IPPR of Traditional Wooden Building Section Based on Deep Learning

Xiaodan Liang[1], Haoming Dong[2*]

[1]*Shanghai Urban Construction Vocational College, Shanghai, China*
[2]*Shanghai Jianke Architectural Design Institute CO.Ltd., Shanghai, China*
**Corresponding author. E-mail: donghaoming76@163.com*

ABSTRACT
In recent years, deep learning (DL) and neural network models have become hot topics in new research directions in the field of machine learning and artificial intelligence. In order to protect our traditional wooden buildings, this paper applies the research of DL in IPPR to traditional wooden building profiles in China through the research of a series of IPPR algorithms such as SSD, SVM, DBN, in several different network training environments such as VGG, DenseNet, and ZF, Research on the IPPR accuracy and recognition speed of traditional wooden building sections. The results show that the SSD algorithm has the highest efficiency when the VGG network training environment and IPPR test cases are about 400, which is higher than before the improvement. The algorithm improves the accuracy by 5-10%, and the recognition speed is also increased by 2-3%.

Keywords: *Deep Learning (DL), Traditional Wooden Architecture, Building Section, Image Recognition*

1. INTRODUCTION

Ancient Chinese architecture is one of the six major components of the world's ancient architecture. The main body of ancient Chinese architecture is wooden construction. Chinese traditional wooden construction has a very high level of wooden construction, which is worthy of our study and inheritance. Speaking of the status of ancient Chinese architecture in the world, we are proud of the architectural miracle of the Chinese nation, and it is a sustenance for the revitalization of modern Chinese wooden architecture.

Protecting traditional wooden buildings is to protect traditional Chinese culture. Based on DL and IPPR (IPPR and Pattern Recognition) research, we can understand the cross-sectional structure of wooden buildings more clearly, and we can better protect and repair traditional wooden buildings. DL the foundation is data.

In medicine and health, Schlesinger W's model was tested using data from 1,000 alumni gathered from 1,000 important and underresearched key stakeholder groups. Due to its direct and indirect influence, the university's brand image was found to be a key driver of the alumni's positive WOM intentions. The study also determined the mediating role of alumni on university recognition and satisfaction [1]. However, they only conducted a lot of research in the medical field based on DL. Regarding the application of DL in the profile IPPR technology of traditional wooden buildings, they did not study and develop it. In the field of traditional wood structure construction, Fujihira M and others conducted a questionnaire survey on the protection of traditional wood structure houses and streetscapes among contractors engaged in construction and woodworking in Nara Prefecture. Contractors are experts in the construction of wooden houses, and they provide support to the residents in the maintenance work of the houses. The survey results show that in order to maintain and protect traditional wooden houses, it is necessary to train technical experts and ensure the use of appropriate materials [2]. However, they only studied the materials used in wooden structures, and did not use IPPR algorithms to conduct in-depth analysis of the sections of traditional wooden buildings. In the field of IPPR based on DL, researchers such as Das R proposed four different techniques to extract multi-view features from images. Efficiency for image classification and retrieval through fusion and data-based feature vectors. Their proposed method outperforms the state-of-the-art in content-based image recognition. And the average accuracy of classification and retrieval increased by 17.71% and 22.78% [3]. However, the improvement and optimization technology of IPPR they studied has not been applied to the traditional wooden building section.

The innovation of this article is: collected and annotated a dataset of 2543 cross-section images of traditional wooden buildings in 3 categories. Based on the task requirements and data sets of traditional timber building profile classification and recognition, improvements have been made from two aspects: network model and algorithm training. In terms of network model, an improved model of DenseNet, UDenseNet, and a hybrid pooling method are designed.

2. EXPLANATION OF METHODS AND RELATED CONCEPTS

2.1. DL

DL [4] is essentially a complex and deep neural network algorithm model built by the human brain.

Convolutional neural network is a network model of DL, which originated from multi-layer perceptrons. It is a feed-forward neural network [5, 6]. The neurons in the network receive the input of the previous layer of neurons, and the input information is nonlinearly transformed [7, 8].

2.1.1. Convolution

Where B represents the input of the nth neuron in the mth layer, and w represents the nth neuron in the mth layer and the m-1th layer. The connection weight of i neurons, b represents the bias term, and the function g(·) is the activation function.

$$b_m^n = g(\sum_i \omega_{mi}^n b_i^{n-1} + c_m^n) \quad (1)$$

Where B represents the nth feature map of the mth layer, the symbol "*" represents the convolution operation, and D represents the nth feature map of the mth layer and the m-1th feature map [9, 10].

$$B_m^n = g(\sum_i D_{mi}^n * B_i^{n-1} + c_m^n) \quad (2)$$

2.1.2. Pooling

Pooling is also called sub-sampling or down-sampling, and its function is to aggregate the features of different pixel positions in adjacent areas of the image [11]. Suppose the input feature mapping of the pooling layer is shown in formula (3):

$$X \in R^{W \times H \times D} \quad (3)$$

Pooling refers to down-sampling each area to obtain a value as a generalization of this area [12, 13]. There are two commonly used pooling operations:

Maximum pooling: Generally, the maximum value of all neurons in a region is taken.

$$G_{m,n}^{[d]} = \max_{i \in R_{m,n}^{[d]}}(x_i) \quad (4)$$

Average pooling: Generally, the average value of all neurons in a region is taken.

$$G_{m,n}^{[d]} = \frac{1}{W/M * H/N} \sum_{i \in R_{m,n}^{[d]}} x_i \quad (5)$$

2.2. Traditional Wooden Buildings

There are three types of wood structure buildings in China, namely hanging beam type, barrel type, and through beam type. Some important wood structure buildings basically adopt the hanging beam type. For example, some folk wood structure buildings use barrel type, and some special areas use pier beams. Formula [14, 15]. The specific cross-sectional diagrams of the three types of wooden buildings are shown in Figure 1:

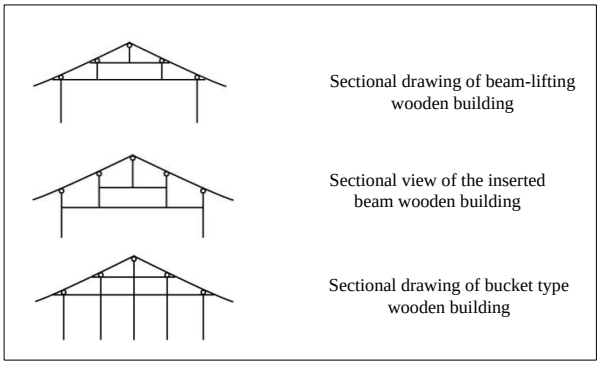

Figure 1. Specific cross-sectional views of three types of wooden buildings

2.3. IPPR and Recognition Technology

IPPR [16, 17] refers to the need to use a computer to process recognizable images and sequence the recognition process. The specific calculation type is shown in formula (6).

$$z(m,n) = \{[y(m+1,n+1) - y(m,n)]^2 + [y(m+1,n) - y(m,n+1)]^2\}^{\frac{1}{2}} \quad (6)$$

(m, n) is a point in the image, y(m, n) is the input image, and z(m, n) is the output image. It can be expressed by formula (7):

$$z(m,n) \approx |y(m+1,n+1) - y(m,n)| + |y(m+1,n) - y(m,n+1)| \quad (7)$$

Since the Roberts differential detector uses the same number of patterns, the gradient width value at (x, y) is actually the intersection value shown in Figure 2, offset by half of the actual position [18, 19].

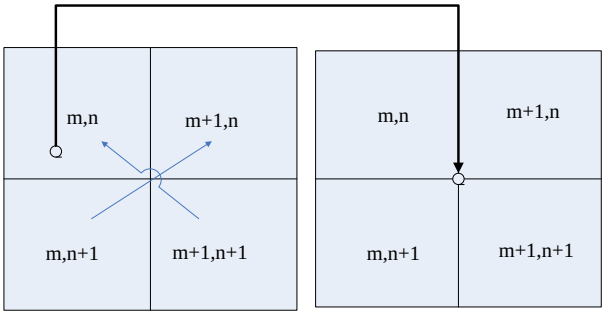

Figure 2. Schematic diagram of roberts operator algorithm

2.3.1. Region segmentation

The principle of regional development is to group segments with the same or similar functions into one site [20, 21]. First, find the original as the seed in each segmentation area, then merge the relative and seed with the same or similar characteristics in turn, then use the newly merged substitution as the seed, and repeat the above process until a result with similar characteristics is found [22, 23].

(1) Global features

1) The formula for calculating the proportion of pixels with a characteristic value of ni in the total pixels of the image is shown in formula (8):

$$g(n_i) = \frac{x(n_i)}{N} = \frac{x(n_i)}{\sum_j x(n_j)} \tag{8}$$

According to formula (8), the feature histogram with feature n can be expressed as shown in formula (9):

$$F(p) = [f(n_1), f(n_2), ... f(n_j)] \tag{9}$$

According to formula (9), we can set an unknown number equation formula (10):

$$\partial(n_i) = \sum_{j=1}^{i} f(n_i) \tag{10}$$

It can be deduced from the above formulas that the cumulative histogram of the feature n in the image P can be expressed as formula (11)

$$\partial(P) = [\partial(n_1), \partial(n_2) ... \partial(n_j)] \tag{11}$$

Assuming that the size of the image is m×n, the spatial autocorrelation function of the observation value g(a,b) at the spatial position y(m,n) can be defined as:

$$s(p,1) = \frac{\sum_{a=0}^{m-1}\sum_{b=0}^{n-1} g(a,b)g(a+p,b+q)}{\sum_{a=0}^{m-1}\sum_{b=0}^{n-1}[g(a,b)]^2} (a = 0,1...m-1; b = 0,1...n-1) \tag{12}$$

(2) Shape feature

First, select an arbitrary point (a, b) on the boundary of the target to be detected as the reference point, and mark the offset vector (a) from the arbitrary point (x, y) on the edge of the shape boundary of the reference point. To do. , B), the angle between r, r and the x axis is Φ, the inclination angle between the (x, y) tangent and the positive direction of the x axis is θ, the coordinates of the figure in the image can be obtained in the following way.

$$x = a + s * r(\alpha) * \cos[\phi(\alpha) + \beta] \tag{13}$$

$$y = b + s * r(\alpha) * \cos[\phi(\alpha) + \beta] \tag{14}$$

3. IPPR EXPERIMENT OF TRADITIONAL WOODEN BUILDING SECTION

3.1. Experimental Model and Data Set

The data sets used in this experiment are all searched on the Internet, and they all come from a total of 2543 jpg format pictures of different types of traditional wooden building sections. The picture size is about 500×400, and the maximum is not more than 500×600. Without losing the image information, in order to reduce the model parameters, it is compressed to 224×224 pixels. The size and comparative analysis of the three traditional wooden building sections were shown in Table 1:

Table 1. Comparative analysis of three traditional wooden building sections

Structure type	Beam lift	Insert beam	Wear bucket
Connection relationship	Rafter-purlin-beam-child column-beam-column, the connection of the components is large, and the connections between the components are close.	Rafter-purlin-child column-beam-child column-beam-column, the association of components Similar to lifting beam.	Rafter-purlin-column, the connection of the components is small, and the independence of each component itself is strong.
Bucket arch	The stigma uses a bucket arch.	The bucket arch is rarely used except for the inserts and diagonal braces.	
Lumber	Due to the large force, the requirements for materials are high, and the amount of materials used is also large.	The amount of material used is between the lifting beam and the bucket	The amount of material used is small.

3.2. Network Training Method and Parameter Configuration

Select 2412 traditional wooden building profile images in the data set as the training set. To use DL algorithms for network training, we need to uniformly convert the pictures into a format of 32 pixels wide by 32 pixels high. So we adopt the method of enhancing the image data set and brightness [24, 25], and then randomly crop it to a size of 32×32. The enhancement result is shown in Figure 3:

Figure 3. Increased brightness of eave tiles in wooden buildings

In order to observe whether fitting occurs during the network training process, we need to extract 5% of the images from the training set for verification. The specific network training and verification process is shown in Figure 4:

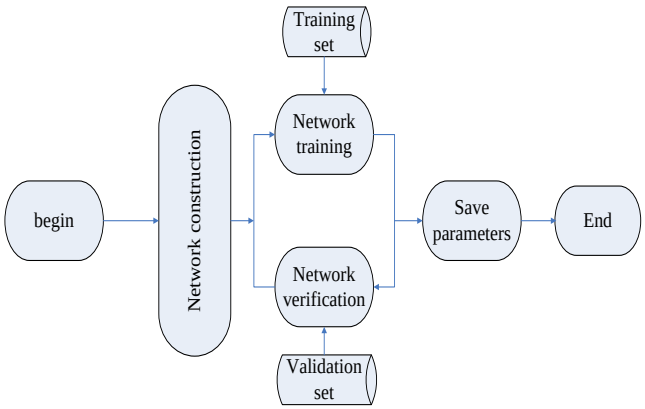

Figure 4. Network training and verification process

Figure 4 shows the process of network training and verification. This stage is mainly to complete the learning and tuning of network parameters. After the network training is completed, we need to save the weights and offsets of the model as a file. In the later stage of traditional wooden building profile recognition, we only need to construct a network parameter questionnaire. The specific parameter settings are shown in Table 2 and just import this parameter file.

Table 2. DL network parameter configuration

Name	Input dimension	Output image	Output dimension	Output image
Convolutional layer 1	2	156×156	2	156×156
Pooling layer 1	32	64×64	32	64×64
Convolutional layer 2	192	32×32	192	32×32
Pooling layer 2	480	8×8	480	8×8
Loss function / classification	832	2×2	8832	2×2
Classification layer	1200	1×1	1000	1×1

4. EXPERIMENTAL RESULTS AND ANALYSIS OF TARGET RECOGNITION FOR WOODEN BUILDING PROFILE

4.1. Recognition Rate of Timber Building Profiles under Different Networks

First, two different networks were trained with a data set of 2500 wooden building profile images to obtain a training model, and then 2100 test images were subjected to classification and recognition tests to obtain the correct identification numbers of the three wooden building profile categories. And the recognition accuracy rate is shown in Table 3:

Table 3. Recognition rate of wooden building profile images under two different networks

Model	Classification	Beam lift	Insert beam	Wear bucket
SSD+DenseNet	Test sample	700	700	700
	Identify the correct number	680	640	650
	Recognition accuracy rate	97%	91%	93%
SSD+VGG	Identify the correct number	690	630	640
	Recognition accuracy rate	98%	90%	91%

It can be seen from Table 3 that under the two networks of VGG and DenseNet, the accuracy of IPPR of wooden buildings based on the SSD algorithm is not much different. Therefore, for IPPR technology,

whether the choice of DL algorithm is correct and reasonable is extremely important.

4.2. Impact of the Number of Test Samples and the Type of Algorithm

This experiment adopts the VGG network environment, takes two kinds of training sample numbers and test sample numbers, chooses different hidden layer nodes, and compares the influence of the number of training samples and algorithm types on the accuracy of image recognition.

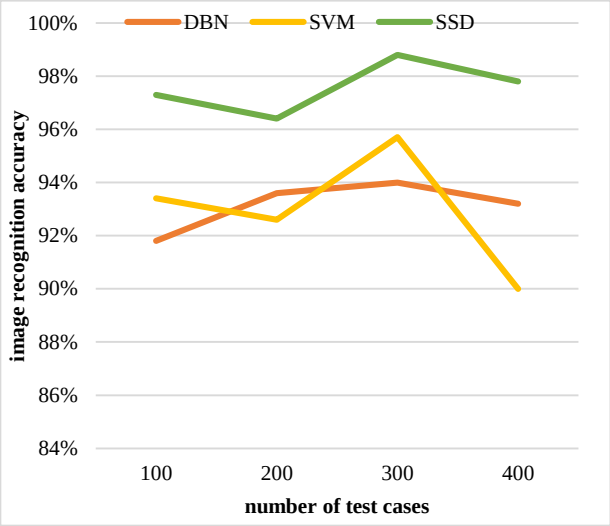

Figure 5. The impact of different test sample numbers and algorithm types on the accuracy of image recognition

It can be seen from the figure 5 that under the same network environment and the same test algorithm, the accuracy of IPPR of wooden building sections fluctuates between 90-99%. It can be seen that the SVM algorithm Suitable for experiments with relatively few test cases.

4.3. Target Recognition Test Based on Wooden Building Profile

After the experiment, the DL target recognition time in different network environments was tested. The total number of images is 400, of which 150 are in the VOC; the data provided in the network search is 250, and the image size is 32×32. The test result is shown in Figure 6.

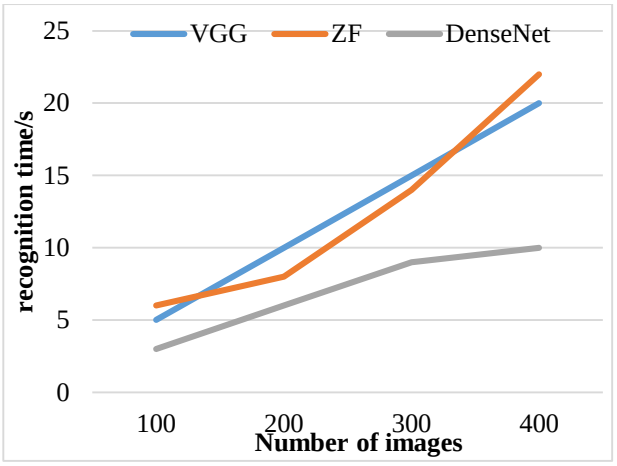

Figure 6. Target recognition test results

It can be seen from Figure 6 that the recognition time of a single image is about 0.13s. When 400 images are tested, the total time is about 23s. The average recognition time of pictures is only about 0.05s, which fully meets the needs of practical applications. Therefore, we can use VGG network training as a network environment for future design experiments.

4.4. Verification of Network Structure Improvement

The DenseNet-220 network contains 3 pooling layers. We first use the DenseNet network as the experimental object to test the impact of different pooling layer configurations on the recognition accuracy of the DenseNet network. The experimental data is given in Table 4.

Table 4. DenseNet pooling experiment comparison table

Verification accuracy	Pooling layer 1	Pooling layer 2	Pooling layer 3
91.47%	max	avg	G-avg
93.5%	avg	mix	G-mix
94.4%	mix	max	G-max

The experimental results in the above table show that using hybrid pooling instead of the mean pooling operation in the middle layer of the network, DenseNet-220 obtains a verification accuracy of 93.26%, an increase of about 0.9%. Then, the effect of hybrid pooling is further verified on DenseNet, and the verification result is shown in Figure 7:

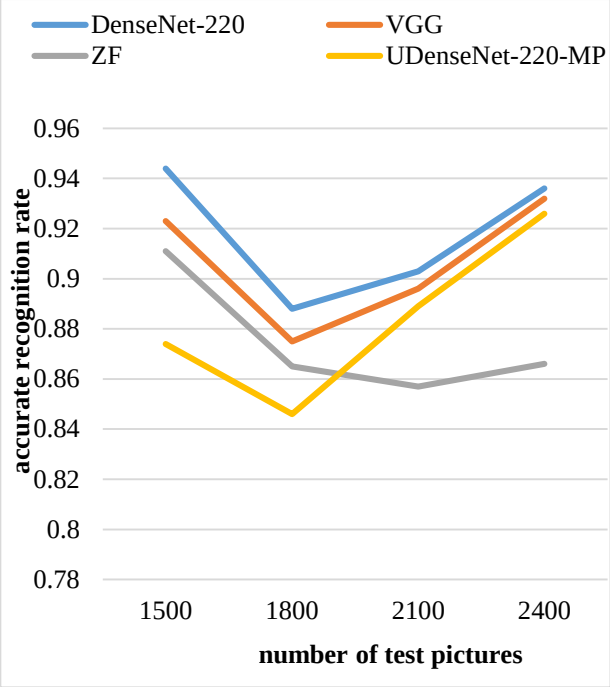

Figure 7. Method verification of mixed pooling effect

It can be seen from the verification results in Fig. 7 that the recognition effect of UDenseNet-220 when the weight parameters are not shared between channels is better than the recognition effect when the weight parameters are shared between channels. DenseNet-220, which uses a hybrid pooling operation that does not share weight parameters between channels, finally achieved a verification accuracy of 94.4%.

5. CONCLUSIONS

This article first introduces the research background and significance of traditional wooden building profile IPPR based on DL, and the current research status at home and abroad. At the same time, it provides an overview of traditional wooden building profile image feature extraction algorithms and compares traditional machine learning image feature extraction techniques. As well as the structural features of several popular convolutional neural networks, through in-depth research on deep convolutional neural networks, it is found that convolutional neural networks play a vital role in feature extraction. The convolutional neural network can effectively make up for this shortcoming. The capacity of the sample set of traditional wooden building profile images in this article is still small, and more application-worthy models need to use more categories and more sample data sets for training and verification.

ACKNOWLEDGMENT

The Research and innovation projects of SRIBS (KY10000440.20210001).

REFERENCES

[1] Schlesinger W., Cervera-Taulet A., Wymer W. The influence of university brand image, satisfaction, and university identification on alumni WOM intentions. Journal of Marketing for Higher Education, 2021(1):1-19.

[2] Fujihira M. A Study on Cooperation of Building Contractors in the Maintenance of Traditional Wooden Houses: Based on a questionnaire survey among building contractors. Journal of Architecture and Planning (Transactions of AIJ), 2017, 82(731):191-199.

[3] Das R., Thepade S., Ghosh S. Framework for Content-Based Image Identification with Standardized Multiview Features. Etri Journal, 2016, 38(1):174–184.

[4] Kermany D S., Goldbaum M., Ca I W., et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based DL. Cell, 2018, 172(5):1122-1131.

[5] Chen Y., Lin Z., Xing Z., et al. DL-based Classification of Hyperspectral Data. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 2017, 7(6):2094-2107.

[6] Wang X., Gao L., Mao S., et al. CSI-based Fingerprinting for Indoor Localization: A DL Approach. IEEE Transactions on Vehicular Technology, 2016, 66(1):763-776.

[7] Ravi D., Wong C., Deligianni F., et al. DL for Health Informatics. IEEE Journal of Biomedical & Health Informatics, 2017, 21(1):4-21.

[8] Hou W., Gao X., D Tao, et al. Blind Image Quality Assessment via DL. IEEE Transactions on Neural Networks and Learning Systems, 2017, 26(6):1275-1286.

[9] Schirrmeister R T., Gemein L., Eggensperger K., et al. DL with convolutional neural networks for decoding and visualization of EEG pathology. Human Brain Mapping, 2017, 38(11):5391-5420.

[10] Albarqouni S., Baur C., Achilles F., et al. AggNet: DL From Crowds for Mitosis Detection in Breast Cancer Histology Images. IEEE Transactions on Medical Imaging, 2016, 35(5):1313-1321.

[11] Tom Y., Devamanyu H., Soujanya P., et al. Recent Trends in DL Based Natural Language Processing [Review Article]. IEEE Computational Intelligence Magazine, 2018, 13(3):55-75.

[12] Zhao W., Du S. Spectral–Spatial Feature Extraction for Hyperspectral Image Classification: A Dimension Reduction and DL Approach. IEEE

Transactions on Geoscience and Remote Sensing, 2016, 54(8):4544-4554.

[13] Zhu X X., Tuia D., Mou L., et al. DL in Remote Sensing: A Comprehensive Review and List of Resources. IEEE Geoscience & Remote Sensing Magazine, 2018, 5(4):8-36.

[14] Takaiwa Y., Matsuno K. An Experimental Study on Shear Stiffness Evaluation under Large Lateral Deformation of Traditional Wooden Building with Mud Wall. Journal of Structural & Construction Engineering, 2017, 82(738):1245-1253.

[15] Lee, Hyunsang. Practice for the Traditional Wooden House Spreading in Kyoto. Review of Architecture and Building Science, 2016, 60(8):28-33.

[16] Rossel G. Wooden Boatbuilding The Sydney Wooden Boat School Manuals. WoodenBoat, 2019(270):96-97.

[17] Takabayashi H., Motoike R., Kado K., et al. Study on tool path generation for wood processing using articulated robot. Aij Journal of TECHNOLOGY and Design, 2016, 22(51):813-816.

[18] Xue, LI, Nobu, et al. The Structure and Timber Connection of Chuandou-system Wooden Houses of Miao Nationality\n-A study on the construction method of chuandou-system wooden houses of minority nationality in Guizhou Province, China-. Journal of Architecture and Planning (Transactions of AIJ), 2017, 82(732):383-391.

[19] Wikantari R R., Narumi K. Architectural Characteristics and the Altering Tendency of Traditional Housing Group of Kudus, Indonesia. Journal of Architecture Planning & Environmental Engineering, 2017, 64(518):213-221.

[20] Seo, Jung-Seung, Kim, et al. Improvement method of process management on new han-ok style public building project. The International Journal of the Korea Institute of Ecological Architecture and Environment, 2017, 17(3):133-140.

[21] Albukhanajer W A., Jin Y., Briffa J A. Classifier ensembles for image identification using multi-objective Pareto features. Neurocomputing, 2017, 238(MAY17):316-327.

[22] Chui T K., Tan J., Li Y., et al. Validating an automated image identification process of a passive image-assisted dietary assessment method: proof of concept. Public Health Nutrition, 2020, 23(15):2700-2710.

[23] Han T. Design and application of multicolor image identification in soil pollution component detection. Arabian Journal of Geosciences, 2020, 13(18):1-9.

[24] Yuan M., Yin D., Ding J., et al. A Multi-image Joint Re-ranking framework with updateable Image Pool for Person Re-identification. Journal of Visual Communication and Image Representation, 2019, 59(FEB.):527-536.

[25] Bingchao, Xu, Xiaofeng, et al. Source camera identification from image texture features - ScienceDirect. Neurocomputing, 2016, 207 (Sep.26): 131-140.