encompass abundant design elements, intricate details, spatial information, and structural attributes. The role of digital tools in the architectural design process is becoming increasingly prominent, not only becoming a supporting element for complex structural design but also providing unprecedented opportunities for designers by combining architectural image recognition technology. Ceylan [2] investigated their level of understanding of architectural image recognition technology. Digital media not only provides architects with more possibilities to execute the design process, but the integration of architectural image recognition technology has also pushed this process to new heights.

The design space for solar energy utilization and other specific design goals that may be obtained from image recognition, such as architectural style, facade materials, etc. In addition, combined with architecture, providing inspiration and reference for future design work [3]. Conventionally, architectural image recognition and optimization techniques relied heavily on manual assessments and expert opinions, posing challenges such as excessive workload, sluggish efficiency, and marked subjectivity. At present, with the continuous advancement of technology, complex analysis, design tools and technologies, as well as data from multiple related disciplines, are becoming increasingly abundant. Fricker et al [4] aimed to engage in in-depth discussions to capture and understand environmental characteristics, thereby designing landscape architecture that is more in line with the natural environment and human needs. Such as architectural form, materials, spatial layout, etc., provide landscape architects with new perspectives and tools. They provide valuable information and inspiration for the innovation of landscape design methods. Building image recognition technology automatically extracts and analyzes key information from images. It is worth noting that advanced technologies such as architectural image recognition have not yet been widely accepted and integrated into landscape architecture design schools. Landscape architecture occupies a core position in landscape design, shaping how we interact with the natural environment and promoting more resilient and adaptable environmental construction. Especially under the urgent challenges brought about by the current climate crisis. In addition to traditional methods, tools, and techniques, this article will particularly emphasize the potential role of architectural image recognition technology in landscape architecture design education. However, the advent of DL (Deep Learning) technology, coupled with the triumphant utilization of DCNN in image recognition, has ushered in novel approaches for architectural image recognition and optimization. Computer-aided design (CAD), is a core tool for modern design and planning, digital large-scale topographic maps, urban planning, and landscape architecture. To overcome these difficulties, Habib and Pradhan [5] not only focus on developing a cost-effective tool but also combine building image recognition technology further to improve the efficiency of data acquisition and processing. This method utilizes a two-dimensional conformal polynomial model and a least-squares fitting algorithm to achieve coordinate transformation between two different grid references. Through image recognition, key features of buildings and terrain can be quickly and accurately extracted from drone aerial photographs, satellite images, or on-site photos. In addition, in order to automate and simplify the operation of the model, we performed polynomial coefficient calculations in the Microsoft Visual Studio environment. The key to this method lies in its ability to reduce reliance on expensive systems and professionals while ensuring feature quality and positional accuracy. These extracted data can be directly used in CAD systems as the basis for establishing or updating topographic maps.

DL models are capable of automatically discerning image features and, through extensive data training and refinement, can accurately recognize and optimize architectural images. This technological breakthrough has brought about transformative changes in the realm of architectural design and urban planning. Lee et al. [6] delved into the use of computer technology and how it has influenced the development of landscape architecture in South Korea over the past thirty years, particularly with the innovative impact of incorporating architectural image recognition technology. In the 1990s, landscape architects began to view computers as a revolutionary tool for landscape research, planning, and design. During this period, landscape architects worked closely with experts in other fields to drive technological innovation. Although it has brought great convenience in data processing and computation, its potential in creative exploration and formal generation has not been fully tapped. They not only developed specialized computer software programs but also explored various landscape analysis and design techniques as computer programmers. However, in early applications, landscape designers primarily viewed computer technology as a tool to replace traditional practical cases. In the 2000s, there were significant changes in the application of computer technology in the field of landscape architecture. Maps and graphic technology have become important visualization tools in the process of landscape beautification, providing designers with more intuitive and accurate visual expression [7]. In addition, the development of graphic software has made the realistic representation of perspective views more crucial. At this time, landscape architects are more like graphic designers, which combines architectural image recognition technology to address the issue.

The aim of this study is to address the aforementioned issues and achieve precise recognition and optimization of architectural images through the development of a computer-assisted model for architectural image recognition and optimization, leveraging DCNN. The specific research contents include:

(1) Build an efficient and accurate building image recognition model to realize automatic recognition and classification of building elements.

(2) Design an effective architectural image optimization algorithm that can intelligently improve the architectural image according to the design requirements and user preferences.

(3) Investigate the performance of DL models in architectural image recognition and optimization, along with potential optimization strategies, to offer theoretical backing and technical direction for real-world implementations.

The primary innovations presented in this paper are as follows:

(1) Introduction of a DCNN-based model for building image recognition, capable of automatically learning feature representations within building images, thereby achieving precise identification of building elements.

(2) An architectural image optimization algorithm based on a generation countermeasure network is designed, which can automatically generate the required architectural images according to the user's requirements and design principles.

(3) The performance optimization strategy of the DL model in architectural image recognition and optimization is deeply explored, which provides a more efficient and accurate technical means for practical application.

In this study, the DL method is used as the main technical means, and a computer-aided architectural image recognition and optimization model based on DCNN is constructed by combining the relevant knowledge in the field of architectural design and computer vision.

The chapter structure of this paper commences with an introduction, outlining the research context and importance. Subsequently, the DCNN model and its theoretical foundation are presented. This is followed by the construction of an image recognition and optimization model rooted in DCNN. Next, the design and execution of the experiments are elaborated upon. The model's practical application is demonstrated through specific case studies. Lastly, the paper concludes by summarizing the findings, identifying the model's constraints, and anticipating future advancements and research paths.

2 RELATED WORK

The complexity of modern science is increasing day by day, accompanied by the emergence of many previously unforeseeable connections between different themes, marking the flourishing development of interdisciplinary science. By analyzing these two knowledge areas and combining professional evaluations from optical and architectural experts. This method aims to utilize advanced computing technology to provide designers in two fields with more efficient and accurate design tools. Livshits et al. [8] proposed an innovative approach to establish heuristic algorithms for computer-aided design between optical design and architectural design. In this context, it is noted that the seemingly distinct fields of optics and architectural design exhibit surprising similarities in certain aspects. This system can not only automatically process a large amount of design data, but also provide personalized design suggestions for designers. BIM technology, with its powerful data integration capabilities and visualization effects, is gradually changing the traditional way of architectural design. By combining our heuristic algorithms with BIM technology, the scope of design can be further expanded from architectural design to optical design, and even more in other related fields. It is worth mentioning that building image recognition, as an important component of BIM technology, provides us with new perspectives and tools.

Image-based 3D reconstruction of ancient Chinese architecture is indeed a complex task in the field of image modelling, and image-matching technology is crucial for constructing accurate 3D models. Nie et al. [9] proposed a new method that combines building image recognition. Considering the unique structural style and rich historical connotations of ancient Chinese architecture. In architectural image recognition, they introduced local region similarity analysis to further distinguish and identify different elements in the building. Based on this similarity analysis, the marked grid connection areas are defined as clusters, each representing a complete building element or structural part. Based on the different densities of key points, we can determine the anchor unit (the area with the highest density of key points), adjacent units (the area adjacent to the anchor unit), and boundary units (the area located at the edge of the image). Extract key points from building images using the SIFT algorithm, which represents prominent features in the image, such as the corners of the roof and the boundaries of window frames. By comparing the feature vectors of key points in adjacent grids, it can be determined whether they belong to the same building element, such as roofs, walls, or windows. After completing the recognition of architectural elements in the image, we use the nearest neighbour distance ratio (NNDR) algorithm to match key points within similar clusters. The experimental results indicate that The FM_GMC-Arch method can significantly improve the feature-matching performance between images. This method compares the distance and angle relationships between key points in different images to find the most matching pair of key points, thereby achieving accurate matching between different images.

Technological media has had an unprecedented impact on the field of landscape design, especially the widespread application of comprehensive software in landscape design. Song and Jing [10] combine architectural image recognition technology to briefly introduce CAD The application process of SketchUp and PS in landscape planning and design. From drawing CAD floor plans to building SketchUp models, and then to Photoshop (PS) processing of renderings. Architectural image recognition technology can automatically extract key information from images, such as building form, materials, spatial layout, etc., providing more accurate data support for landscape designers. These technologies not only enable landscape designers to have more intuitive and in-depth creative ideas but also become the best way to showcase landscape design. In the construction of SketchUp models, image recognition technology can assist designers in converting two-dimensional images into three-dimensional models, achieving more realistic preview effects. By identifying elements such as colour, texture, and lighting in an image, designers can more accurately adjust the image effect to better match the actual scene and design requirements. By digitizing design images and converting them into standard formats, designers can more conveniently share design results with other team members, achieving efficient collaboration. In the drawing of CAD floor plans, image recognition technology can help designers quickly identify and extract building contours and terrain information from images, thereby accelerating the design process and improving accuracy. The importance of protecting historical heritage has long been emphasized by global consensus, and many countries have invested significant resources to achieve this goal. The current data on historical heritage archiving mainly focuses on the physical characteristics of buildings, while insufficient attention is paid to the spatial perception and visual experience of tourists. Therefore, using drawings, photos, and advanced digital technology to record these heritage buildings is particularly crucial for the archiving and protection of cultural heritage. Historical buildings often face the risk of damage caused by natural or human factors. In their research, Tai and Sung [11] explored the hypotheses of sequence and maximum information, aiming to determine the normative perspective of architectural space. With the advancement of digital imaging technologies such as photogrammetric modelling, the measurement accuracy of architectural heritage has been significantly improved. Through behaviour mapping, we can identify the scenes that tourists watch the most frequently. Spatial syntax analysis helps us understand the importance of these scenes in the spatial structure of architecture. By combining this information with the results of behavioural mapping and spatial syntax analysis, it can more accurately capture and record the visual perception and experience of tourists in the architectural space. The results of perception research show that when participants are asked to recall architectural spaces, the most frequently viewed scenes identified through behaviour mapping and architectural image recognition techniques are the most representative.

Tastan et al. [12] Analyzed the role of architectural image recognition technology in improving modelling efficiency and accuracy. Explored user preferences for different interaction methods and virtual environment settings during the modelling process, particularly in interface design combined with architectural image recognition technology. This method can use a new keyboard for numerical input, and utilize more advanced mobile technology to improve modeling efficiency and accuracy. The research results indicate that the DM modelling method combined with building image recognition technology has significant advantages in the process of building modelling. This includes automatically identifying building elements, extracting dimensions and proportions, and assisting in creating accurate 3D models. Especially when the interface is combined with building image recognition technology, how to optimize the interface to improve user operation efficiency and satisfaction. The user-friendly nature of technology is crucial for its widespread application. By optimizing the interface and algorithms, the complexity of user operations can be reduced, making it easy for non-professionals to use. At the same time, provides rich data visualization tools to help users better understand and analyze identification results. Ensuring data security and user privacy is crucial when processing large amounts of architectural image data. Encryption technology and access control policies are needed to protect data. Yu et al. [13] not only used computer parameterization technology to model the design object but also combined architectural image recognition technology to enhance the accuracy and efficiency of the design. Furthermore, it systematically proposes a dynamic environmental landscape parameterization design method that combines architectural image recognition. Building image recognition technology can automatically extract key features such as the shape, size, and texture of building elements. By combining architectural image recognition technology, designers can adjust and optimize design parameters in real time. In landscape design, using the optimization parameter method to simulate the system dynamics process of dynamic changes in the environment has faster speed and higher accuracy. These features provide accurate data support for subsequent parametric design, making the design results more in line with the needs of actual building scenarios. This is because the optimization parameter method can more accurately describe the patterns and trends of environmental changes, thereby generating more realistic simulation results. By identifying the features in actual building scenes, designers can modify parameter values specifically to better match actual needs. In environmental landscape design, optimizing parameters is usually aimed at improving the visual effect, ecological sustainability, or functionality of the design. By adjusting the pseudo velocity and uniformity of the parameter curve, these optimization objectives can be achieved. Parameter optimization can adopt various methods, such as genetic algorithms, simulated annealing, particle swarm optimization, etc. Taking terrain simulation in landscape design as an example, optimizing the pseudo velocity and uniformity of parameter curves can generate terrain models that are more in line with natural terrain characteristics. These methods iteratively search for the optimal solution and continuously adjust parameter values to achieve optimization goals. At the same time, combined with architectural image recognition technology, key features in the terrain (such as mountains, rivers, etc.) can be automatically extracted, providing accurate data support for parameter optimization.

Colour plays a crucial role in landscape design, not only enhancing the artistic and aesthetic aspects of the design but also generating deep resonance with the audience through emotional guidance. Based on a computer-aided collaborative design system, Zhang and Deng [14] established a landscape design model. Starting from the colour extraction of landscape architecture, with the help of architectural image recognition technology, it can efficiently and accurately identify and extract key colour elements in the building. Secondly, they conducted in-depth research on the emotional structure of the environment and explored the emotional guidance role of colour in street space design. On this basis, combined with architectural image recognition technology. This scale quantifies the emotional effects of different colour combinations in landscape design, providing designers with a more scientific basis for colour matching. Furthermore, we used the SD method (semantic difference method) for colour space perception evaluation, combined with colour data obtained from architectural image recognition technology, to generate a scale for colour emotional effects. Through architectural image recognition technology, we can capture and analyze the practical application effects of different colour combinations in street space, as well as how they affect the emotional experience of the audience. Architectural landscape carries multiple functions such as ecology, culture, society, and aesthetics in modern urban development. Scientific and reasonable organization of garden landscapes plays an irreplaceable role in improving urban quality. Zhao [15] proposed an innovative method that combines architectural image recognition technology to achieve efficient and accurate construction and display of garden landscapes. These features not only include the appearance and size of the building, but also the type, distribution, and density of vegetation. Next, it uses 3D CAD technology to construct accurate garden vegetation and build models based on image recognition extracted data. And high workload in constructing garden vegetation and architectural models. Using architectural image recognition technology to extract key architectural and landscape features from a large number of real-life photos or high-definition images captured by drones. Through image recognition, we can quickly and accurately obtain the basic data required to build a model, greatly reducing the workload of manual input. Then, we use virtual reality (VR) and immersive technology to visually display the constructed garden landscape model. These technologies can dynamically adjust the model's precision and rendering effects based on the audience's perspective and distance, thereby improving rendering efficiency and real-time rendering effects while ensuring visual effects.

3 DCNN THEORETICAL BASIS

3.1 DCNN Foundation

NN, short for Neural Network, is a mathematical model designed to mimic the way a biological nervous system processes information. Comprised of numerous interconnected neurons, it processes input data and generates output by adjusting the connection weights and activation functions among these neurons. The core concept behind NN is its ability to automatically fine-tune network parameters using training data, allowing the network to approximate a specific input-output mapping closely. This mapping can represent various functions, including classification, regression, or more intricate relationships. Typically, a NN's structure encompasses an input layer, one or more hidden layers, and an output layer. Refer to Figure 1 for a basic visualization of a NN model.

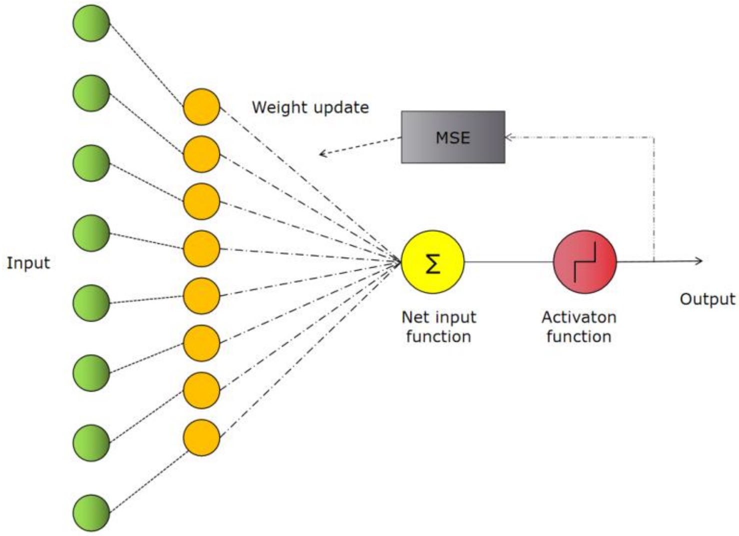

Figure 1: NN model diagram.

CNN is a specialized NN architecture tailored for image data processing. Unlike conventional NN, CNN utilizes local connectivity and weight sharing, significantly decreasing network parameter counts and enhancing computational efficiency. Convolutional layers employ convolution kernels to extract local features from the input image, yielding a feature map. Pooling layers then downsample this feature map, reducing its dimensionality while preserving essential features. Following numerous convolution and pooling operations, the final prediction is ultimately output. A simple convolution operation can be expressed as:

$$Output[i][j] = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} Conv\_weights[m][n] \cdot Input[i+m][j+m] + bias[i][j] \tag{1}$$

Among them:

$Output[i][j]$ is a pixel value on the feature map after the convolution operation.

$Conv\_weights[m][n]$ is a weight on the convolution kernel.

$Input[i][j]$ is a pixel value on the input image.

$bias[i][j]$ is an offset.

M and N are the moving steps of the convolution kernel in the height and width of the input image, respectively.

CNN exhibits evident strengths in the realm of image recognition. Firstly, CNN possesses the capability to autonomously acquire feature representations from images, eliminating the need for manual feature extractor design. Secondly, by utilizing local connectivity and weight sharing, CNN minimizes network parameters and enhances computational efficiency. Lastly, CNN's hierarchical structure facilitates the extraction of abstract features from images in a layered manner, enabling a profound comprehension of the image content.

3.2 The Classic Model of CNN

Throughout the evolution of CNNs, numerous iconic models have emerged, leaving a significant impact on the field of image recognition. Listed below are some exemplary CNN models:

AlexNet stands out with its 8 learning layers, encompassing 5 convolution layers and 3 fully connected layers. To prevent overfitting, it incorporates the ReLU activation function, Dropout, and LRN. AlexNet's success underscores the remarkable capabilities of DCNN in image recognition.

VGGNet, on the other hand, constructs a deeper network architecture by stacking multiple 3×3 convolution kernels. Building image recognition and optimization model construction

3.3 Model Design

When constructing the model of building image recognition and optimization, this study chooses DCNN as the basic structure, which gives full play to its powerful feature learning and representation ability. The model consists of multiple levels, forming an efficient and accurate processing flow. Initially, the input layer assumes the role of receiving the preprocessed architectural images, guaranteeing the quality and uniformity of the image data. Thereafter, the convolution layer employs multiple convolution kernels to capture the local features inherent in the image. These convolution kernels, like feature detectors, can capture key information such as edges, textures and shapes in the image. Then, the pool layer downsamples the feature map, reducing the data dimension and retaining the main features, while enhancing the robustness of the model to image changes. The fully connected layer further flattens the feature map into a one-dimensional vector, learns the global feature representation, and finally outputs the recognition results or optimized parameters through the output layer. The structure of the DCNN model is shown in Figure 2.

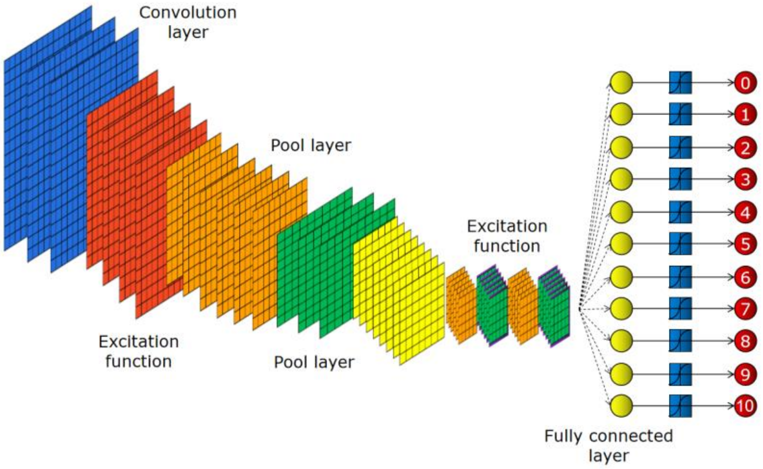

Figure 2: DCNN model structure.

The DCNN model adopts a specific filtering algorithm in the training process, and describes the structural risk minimization function of the model through a specific formula:

$$\beta = \min \sum_{i=1} f \ x \ - y^{2} + \alpha \left\| \theta \right\|^{2} \tag{2}$$

This formula serves as a guideline for optimizing model parameters during training, aiming to minimize both prediction error and structural risk, thereby enhancing the model's generalization capability and overall performance.

When the image data is input to the DCNN model, it will undergo a linear transformation process, which can be specifically described by the following formula:

$$\partial = \sum_{i} \beta_i \omega \ x_i \tag{3}$$

$$\theta = \theta \;\; x_i, x_j \tag{4}$$

$$f \;\; y \;\; = \sum \beta_i \theta \;\; y, x_i \tag{5}$$

In this paper, the input features are linearly transformed to obtain a fractional vector. Then, this score vector will be normalized by the Softmax function and transformed into a probability distribution. The Softmax function converts each element of the score vector into a probability value, which indicates the possibility that the input sample belongs to the corresponding category:

$$y \;\; x_i \;\; = \frac{\exp \;\; x_i}{\sum_{i=1}^{M} \exp \;\; x_i} \tag{6}$$

Where x is the eigenvector output by the fully connected layer; M is the number of categories of classification.

When designing the model, choosing suitable parameters is crucial to guarantee its performance and ability to generalize. Parameters such as the size and quantity of convolution kernels, step size, filling technique, pooling type, pooling window size, and step size are all important considerations. After conducting numerous experiments and fine-tuning, we have identified the optimal parameter combination tailored for the task of building image recognition and optimization. These optimal parameters are detailed in Table 1.

Parameter type	*Parameter name*	*Numerical value*
Convolution layer	Convolution kernel size	5*5
	Number of convolution kernels	128
	Stride	2
Pool layer	Pool type	Max Pooling
	Pool window size	2*2
	Pool step size	2
Other	Activation function	ReLU
	Optimizer	Adam
	Learning rate	0.001
	Batch Size	64
	Epochs	100
	Loss function	Cross entropy

Table 1: Parameter list of building image recognition and optimization model.

The model design in this paper is based on the basic principle of DL and the characteristics of CNN. By introducing advanced technologies such as residual connection and batch normalization, the stability and convergence speed of the model are improved. In addition, according to the characteristics of architectural images, this study also designed a specific convolution kernel and pooling method to extract the feature information from architectural images better. These innovations make the model of this study achieve excellent performance in the task of building image recognition and optimization.

3.4 Model Training and Optimization

$$v_t = \mu \cdot v_{t-1} - \eta \cdot \nabla_\theta J \;\; \theta \tag{7}$$

$$\theta_{t+1} = \theta_t + v_t \tag{8}$$

To avoid model overfitting and enhance generalization capabilities, this study incorporates the L2 regularization technique. L2 regularization punishes the larger weight value by adding the sum of squares of weights to the loss function, which makes the model pay more attention to global features than local noise in the training process:

$$loss = -\sum_i \sum_j y_i^j \log \hat{y}_i^j + \lambda \left\| \theta \right\|^2 \tag{9}$$

Here, y represents the expected value, $\hat{y}$ denotes the predicted value, λ signifies L2 regularization, and θ stands for the parameter set of the neural network NN. By choosing the appropriate regularization coefficient λ, the model performance can be guaranteed and the risk of over-fitting can be reduced.

4 EXPERIMENTAL DESIGN AND RESULT ANALYSIS

4.1 Construction of Experimental Environment

Before the experiment of building an image recognition and optimization model, this paper first needs to build a suitable experimental environment. This includes selecting the appropriate hardware equipment and configuring the corresponding software environment.

Hardware environment: The hardware environment needed for the experiment mainly includes high-performance computers or servers equipped with powerful graphics processors to speed up the training of the DL model.

The recognition accuracy stands as a crucial benchmark for assessing an image recognition model's performance. Refer to Figure 3 for a visual representation of the DCNN model's recognition accuracy across various datasets.

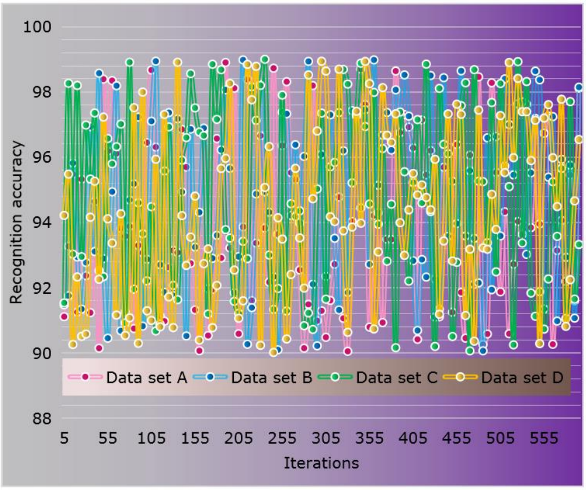

Figure 3: Recognition accuracy of the DCNN model on different data sets.

The horizontal axis in Figure 3 represents the number of iterations, and the vertical axis represents the recognition accuracy. In Figure 3, we can see the variation in recognition accuracy of the DCNN model with the number of iterations on four different datasets (datasets A, B, C, and D). Each curve represents the change in recognition accuracy on a dataset. Usually, as the number of iterations increases, there is a significant fluctuation in the recognition accuracy of the model, but the recognition accuracy is always at a relatively stable level of 90% -98%. The DCNN model has improved the recognition accuracy, which shows that the model has advantages in image recognition tasks.

Figure 4 shows the RMSE of the DCNN model on different data sets. By observing the numerical value and changing trend of RMSE, we can understand the error distribution of the model on different data sets.

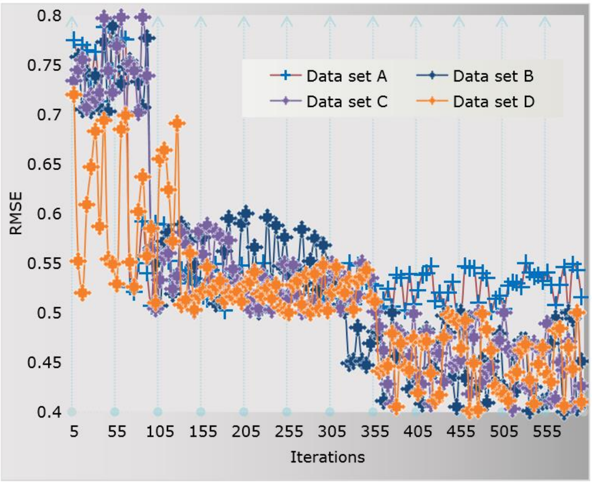

Figure 4: RMSE of DCNN model on different data sets.

The horizontal axis in Figure 4 represents the number of iterations, while the vertical axis in Figure 4 represents the root mean square error. RMSE is a commonly used indicator to measure the difference between predicted and actual values, reflecting the accuracy of predictive models. The smaller the RMSE value, the closer the predicted value of the model is to the actual value, indicating better model performance. Figure 4 shows the trend of RMSE variation when the DCNN model is trained on different datasets. By observing the curves in the graph, we can understand the error distribution of the model on different datasets. The RMSE value gradually decreases and stabilizes as the number of iterations increases, indicating that the model has good predictive performance and strong predictive ability on this dataset.

In the task of building image recognition and optimization, a short response time can ensure that the model can process a large number of image data in time. Figure 5 shows the response time of the DCNN model to different datasets.

The horizontal axis in Figure 5 represents the transaction set, and the vertical axis represents the response time. Figure 5 shows the response time of the DCNN model on different transaction sets (or datasets). By observing the bar or line graphs in the graph, we can understand the processing speed of the model on different datasets. The shorter response time of dataset C as shown in the figure indicates that the model has a higher processing efficiency on this dataset. The response time of dataset A is relatively long, so it is necessary to consider optimizing the model or dataset to improve processing speed.

5 APPLICATION CASE OF ARCHITECTURAL IMAGE RECOGNITION AND OPTIMIZATION

In order to show the effect of building image recognition and optimization technology in practical application, this paper takes intelligent building design as an example for case analysis. In this case, we use the DCNN model constructed in the previous chapter to assist the intelligent building design. Firstly, this paper collects a large number of architectural image data sets and uses data enhancement technology to preprocess them. Then, the DCNN model is used to extract and classify architectural images, and different types of architectural styles and elements are identified.

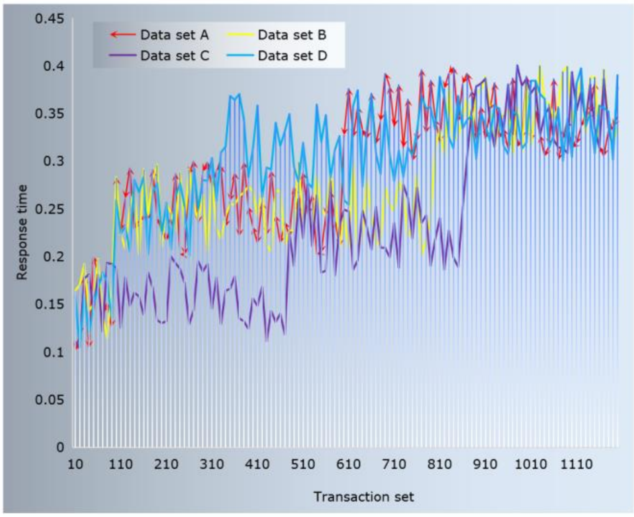

Figure 5: Response time of DCNN model on different data sets.

Then, according to the user's needs and design principles, these features and classification results are used to generate the initial architectural design scheme. After generating the initial design scheme, this paper uses the architectural image optimization technology to further adjust and optimize the scheme. Through modifications to the building's configuration, hues, materials, and other variables, we have produced numerous potential design plans. Ultimately, we utilize evaluation criteria to assess and compare these candidate designs, selecting the optimal plan (illustrated in Figure 6). The design score before and after optimization is shown in Figure 7.

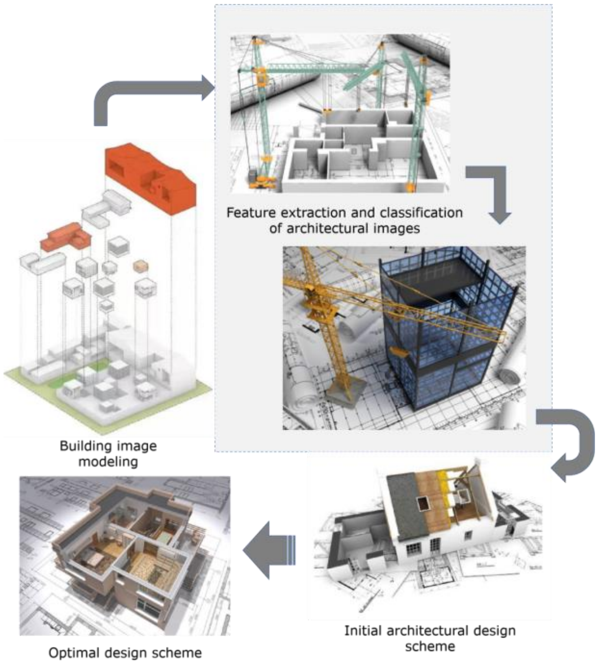

Figure 6: Design scheme display.

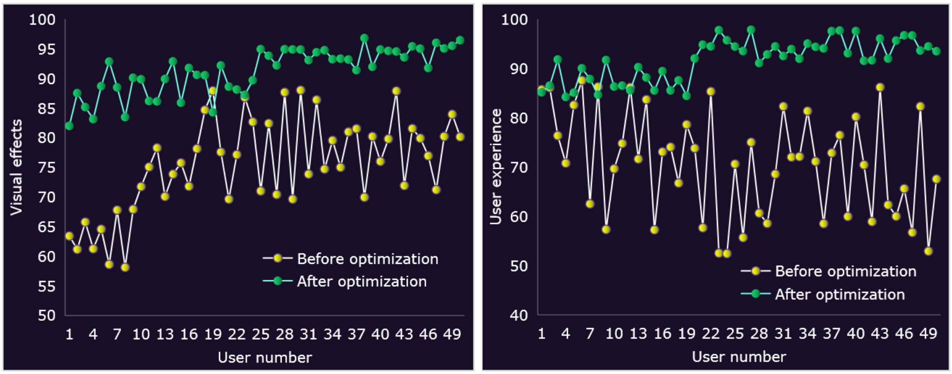

Figure 7: Design scheme score before and after optimization.

The horizontal axis in Figure 5 represents the number of users, which refers to the number of users participating in the evaluation of architectural design schemes. The vertical axis represents visual effects and user experience, which is a comprehensive evaluation indicator. Used to measure the performance of architectural design schemes in terms of visual appeal and user experience. Figure 5 shows the scores of different design schemes in user evaluation before and after using the DCNN model for image recognition and optimization in intelligent building design. There are two curves in the graph, representing the scores of the design scheme before and after optimization. By comparing the positions and trends of these two curves, we can clearly see whether the optimized design scheme has improved in terms of visual effects and user experience. The position of the optimized curve on the vertical axis is generally higher than that of the pre-optimized curve, indicating that the optimized design scheme has achieved higher scores in evaluation, i.e. improved visual effects and user experience.

6 CONCLUSIONS

This article proposes a key point extraction method for building images based on the traditional algorithm DCNN model to address the aforementioned issues between building images. At present, traditional algorithms and DCNN models are used to handle the relationship between images, but existing DCNN models and traditional algorithms extract image features from different perspectives. The foundation of the research is to handle the relationship between extracted image features, in order to solve problems in fields such as image recognition and image retrieval. Using a single method alone cannot provide a comprehensive feature description of an image, and how to enhance the contextual information in the image is also a key issue. The information in images can be transformed into a simple and understandable form, making it easier for people to solve problems in the field of images. In the field of image keypoint extraction, architectural images cannot use a single method to improve the accuracy of matching due to the complexity of the information they contain. On this basis, an attention mechanism was added to improve accuracy. The model exhibits extraordinary generalization ability in handling complex and ever-changing architectural images, making it highly practical in real-world applications.

Haohao Guo, https://orcid.org/0009-0009-1328-7040
Yang Lin, https://orcid.org/0009-0001-2798-7410
Qing Yang, https://orcid.org/0009-0009-9891-2131
Minhui Sun, https://orcid.org/0009-0001-6896-9219
Peng Wang, https://orcid.org/0009-0008-1968-4170

REFERENCES

[1] Bar, S.-K.-L.; Shaked, T.; Fettahoglu, E.-S.; Krimm, J.; Boucsein, B.: Embedding acoustic analysis in landscape architecture design processes: a case study of Munich airport, Buildings, 13(1), 2023, 143. https://doi.org/10.3390/buildings13010143

[2] Ceylan, S.: A case study on the change of students' perception of architectural design based on their knowledge of digital tools, The International Journal of Design Education, 14(2), 2019, 1-16. https://doi.org/10.18848/2325-128X/CGP/v14i02/1-16

[3] Chen, K.-W.; Choo, T.-S.; Norford, L.-K.: Enabling algorithm-assisted architectural design exploration for computational design novices, Computer-Aided Design and Applications, 16(2), 2019, 269-288. https://doi.org/10.14733/cadaps.2019.269-288

[4] Fricker, P.; Hayek, U.-W.; Monacella, R.: Digital landscape architecture education–where do we stand and where should we go? Journal of Digital Landscape Architecture, 2023(8), 2023, 576-584. https://doi.org/10.14627/537740060

[5] Habib, M.; Pradhan, B.: A low-cost spatial tool for transforming feature positions of CAD-based topographic mapping, Geodesy and Cartography, 45(4), 2019, 161-168. https://doi.org/10.3846/gac.2019.10322

[6] Lee, M.-J.: Transformation of discourse on uses of computer technology in Korean landscape architecture-focused on Journal of the Korean Institute of Landscape Architecture and Environmental & Landscape Architecture of Korea, Journal of the Korean Institute of Landscape Architecture, 48(1), 2020, 15-24. https://doi.org/10.9715/KILA.2020.48.1.015

[7] Liu, C.; Lin, M.; Rauf, H.-L.; Shareef, S.-S.: Parameter simulation of multidimensional urban landscape design based on nonlinear theor, Nonlinear Engineering, 10(1), 2022, 583-591. https://doi.org/10.1515/nleng-2021-0049

[8] Livshits, I.-L.; Glebovskyi, A.-S.; Protsuto, M.-V.: Interdisciplinary approach for simulation of starting points for optical and architectural design, Advanced Optical Technologies, 8(2), 2019, 135-144. https://doi.org/10.1515/aot-2018-0062

[9] Nie, Y.; Hu, L.; Zhang, J.: Feature matching based on grid and multi-density for ancient architectural images, Journal of Computer-Aided Design and Computer Graphics, 32(3), 2020, 437-444. https://doi.org/10.3724/SP.J.1089.2020.17835

[10] Song, Y.; Jing, Y.: Application prospect of CAD-SketchUp-PS integrated software technology in landscape planning and design, Computer-Aided Design and Applications, 18(S3), 2020, 153-163. https://doi.org/10.14733/cadaps.2021.S3.153-163

[11] Tai, N.-C.; Sung, L.-W.: Digital archiving of perceptual experiences of an architectural space with computer-aided methods, Computer-Aided Design and Applications, 17(3), 2019, 585-597. https://doi.org/10.14733/cadaps.2020.585-597

[12] Tastan, H.; Tong, T.; Tuker, C.: Using handheld user interface and direct manipulation for architectural modeling in immersive virtual reality: An exploratory study, Computer Applications in Engineering Education, 30(2), 2022, 415-434. https://doi.org/10.1002/cae.22463

[13] Yu, C.; Altalbe, A.; Du, N.: Environmental landscape art design using dynamic nonlinear parameterization, Fractals, 30(02), 2022, 2240077. https://doi.org/10.1142/S0218348X22400771

[14] Zhang, M.; Deng, X.: Color effect of landscape architecture design under computer aided collaborative design system, Computer-Aided Design and Applications, 19(S3), 2021, 13-22. https://doi.org/10.14733/cadaps.2022.S3.13-22

[15] Zhao, X.: Application of 3D CAD in landscape architecture design and optimization of hierarchical details, Computer-Aided Design and Applications, 18(S1), 2020, 120-132. https://doi.org/10.14733/cadaps.2021.S1.120-132