WARSAW UNIVERSITY OF TECHNOLOGY Index 351733
FACULTY OF CIVIL ENGINEERING
COMMITTEE FOR CIVIL AND WATER ENGINEERING
POLISH ACADEMY OF SCIENCES ISSN 1230-2945

DOI: 10.24425/ace.2025.154126
ARCHIVES OF CIVIL ENGINEERING
Vol. LXXI ISSUE 2 2025
pp. 363–376

Research paper

Architectural design methods based on image recognition technology and virtual VR

Feng Li[1]

Abstract: To enhance the conceptual expression of architectural design and enhance the interactive experience, and to strengthen the application of computer technology in architectural design drawings, this study explores architectural design methods on the ground of image deep learning image recognition technology and augmented reality technology. The experiment demonstrated that when the accuracy of the improved You Only Look Once version 4 (YOLOv4) model was 0.9, the recall rate was 0.98, and the curve area was 0.93. The model loss function curve converged to the minimum value of 0.04, with the fastest convergence speed and the highest model recognition efficiency. Its time consumption was decreased by as much as 70.06%, indicating better overall performance. Meanwhile, the clustering strategy design of the model was relatively optimal, with the highest values of purity, standard mutual information, and Rand coefficient reaching 0.944, 0.931, and 0.942, respectively. In practical analysis of architectural design, the average accuracy and intersection over union of the improved YOLOv4 model confirmed the good detection performance of this method. The application of virtual reality technology in building information models has significantly improved the visualization delay rate, and the subjective evaluation of users was relatively high. The combination of visible image recognition and augmented reality can achieve intelligent processing and application of drawing information, improve design efficiency and quality, and optimize design experience.

Keywords: virtual reality, deep learning, YOLOv4, image recognition, architectural design

[1]Prof., MSc., Shanxi Vocational University of Engineering Science and Technology, College of Architecture and Design, 030619 Jinzhong, China, e-mail: jzzsxlf69@163.com, ORCID: 0009-0008-7784-4566

1. Introduction

The traditional architectural design methods mainly include hand drawn images and computer-aided design. It has low efficiency in transmitting design information, insufficient expression of design information, and less intuitive design decisions [1]. With the progress and development of technology, the architectural design is gradually transitioning towards digitalization and informatization. At present, architectural design work is on the ground of Building Information Modeling (BIM) technology, Virtual Reality (VR) and Augmented Reality technology, as well as automation and intelligent tools [2]. Compared to traditional design techniques, the VR technology can present the BIM in the form of VR, which facilitates designers and clients to better exchange opinions, adjust design plans, and improve efficiency and accuracy. VR technology improves the real-time updating ability and data presentation ability of BIM [3, 4]. Faced with some complex architectural design schemes, such as a large number of architectural design drawings, complex building structures, and cumbersome design projects, Compared with directly using AutoCAD or BIM data, Image Recognition (IR) can more quickly identify components in building drawings, accelerating BIM modeling and rendering speed [5]. This study designs a recognition method for architectural design drawing images on the ground of object detection technology You Only Look Once version 4 (YOLOv4), and optimizes the aiming frame positioning and feature fusion module of traditional YOLOv4 technology. The optimization of architectural design methods is achieved by integrating BIM and VR technology. The research contains four. The first summarizes the current status of technologies related to architectural design methods both domestically and internationally. The second part elaborates on the improved YOLOv4 design IR model. The third part conducts performance testing on the designed IR algorithm. The fourth part summarizes the research experimental results. This study is expected to promote the digital development of the construction industry and promote the improvement of architectural design methods and schemes.

2. Literature review

To further enhance the effectiveness and efficiency of architectural design, scholars around the world have conducted a series of studies on digital architectural design. In 2023, Fernández Rodríguez leveraged the existing experience of BIM architectural design execution to optimize mechanical, electrical, and plumbing facilities as well as architectural design by constructing traditional BIM models. The study found that BIM helped to coordinate and optimize facilities, and promoted collaborative design among different entities. Promoting the BIM technology diffusion and simplifying the BIM modeling process are important research directions for BIM in the future [6]. The architectural design process provides alternative structural layout options for building development. In 2022, Sherif et al. improved the BIM model by integrating structural design, BIM modeling, and computer programming concepts. This optimization framework could effectively save 15% of structural material costs by reducing the cost of architectural design [7]. The design process of the BIM technology architectural floor tile layout plan is still deficient. In 2022, Wu et al. designed an optimization algorithm for automatic

generation of floor tile layouts based on industry experience, and improved the BIM modeling system by combining it with the parametric design platform. The results showed that the method effectively saved building design materials and improved the modeling calculation efficiency [8]. The reconstruction of historical buildings using BIM models through laser scanning, photography, and measurement is more subjective and time costly. In 2023, Croce et al. designed a high-level automated BIM reconstruction method based on artificial intelligence technology. It could import annotated data and classify and propagate geometrical shapes in the BIM modeling process. The experimental results verified the reproducibility of the method [9].

Automatic Computer Aided Design (AutoCAD) software has generated a massive amount of digital architectural design information. Traditional BIM is difficult to generate from complex building plans. In 2023, Urbana et al. generated BIM models on the ground of artificial intelligence algorithms by parsing architectural and structural drawings, and identified structural elements in the blueprint using the Mask R-CNN framework. This method was applied to the design of multi-story buildings [10]. During the architectural design process, it is necessary to refer to building drawings that are similar to the spatial planning of the design project. However, existing search methods are mostly on the ground of highly shared similarity of specific details. To search for spatial elements, in 2022, Kim et al. used BIM and Dynamo algorithms to detect drawing elements. Then a spatial relationship database was established by analyzing the similarity of spatial relationships in existing architectural drawings, achieving an automatic spatial planning model for architectural design [11].

A multi-layer fruit IR model was designed for MobileNetV2 architecture with migration learning. The results of the study showed that the recognition accuracy was as high as 99%, and the performance achieved a more significant improvement relative to the baseline model [12]. Chen et al. proposed a local feature extraction method based on the Tensorflow quantum framework and quantum Convolutional Neural Network (CNN) for binary IR classification in 2023. Experimental results showed that the designed method outperformed traditional CNN in terms of accuracy [13]. In order to solve the over-fitting problem in the deep learning, Liang et al. proposed an improved multi-stage augmented hybrid method in 2023, which was integrated into the IR classification algorithm. Experimental results showed that this method significantly improved the model classification performance without incurring a large computational overhead [14].

In summary, there have been many studies on the joint application of architectural design and computers. Various intelligent technologies to solve the challenges and deficiencies faced by BIM models has achieved better results. With the continuous development of IR technology, various fields have achieved more in-depth research results with the help of IR and classification of computer vision, but the research on the identification and classification of BIM model for building components is still in the blank. There is relatively little research on introducing IR to improve the quality and efficiency of BIM modeling. This study conducts optimization research on architectural design methods on the ground of IR and VR.

3. Architectural design methods on the ground of IR and VR

In order to optimize architectural design methods, this study conducts research on the optimization of architectural design methods based on IR and VR technology.

3.1. Construction of BIM architecture aided design model on the ground of YOLOv4 and VR

BIM technology is widely used in architectural design, improving the efficiency of design and construction. The BIM design process is shown in Fig. 1 [15, 16]. However, faced with complex architectural design schemes, a large number of architectural drawings, images, and complex building structures, the modeling process of BIM becomes complex and difficult. In order to effectively integrate different architectural design image data and maintain consistency and collaboration within the design project cycle, this study utilizes IR assisted BIM technology to analyze architectural design images.

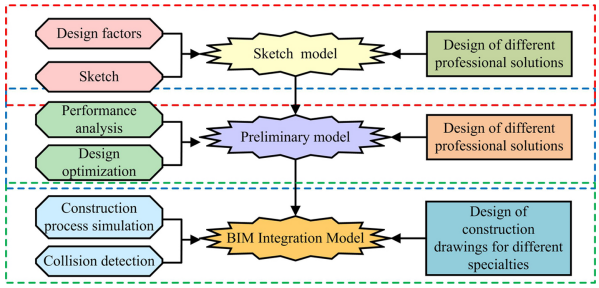


Fig. 1. BIM architecture design process diagram

The implementation method of BIM is relatively limited, and there are shortcomings in the degree and scope of visualization [17, 18]. VR technology generates digital environments through computers to achieve user environment interaction and immersive experiences [19]. Architectural designers can experience the functional space of the virtual scene in real time with the help of various image processing software, modeling software, and head-mounted immersion equipment for architectural creation and design. The architectural aided design digital tool requires a series of work, including architectural design demand analysis, BIM model design, scene construction, software technology and hardware equipment support, and multiple interaction test optimization.

Architectural design drawings are large in volume, while CAD and BIM software can only realize the drawing and editing of building components. They cannot identify component information. At the same time, there are many types of building components, and complex spatial relationships. Computer intelligence technology can greatly improve design efficiency and reduce labor costs by identifying and classifying building components in building drawings. YOLOv4 is studied for IR. Using YOLOv4 to identify the characteristics of doors, windows, stairs, equipment, and inspection components in architectural design drawings can provide

more intuitive visualization information and more accurate data support for architectural design [20–22]. YOLOv4 is divided into input layer, Backbone layer, Neck layer, and Head layer, as shown in Fig. 2.

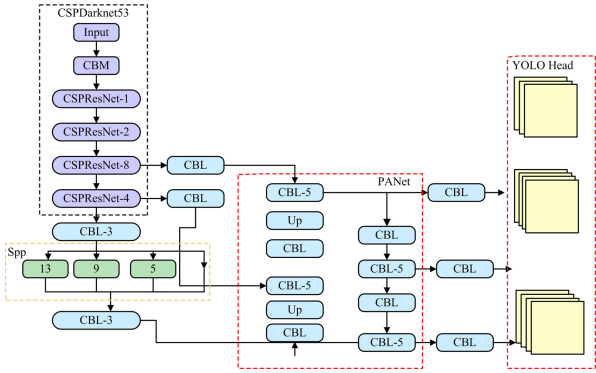


Fig. 2. YOLOv4 network structure diagram

The image feature extraction module of YOLOv4's Backbone layer uses CNN. CNN is a deep learning model including four layers. The convolutional operation obtains image feature information through several matrix calculations. The spatial size W_{output} of the output image from single channel convolution is calculated, as shown in Equation (3.1).

$$W_{\text{output}} = \frac{W_{\text{input}} - F + 2p}{s} + 1 \tag{3.1}$$

In Equation (3.1), W_{input} represents the size of the input Feature Map (FM). F represents the convolutional kernel dimension. p represents the filling unit of the FM. s represents the sliding step size of the convolutional window. By compressing the feature information of the activation function and pooling layer, the fully connected layer connects the output FM to one or more fully connected layers for final classification. The forward propagation calculation of convolutional layers, pooling layers, and fully connected layers is shown in Equation (3.2).

$$\begin{cases} a^l = \sigma\left(\sum_{k=1}^{M} a_k^{l-1} * W_k^l + b^l\right) \\ Q = \left(\frac{N}{K}, \frac{N}{K}\right) \\ a^l = \sigma\left(W^l a^{l-1} + b^l\right) \end{cases} \tag{3.2}$$

In Equation (3.2), l represent the number of layers. σ represent the activation function. $*$ represent the convolution operation. b represent the offset. W represent the neuron weight. k represent the neuron number. Q represent the spatial structure of the output matrix. N represent the input matrix of the pooling layer. K represent the spatial structure of the pooling layer.YOLOv4 divides the image into $s \times s$ equally sized grids. If the object center to be detected is located within a certain grid, the planar position detection of the object is completed.

The confidence level and intersection over union of the predicted bounding box are shown in Equation (3.3).

$$\begin{cases} \zeta = \Pr(\text{object}) * IoU_{\text{pred}}^{\text{teuth}} \\ \text{IoU}\left(\widehat{b}, b\right) = \dfrac{\left|\widehat{b} \cap b\right|}{\left|\widehat{b} \cup b\right|} \end{cases} \tag{3.3}$$

In Equation (3.3), $\Pr(\text{object})$ represents the probability of including the detected object. $\widehat{b}$ represent predicted bounding boxes, b represent actual bounding boxes. ζ represents the confidence level. IoU represents the intersection over union. Architectural design images often include multiple objects to be detected, and grids can predict the confidence level of different object categories. The final confidence calculation process is showcased in Equation (3.4). The Neck layer of YOLOv4 mainly includes Space Pyramid Pool (SPP) and Path Aggregation Network (PANet).

$$\zeta = \Pr(\text{Class}_i|\,\text{object}) * \Pr(\text{object}) * IoU_{\text{pred}}^{\text{teuth}} \tag{3.4}$$

3.2. Design of architectural drawing recognition and analysis algorithm on the ground of improved YOLOv4

Architectural drawings involve a large number of building components. There are many overlapping and blocking phenomena between building components, which leads to the fact that the traditional YOLOv4 model can't meet the needs of analyzing architectural drawings. In order to better extract feature information from architectural design images, the traditional YOLOv4 has been optimized. The improved YOLOv4 designed in the study is based on the characteristics of architectural drawings. It is improved according to the three aspects of the traditional YOLOv4: the aiming frame, the feature extraction mechanism, and the loss function. Cluster target boxes are used to distinguish overlapping building components, and a feature extraction architecture is adopted to improve the recognition accuracy of complex building components. Finally, the loss function is improved to accelerate the convergence speed of the model. The improvement of YOLOv4 is based on the YOLOv4 model functional module, which fully combines the distribution and spatial position characteristics of building drawing image components. Compared with traditional object detection algorithms, it has stronger adaptability and relevance.

The traditional YOLOv4 prior box calculation uses K-means clustering algorithm for clustering the size and shape of the target on the training set. The objective function for K-means clustering is showcased in Equation (3.5).

$$J(X, \pi) = \sum_{j=1}^{k} \sum_{i \in \pi_j} \left\| x_i - m_j \right\|^2 \tag{3.5}$$

In Equation (3.5), π_j represents the j-th class and its center is m_j, x_i represents a data point. In architectural design, door and window components often intersect or have overlapping centers with walls. Traditional clustering methods can easily lead to errors in center overlap between prior frames of different components. Therefore, it is necessary to separate the structures and perform prior frame detection. The study introduces the Kernel K-means clustering method to improve the prior box calculation of YOLOv4. The kernel function in this study is Gaussian kernel function, as shown in Equation (3.6).

$$k\left(x, x'\right) = e^{-\frac{\|x-x'\|^2}{2\sigma^2}} \tag{3.6}$$

In Equation (3.6), x, x' represent different spatial feature vectors. $\|x - x'\|$ serves as the Euclidean distance between vectors. σ serves as the hyper-parameter of the kernel function. The objective function of Kernel K-means is showcased in Equation (3.7).

$$J\left(X, \pi\right) = \sum_{j=1}^{k} \sum_{i \in \pi_j} \left\|\partial\left(x_j\right) - m_j\right\|^2 \tag{3.7}$$

In Equation (3.7), $\partial\left(x_j\right)$ represents the kernel function. The common technique for handling multi-scale is the Feature Pyramid Network (FPN). On the ground of the FPN structure, PANet combines the image feature levels from the FPN backbone model in a certain order to construct a pyramid with top-down and left-right connections. The PANet system architecture of YOLOv4 is shown in Fig. 3.

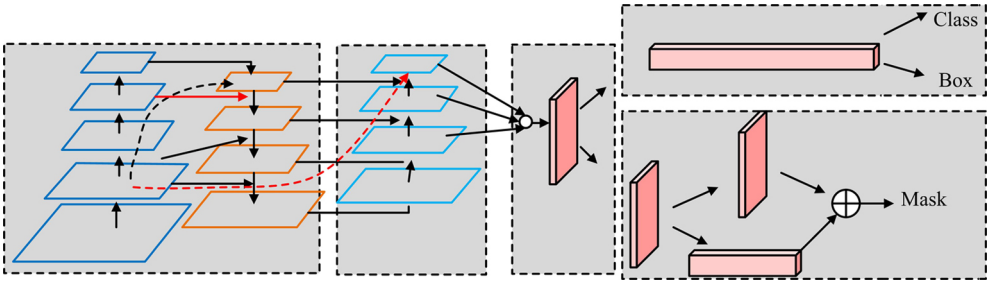


Fig. 3. PANet system architecture diagram

There is a significant similarity in the presentation of different building components in architectural design images. The feature recognition ability of PANet is relatively insufficient, and it cannot comprehensively recognize the all-round features of the target. The study introduces Neural System Architecture Search (NAS) to form NAS-FPN. The system architecture of NAS-FPN is shown in Fig. 4. NAS is a technical method for automatically searching for the optimal structure of neural networks, used to solve the incomplete defined system architectures.

However, NAS increases search space and consumes more computing resources. Therefore, this study introduces reinforcement learning to enhance the scalability of NAS-FPN. It uses

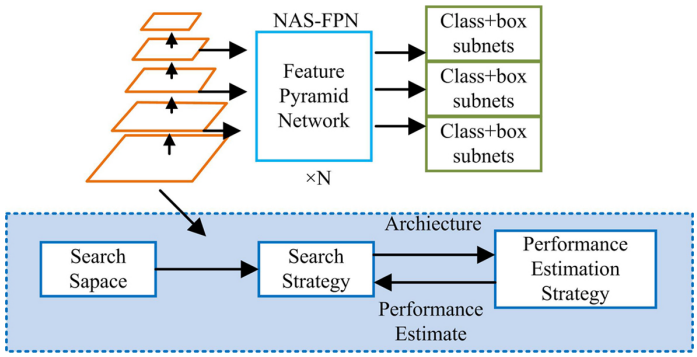


Fig. 4. NAS-FPN structure and NAS system architecture search principle

merged cells to process feature combinations in the NAS-FPN network structure. The loss function Complete IoU Loss (CIOU Loss) of YOLOv4 is derived from the YOLO series of loss functions, and the calculation process is shown in Equation (3.8).

$$L_{\mathrm{CIoU}} = 1 - \mathrm{IoU} + \frac{\rho^2\left(b, b^{gt}\right)}{c^2} + \alpha v \tag{3.8}$$

In Equation (3.8), α represent the aspect ratio penalty term. v represent the aspect size matching degree. $\rho\left(b, b^{gt}\right)$ serves as the Euclidean distance between the predicted bounding box center coordinate b and the true target bounding box center coordinate b^{gt}. c represents the minimum diagonal length of the bounding rectangle between two rectangular boxes. The α, v are shown in Equation (3.9).

$$\begin{cases} v = \dfrac{4}{\pi^2}\left(\arctan\dfrac{w^{gt}}{h^{gt}} - \arctan\dfrac{w}{h}\right)^2 \\ \alpha = \dfrac{v}{(1 - IoU) + v} \end{cases} \tag{3.9}$$

In Equation (3.9), w and h serve as the height and width of the predicted box. w^{gt} and h^{gt} serve as the height and width of the true box. However, the CIOU loss does not truly reflect the true aspect ratio difference of the bounding box. This study calculates the length and width dimensions of the target box and the prior box separately. The improved Efficient IoU loss function calculation is shown in Equation (3.10).

$$L_{\mathrm{ECIoU}} = 1 - IoU + \frac{\rho^2\left(b, b^{gt}\right)}{c^2} + \frac{\rho^2\left(w, w^{gt}\right)}{c_w^2} + \frac{\rho^2\left(h, h^{gt}\right)}{c_h^2} \tag{3.10}$$

In Equation (3.10), c_w and c_h serve as the width and height of the minimum bounding box that covers two rectangular boxes, respectively. EIoU includes three parts: IoU loss, distance loss, and aspect loss. In addition, the Efficient IoU loss function introduces a focus on high-quality prior boxes to distinguish the quality of the prior boxes. The penalty is calculated in Equation (3.11).

$$L_{\mathrm{Focal-EIoU}} = IoU^y L_{\mathrm{EIoU}} \tag{3.11}$$

In Equation (3.11), γ represents the effective parameter for suppressing outliers. The improved loss function accelerates the convergence rate of the model. Finally, the study introduces the Non-Maximum Suppression (NMS) algorithm to select multiple prediction prior boxes.

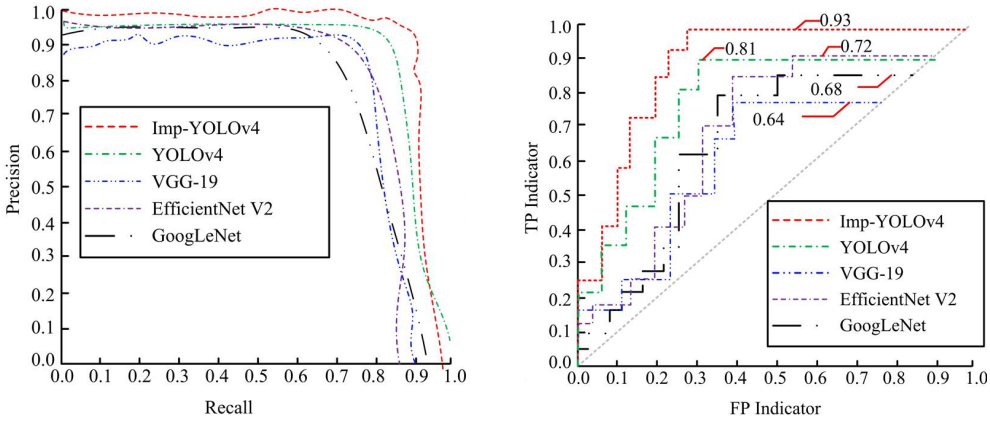
4. Application effect of joint IR and VR in architectural design methods

To test the application effect of the architectural design method in the study, performance testing and design effect analysis experiments of IR algorithms are designed. The results are analyzed and discussed.

4.1. Performance testing of improved YOLOv4 architectural drawing recognition and analysis algorithm

It uses Houses dataset, CAD image symbol automatic recognition dataset, CIFAR-10, and Visual Object Classes dataset as performance testing datasets. The selected image data for the experiment is divided into a training set and a testing set in an 8:2 ratio. The experimental environment used in the research institute is an NVIDIA GeForce GTX 1080Ti 11G graphics card, with a computer operating system of Windows 7, and an improved YOLOv4 network framework built on the ground of Python.

The performance of the designed improved-YOLOv4 model (denoted as Imp-YOLOv4) is compared with existing advanced IR models, including the traditional YOLOv4 model, deep learning model Visual Geometry Group 19 (VGG-19), image classification model EfficientNet V2, and GoogLeNet. Firstly, the Precision Recall curve (PR) of the IR model is compared with the Receiver Operating Characteristic Curve (ROC), which can comprehensively estimate the IR model. The experimental outcomes are illustrated in Fig. 5. Fig. 5(a) showcases that the

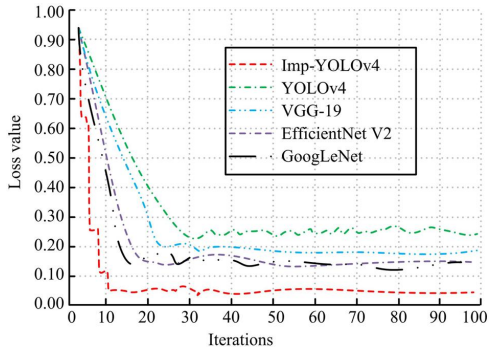


(a) PR (b) ROC

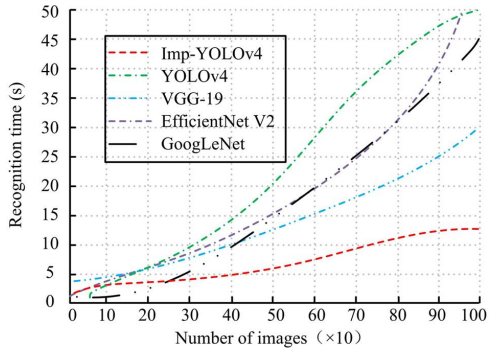
Fig. 5. Comparison of comprehensive performance for different image recognition models

Imp-YOLOv4 model designed in the study had the maximum Area Under the Curve (AUC) under the PR curve. When the accuracy was 0.9, the recall rate was at the highest level of 0.98, which was 0.3 higher than of the GoogLeNet model. Fig. 5(b) shows that the ROC of the Imp-YOLOv4 model was located at the top left corner of the coordinate axis, with an AUC of 0.93, which was 0.29 higher than that of the VGG-19 model.

The comparison results of the loss function curves and recognition efficiency of different recognition models are showcased in Fig. 6 Fig. 6(a) showcases that the loss function curve of the Imp-YOLOv4 model converged to the minimum value of 0.04, with the smallest number of iterations at convergence. After 10 iterations, the convergence curve of the model reached stability. The convergence speed of the loss curve was the fastest. Fig. 6(b) shows that the Imp-YOLOv4 model possessed the shortest recognition time and the highest recognition efficiency. When the number of IR samples reached 1000, the time consumption was within 15 seconds, which was significantly different from other models, with a maximum reduction of 70.06%.



(a) Loss function curve



(b) Recognition efficiency

Fig. 6. Comparison of damage curves and recognition efficiency of different image recognition models

The purity, Normalized Mutual Information (NMI), and Rand Index (RI) indicators are compared in different clustering methods. The statistical outcomes are illustrated in Table 1. As shown in Table 1, the Kernel K-means clustering method performed well in purity and NMI,

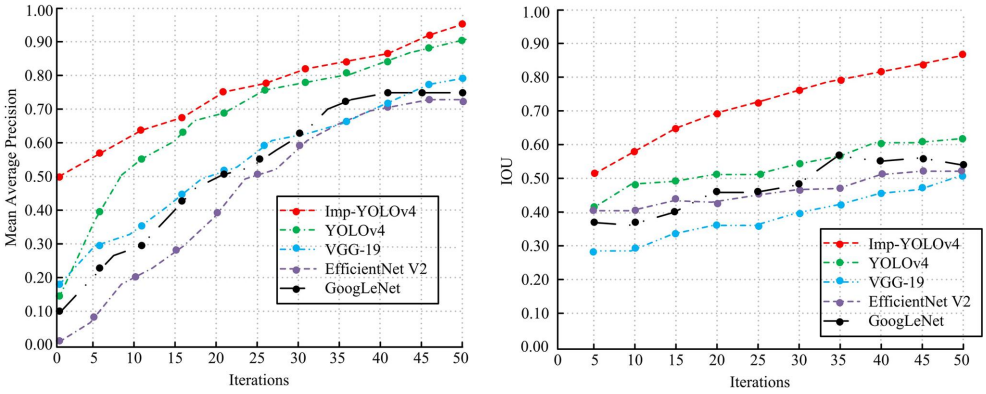
with a minimum purity value of 0.894 and a maximum value of 0.944 on different datasets. The minimum value of NMI reached 0.832, and the maximum value reached 0.931, which was better than other clustering algorithms, and the clustering effect is better. The RI value of simultaneous projection pursuit dynamic clustering is relatively large, with a maximum value of 0.942, which has higher fitting with real building components.

Table 1. Evaluation of indicators for different clustering methods

Model	**Index**	**Housesdataset**		**CAD image symbols**		**CIFAR-10**		**Visual Object Classes**	
		Test	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**	**Training**
Kernel K-means	Purity	0.926	0.944	0.909	0.899	0.921	0.918	0.894	0.901
	NMI	0.897	0.866	0.851	0.844	0.876	0.832	0.892	0.931
	RI	0.942	0.922	0.806	0.815	0.814	0.826	0.925	0.891
K-means	Purity	0.682	0.744	0.743	0.726	0.746	0.732	0.742	0.731
	NMI	0.719	0.720	0.714	0.721	0.632	0.643	0.652	0.642
	RI	0.726	0.713	0.607	0.613	0.625	0.602	0.673	0.688
DBSCAN	Purity	0.694	0.709	0.683	0.654	0.672	0.711	0.692	0.701
	NMI	0.753	0.747	0.713	0.742	0.694	0.706	0.627	0.636
	RI	0.665	0.607	0.627	0.608	0.708	0.698	0.707	0.697

4.2. Analysis of the application effect of mixed architectural design methods

It converts the DWG format files of architectural CAD design images from a certain design institute into architectural design image data of different sizes. Then, it uses data augmentation techniques to expand the dataset through changes including rotation, enlargement, or reduction, and analyzes the effectiveness of architectural design methods. The Mean Average Precision (MAP) and IOU results of different models on architectural design drawings are showcased in Fig. 7. From Fig. 7(a), the MAP value of the Imp-YOLOv4 model was at its highest level, with the highest MAP value reaching 0.96 as the number of iterations increased. This indicates that the Imp-YOLOv4 model has higher recognition accuracy than other models, and the model performs better in detecting targets. From Fig. 7(b), the Imp-YOLOv4 model had the highest IOU value curve, all above the 0.5. It can be considered that the detection performance for different building components is good, and the detection results have the highest degree of overlap with the true target position. The IOU values of the other four models are mostly below 0.5, indicating poor detection performance.



(a) MAP (b) IOU

Fig. 7. Comparison of recognition effects of architectural drawings using different methods

5. Conclusions

To solve the shortcomings of BIM model that cannot identify and classify building components, this study integrates IR and VR technology to design a new architectural design method. The experiment demonstrated that the PR and ROC of the Imp-YOLOv4 model performed better. When the accuracy was 0.9, the recall rate was 0.98, which was 0.3 higher than the GoogLeNet model. The AUC value was 0.93, which was 0.29 higher than the VGG-19 model. The loss function curve of the Imp-YOLOv4 model converged to 0.04, and the model recognition time was the shortest, with a maximum time reduction of 70.06%. The clustering strategy of this method was improved appropriately, with good performance in clustering index purity, NMI and RI. The maximum initial reading value on different datasets was 0.944, the maximum NMI value was 0.931, and the maximum RI index value was 0.942. Combining the IR with VR, the average recognition accuracy and IOU value of architectural design drawings were at the highest level, with a maximum MAP value of 0.96 and IOU value curves above 0.5. The recognition effect and visualization delay rate were improved, and the subjective application evaluation score achieved good results. Through the YOLOv4 model, BIM and VR technology are integrated to realize the architecture design process from component identification to rapid modeling and analysis of architectural drawings. However, further research is needed to identify the types of building components in this study, and the robustness of uneven and incomplete building design drawings can be further optimized. Further research on BIM high-level automatic analysis of building components should be carried out in the future.

References

[1] P.P.W. Aung, W. Choi, A.S. Kulinan, G. Cha, and S. Park, “Three-dimensional engine-based geometric model optimization algorithm for BIM visualization with augmented reality”, *Sensors*, vol. 22, no. 19, pp. 7622–7634, 2022, doi: 10.3390/s22197622.

[2] Y.Y. Al-Ashmori, I. Othman, Y. Rahmawati, Y.H. Mugahed Amran, S.H. Abo Sabah, A. D. Rafindadi, and M. Mikić, "BIM benefits and its influence on the BIM implementation in Malaysia", *Ain Shams Engineering Journal*, vol. 11, no. 4, pp. 1013–1019, 2020, doi: 10.1016/j.asej.2020.02.002.

[3] A.M. Usman and M.K.Abdullah, "An assessment of building energy consumption characteristics using analytical energy and carbon footprint assessment model", *Green and Low-Carbon Economy*, vol. 1, no. 1, pp. 28–40, 2023, doi: 10.47852/bonviewGLCE3202545.

[4] Z. Zhuang, "Optimization of building model based on 5G virtual reality technology in computer vision software", *Mathematical Biosciences and Engineering*, vol. 18. no. 6, pp. 7936–7954, 2021, doi: 10.3934/mbe.2021393.

[5] M. Masana, X. Liu, B. Twardowski, M. Menta, A.D. Bagdanov, and D.W.J. Van, "Class-incremental learning: survey and performance evaluation on image classification", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 5, pp. 5513–5533, 2022, doi: 10.1109/TPAMI.2022.3213473.

[6] J.F. Fernández Rodríguez, "Implementation of BIM virtual models in industry for the graphical coordination of engineering and architecture projects", *Buildings*, vol. 13, no. 3, pp. 743–765, 2023, doi: 10.3390/buildings13030743.

[7] M. Sherif, K. Nassar, O. Hosny, S. Safar, and I. Abotaleb, "Automated BIM-based structural design and cost optimization model for reinforced concrete buildings", *Scientific Reports*, vol. 12, no. 1, pp. 21616-21627, 2022, doi: 10.1038/s41598-022-26146-6.

[8] S. Wu, N. Zhang, Y. Xiang, D. Wu, D. Qiao, X. Luo, and W.Z. Lu, "Automated layout design approach of floor tiles: based on building information modeling (BIM) via parametric design (PD) platform", *Buildings*, vol. 12, no. 2, pp. 250–272, 2022, doi: 10.3390/buildings12020250.

[9] V. Croce, G. Caroti, A. Piemonte, L. De Luca, and P. Véron, "H-BIM and artificial intelligence: classification of architectural heritage for semi-automatic scan-to-BIM reconstruction", *Sensors*, vol. 23, no. 5, pp. 2497–2536, 2023, doi: 10.3390/s23052497.

[10] M. Urbieta, M. Urbieta, T. Laborde, G. Villarreal, and G. Rossi, "Generating BIM model from structural and architectural plans using artificial intelligence", *Journal of Building Engineering*, vol. 78, no. 17, pp. 107672–107701, 2023, doi: 10.1016/j.jobe.2023.107672.

[11] G.J. Kim, H. Gu, H. Park, and S.Y. Choo, "Development of spatial adjacency graph extraction algorithm for improving pre-design efficiency in architectural design process", *Journal of the Architectural Institute of Korea*, vol. 38, no. 1, pp. 67–74, 2022, doi: 10.5659/JAIK.2022.38.1.67.

[12] Y. Gulzar, "Fruit image classification model based on MobileNetV2 with deep transfer learning technique", *Sustainability*, vol. 15, no. 3, pp. 1906–1919, 2023, doi: 10.3390/su15031906.

[13] G. Chen, Q. Chen, S. Long, W. Zhu, Z. Yuan, and Y. Wu, "Quantum convolutional neural network for image classification", *Pattern Analysis and Applications*, vol. 26, no. 2, pp. 655–667, 2023, doi: 10.1007/s10044-022-01113-z.

[14] W. Liang, Y. Liang, and J. Jia, "MiAMix: enhancing image classification through a multi-stage augmented mixed sample data augmentation method", *Processes*, vol. 11, no. 12, pp. 3284–3302, 2023, doi: 10.3390/pr11123284.

[15] R. Yu, N. Gu, G. Lee, and A. Khan, "A systematic review of architectural design collaboration in immersive virtual environments", *Designs*, vol. 6, no. 5, pp. 93–115, 2022, doi: 10.3390/designs6050093.

[16] B. Yang, T. Fang, X. Luo, B. Liu, and M. Dong, "A bim-based approach to automated prefabricated building construction site layout planning", *KSCE Journal of Civil Engineering*, vol. 26, no. 4, pp. 1535–1552, 2022, doi: 10.1007/s12205-021-0746-x.

[17] S. S. Pibal, K. Khoss, and I. Kovacic, "Framework of an algorithm-aided BIM approach for modular residential building information models", *International Journal of Architectural Computing*, vol. 20, no. 4, pp. 777–800, 2022, doi: 10.1177/14780771221138320.

[18] L. Zarantonello and B.H. Schmitt, "Experiential AR/VR: a consumer and service framework and research agenda", *Journal of Service Management*, vol. 34, no. 1, pp. 34–55, 2023, doi: 10.1108/josm-12-2021-0479.

[19] V.S. Chan, H.N.H. Haron, M.I.B.M. Isham, and F.B. Mohamed, "VR and AR virtual welding for psychomotor skills: a systematic review", *Multimedia Tools and Applications*, vol. 81, no. 9, pp. 12459–12493, 2022, doi: 10.1007/s11042-022-12293-5.

[20] R. Gai, N. Chen, and H. Yuan, "A detection algorithm for cherry fruits based on the improved YOLO-v4 model", *Neural Computing and Applications*, vol. 35, no. 19, pp. 13895–13906, 2023, doi: 10.1007/s00521-021-06029-z.

[21] C. Dewi, R.C. Chen, X. Jiang, and H. Yu, "Deep convolutional neural network for enhancing traffic sign recognition developed on Yolo V4", *Multimedia Tools and Applications*, vol. 81, no. 26, pp. 37821–37845, 2022, doi: 10.1007/s11042-022-12962-5.

[22] M. Goncikowski, "Research by design: architectural and structural solutions allowing the integration of the skyscraper complex with the urban space in Warsaw", *Archives of Civil Engineering*, vol. 4, no. 4, pp. 21–36, 2023, doi: 10.24425/ace.2023.147645.

Received: 2024-03-26, Revised: 2024-06-11