International Journal for Housing Science and Its Applications
Publish August 4, 2025. Volume 46, Issue 3 Pages 1872-1881

https://doi.org/10.70517/ijhsa463148

Gene Recognition and Genealogy Construction of Ancient Village Architecture in Henan Province Based on Artificial Intelligence Algorithm under the Framework of Digital Intelligent Countryside

Yanjun Wang[1,*]
[1] School of Art and Design, Pingdingshan University, Pingdingshan, Henan, 467000, China
Corresponding authors: (e-mail: wangjianli0391@163.com).

Abstract In order to realize the digital inheritance and development of ancient village buildings, this paper considers the application of artificial intelligence algorithms to ancient village buildings. Taking the ancient village buildings in Henan Province as an example, deep learning methods such as U²-Net and YOLO, convolutional neural network, etc. are used to identify the overall recognition and color recognition of the ancient village buildings in the place, so as to realize the extraction of the genes of the ancient village buildings in Henan Province. On the basis of the results of building recognition, the factors such as façade styles, roof shapes, decorative patterns, color matching, and building materials of the ancient village buildings are analyzed. Through quantitative analysis of the statistical results, the architectural genealogy of ancient villages in Henan Province was constructed. The genes of façade style, roof shape, decorative pattern, color matching, and building materials of ancient villages in Henan Province are closely related to each typical building, and the characteristics of ancient villages with different buildings are different. The architectural layout of the ancient villages in Henan Province is dominated by 6 rooms with broad faces. 6-room a zigzag with three bright, two dark and two dumping sleeves plan layout is the architectural prototype, and the number of two-story attic variants is 1.5 times of the prototype.

Index Terms artificial intelligence algorithm, U²-Net, deep learning, architectural gene, genealogy

I. Introduction

As an important part of cultural heritage in rural areas, ancient village architecture assumes the role of cultural inheritance, and Henan's unique geographic location, climate, and building material resources make Henan's traditional architecture present a magnificent image and rich regional types [1]-[4]. It can be said that Henan is the best physical repository to study the generation and development of Chinese traditional architecture [5]. Under the perspective of Digital Intelligent Countryside, drawing on advanced conservation and development experiences at home and abroad, and giving full play to the regional characteristics of Henan Province, the gene identification and genealogy construction of ancient villages' architecture is an important way to solve the various problems existing in the conservation of ancient villages' architecture in Henan Province in the new era, and to realize the win-win situation of rural revitalization and the conservation of ancient villages' architecture [6]-[9].

Architectural gene recognition, derived from the concept of landscape genes inspired by the principles of biogenetics, Western scholars creatively put forward the concept of cultural genes, which are used as the basic unit for the propagation of cultural representations [10]-[12]. Through the comparative analysis with biological genes, the basic unit of cultural transmission and its mutation law are explored, opening up a new direction of cultural research [13], [14]. And the construction of landscape gene mapping as a systematic process, derived from geomatics information mapping, visualizes the regular characteristics of landscape genes in graphic language, and systematically expresses the cultural connotation it contains [15]-[17]. Ancient village architecture in Henan Province is characterized by many distinctive regional characteristics and a long history. By analyzing the landscape genes of ancient village architecture in Henan Province, deeply excavating and extracting the local characteristics of its architecture, and constructing a prototype gene map of its village architecture, it can help the sustainable development of the village architectural landscape [18]-[21].

This paper studies architectural genes and their architectural genealogies in depth, and tries to extract architectural genes and construct genealogies from the aspect of artificial intelligence. Taking the ancient village buildings in Henan Province as the research object, the U²-Net model is introduced and improved by adding the ECA attention mechanism on the basis of the original model. The improved U²-Net model is used in the recognition of ancient village buildings. Secondly, the colors of ancient village buildings are carefully extracted and analyzed by deep learning algorithms such as YOLO, convolutional neural network and K-mean clustering algorithm. Twenty-four typical ancient villages were selected, and five factors, namely, façade style, roof shape, decorative patterns, color matching, and building materials, were extracted and analyzed by regression. Using the extraction results and the quantitative results of the architectural layout of ancient villages in Henan Province, the architectural genealogy is constructed.

II. Ancient Village Architecture Gene Recognition

II. A. Architectural genetics

The architectural forms of different regions are jointly influenced by multiple factors such as nature, economy and humanities. In response to these complex factors, architecture, like genes in a living organism carrying genetic information, steadily displays its unique architectural genetic qualities. These influencing factors are intertwined with each other and work together in the generation of architectural form, and through the adaptive adjustment of the environment, the building is prompted to grow and evolve continuously. As a result, the generation and growth law of architecture and the regulation mechanism of biological genes present the phenomenon of "heterogeneous isomorphism". In the context of the pursuit of sustainable development, we regard architecture as an organic living organism in the natural world, and integrate the biogenetic theory into architectural research. The inherent law that architecture generates and grows can be regarded as the self-regulation mechanism of architecture. People's response to various environmental factors is called "architectural genes" [22]. This initiative aims to explore the core mechanism of architectural generation and growth, so as to provide a scientific method and basis for the adaptive renewal and sustainable development of buildings.

Genes carry genetic information, which is transmitted to the offspring through a mechanism that ensures that the offspring present traits similar to those of their parents. In fact, a gene-like phenomenon often occurs in the process of cultural inheritance and transmission. By virtue of its inherent position and nature, a certain culture is able to maintain its uniqueness through transmission and dissemination. At the same time, certain mutations will occur in the process to adapt to environmental changes in order to obtain a better form of inheritance and dissemination. Taking traditional residential buildings as a typical representative, traditional residential buildings in a certain region show such similarity because they, as carriers of culture, maintain the characteristics of their cultural genes through inheritance and dissemination. However, due to the changes in time and space, the traditional residential architecture genes will also make necessary mutations to adapt to the environment while maintaining their cultural genes. This is both the basic law of biological inheritance and the internal logic of regional architectural evolution. Although biology and architecture have different attributes, they have similar inheritance principles, and the life-like architecture gene analysis diagram is shown in Figure 1. In this paper, we take the gene as the perspective to conduct in-depth research and excavation of ancient village architecture in Henan Province.

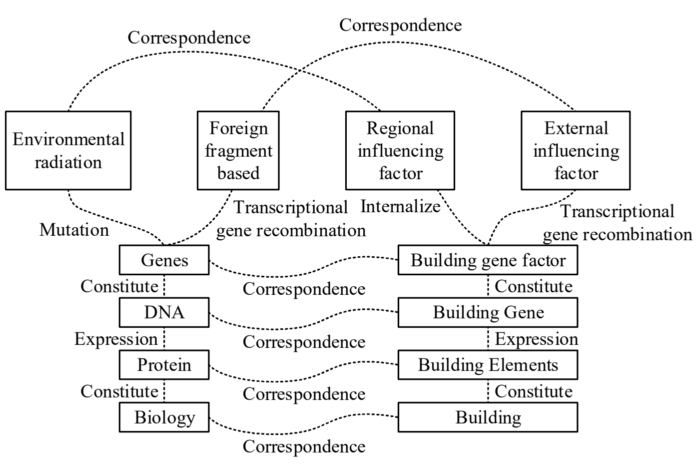

Figure 1: Genetic analysis of like-life buildings

Due to the different expression of architectural genes in ancient villages in Henan Province and the different patterns of expression, we call it architectural gene mapping. This atlas not only reflects the inner law of the architectural genes of the ancient villages in Henan Province, but also accurately reflects the logic and order of their genes. An atlas is an atlas that aims to understand things more deeply and organize them by categories through graphical expression. It is based on the basic principles of biogenetic mapping, and systematically expresses the external forms, internal features and their patterns of architectural genes by means of scientific illustrations. The atlas summarizes, summarizes and abstracts the architectural features in a schematic way so that we can understand the architectural features more intuitively. It plays an important role in the research of digitization of ancient villages in Henan Province. The concept of “architectural gene mapping” is inspired by “landscape gene mapping”. The architectural gene mapping of ancient villages in Henan Province can include various forms of expression, such as architectural form and structure mapping, architectural component mapping, functional layout mapping and detail decoration mapping.

II. B.Recognition of ancient village buildings based on artificial intelligence algorithm

II. B. 1) Building Recognition Based on U²-Net Modeling

U²-Net has significant advantages in processing semantic information of complex images [23], U²-Net has strong feature extraction ability, but it does not pay enough attention to edge and detail information when processing complex images. In the feature fusion stage, there is a lack of effective inter-channel interaction mechanism, which affects the model's ability to extract important features.

To further enhance the model performance, U²-Net introduces the ECA attention mechanism in the final feature fusion stage. This mechanism effectively enhances the network's attention to the edge semantic information, and the improved U²-Net network structure is shown in Figure 2.

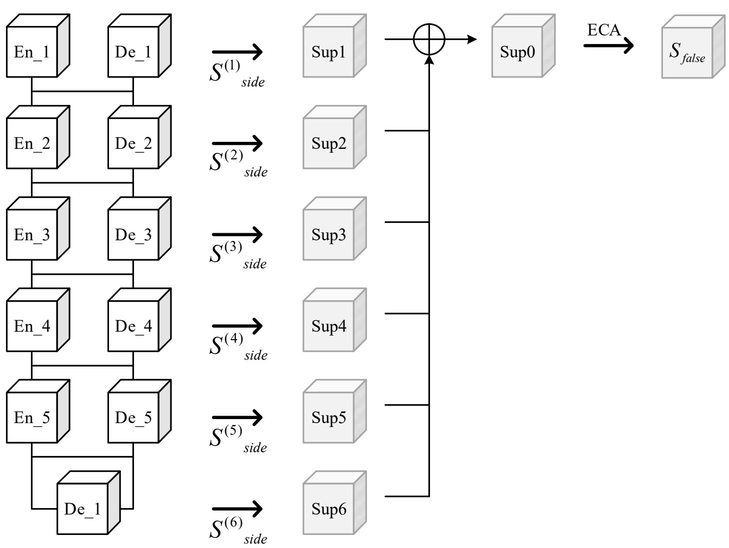

Figure 2: Improve the U²-Net network architecture diagram

The ECA attention mechanism can adaptively adjust the channel weights so as to highlight the important features and help the model better recognize the key information in the complex background. By weighting the channels, it can effectively reduce the influence of background noise and improve the accuracy of target extraction. The ECA attention mechanism enables the model to dynamically select the importance of different feature channels and adapt to the diversity of complex scenes. Enhancing the expression ability of local features makes the extraction of detailed parts in the complex background more accurate.

When the input features $X \in R^{H\times W\times C}$ are compressed once through an averaging pooling layer, $X_{avg} \in R^{1\times 1\times C}$ is obtained. The expression within each channel is shown in equation (1):

$$F^{'} = AvgPool(X) \tag{1}$$

A fast one-dimensional convolution of size k is used to realize the information interaction between channels. In order to determine the size k value of the convolution kernel, an adaptive function is introduced, which can be dynamically adjusted as needed to specify the coverage of the locally interacting information channels. This design allows cross-channel interactions to be performed more frequently and efficiently in layers with a larger number of channels, thus improving the performance of the model. The k value adaptive function is shown in equation (2):

$$k = \psi(C) = \left| \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right|_{odd} \tag{2}$$

$|\cdot|_{odd}$ denotes its closest odd number, and γ and b represent two adjustable hyperparameters that control the scaling factor of the exponential term. Through the mapping ψ, interactions between high-dimensional channels are realized in a wider domain, while the low-dimensional channels localize their interactions through a nonlinear mapping. This design facilitates the exchange of information between channels of different dimensions. The weights M_c of the feature maps are obtained through Sigmoid function activation, as shown in equation (3):

$$M_C(F^{'}) = \sigma(CID_k(F^{'})) \tag{3}$$

C is the number of channels, CID_k denotes a one-dimensional convolution of size k, and σ denotes the Sigmoid activation function.

The weight coefficients $M_C(F^{'})$ are multiplied element-by-element with the input features X to obtain the feature map of the channel's attention $\tilde{X}$, and the expression is shown in equation (4):

$$\tilde{X} = M_C(F^{'}) \otimes X \tag{4}$$

The ECA attention mechanism is introduced to adaptively adjust the weights of each channel so that it can better reflect the task-related features. Since the method only involves the channel dimension and does not need to operate on the spatial dimension, the mechanism not only improves the expressive ability of the model, but also enhances the generalization ability of the model, thus improving the computational efficiency.

In order to improve the training speed of the U²-Net model, this chapter introduces random deactivation in the channel convolution layer to speed up the convergence of the model. The neuron random deactivation is shown in Figure 3. The risk of overfitting is effectively reduced by randomly discarding some neurons during training. The generalization ability of the model is improved, the robustness of the model is enhanced, and the model performs more stably in the face of noise and changes.

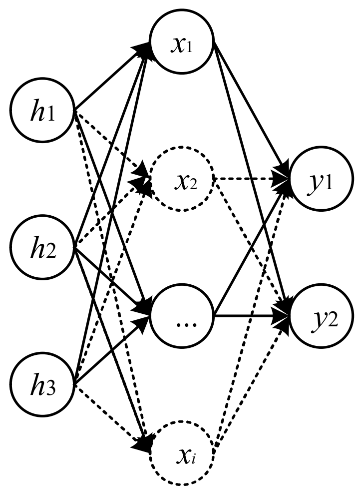

Figure 3: Schematic diagram of random inactivation of neurons

II. B. 2) Deep Learning Based Building Color Recognition

Based on the building recognition framework, buildings are categorized into building regions and panoramic regions. The original image is segmented and stored according to the building recognition framework. Thus, the images are stored as building image and panoramic image. After determining the building color and panorama color, the system extracts the color from both images using K-value setting and assigns the corresponding color label to each image.

The main goal of YOLO is feature extraction [24]. YOLO-V5 builds on YOLO-V4 to gradually improve performance. Given the requirements of the article on building recognition, YOLO-V5 is chosen to deploy deep learning and recognize building images.

(1) Convolutional neural network algorithm

1) Single-layer neural network

The deep learning based building color recognition algorithm is implemented through a single layer network structure. During the training process, the algorithm captures many feature values (e.g. $x_1 x_2 ...... x_n$) and applies the function $a = g(z)$ to evaluate the features of various objects. This allows for the detection of buildings and helps to distinguish between different building classes. Eq. (5) illustrates this process, where $z = WT + b$ denotes the linear regression prediction, which is a prerequisite required for the introduction of the Sigmoid function:

$$z = W^{\wedge}Tx + b(x = x_1 x_2 \cdots\cdots x_n) \tag{5}$$

2) Activation

The article introduces nonlinear functions in building color recognition algorithms to learn complex functions in depth. In the article, Sigmoid and tanh activation functions are used, both of which exhibit S-shaped saturation features. The inclusion of the ReLU slope function further improves the effectiveness of neuron screening.The Sigmoid function, denoted as $f(x) = 1/(1+e^{-x})$, produces constant positive output values. As a result, the weight updates are restricted to the same direction, affecting the convergence speed; on the contrary, the zero-centered tanh function $e^x - e^{-x} / e^x + e^{-x}$ contributes to a fast convergence without a significant impact on the loss values; moreover, the ReLU function $f(x) = \max/(0, x)$ allows a subset of neurons to produce zero output, forming a sparse network. This mechanism helps to solve the overfitting problem and reduces the interdependence between parameters:

$$f(x) = \frac{1}{1+e^{-x}} \tag{6}$$

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{7}$$

$$f(x) = \begin{cases} \max(0, x) & x \ge 0 \\ 0 & x < 0 \end{cases} \tag{8}$$

3) Gradient descent based on logistic regression

Nonlinear functions need to be used for gradient descent using logistic regression. In the logistic regression method, a suitable prediction function needs to be selected such as Sigmoid function, tanh function or ReLU function as described above. These three functions divide the values into two different regions. Then the loss function is constructed and the logistic regression gradient is trained as shown in Figure 4 to quantify the difference between the prediction function and the training data labels. This loss reflects the difference between the predicted and actual choices in the training set. By considering the "loss" of all the training data, the deviation of the training data from the actual category is evaluated. Through the optimization process, the best parameter values are determined; the smaller the parameter values, the more accurate the prediction function and the better the training results.

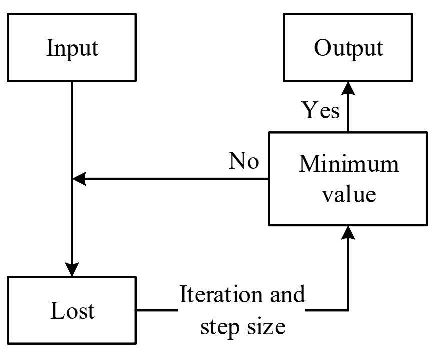

Figure 4: Logistic regression gradient training

(2) Analysis of K-means clustering algorithm

The core of K-means clustering algorithm lies in the process of dividing the data into K clusters. Initially, K objects are randomly selected as cluster centers and then, the algorithm starts calculating the distance between the data points and K. To calculate the distance, the article uses the Euclidean distance measure, which quantifies the distance between the data points and K.

To calculate the distance, the article uses the Euclidean distance metric, which quantifies the distance between the clustering centers and the distributed data. The distance between the data objects and the clustering center in the determined space is shown in equation (9) below:

$$(x,y)=\sqrt{((x_1-y_1)^2+(x_2-y_2)^2+\cdots+(x_n-y_n)^2)} = \sqrt{\sum_{i=1}^{n}(x_i-y_i)^2} \quad (9)$$

In equation (9), x denotes the data object, y denotes the center of clustering, and n denotes the dimensionality of the data object. The clusters generated by this classification method calculate the distance to the point with the smallest value of the other uniform values within the cluster and consider the center of mass as a reference point until it no longer changes.

III. Extraction and analysis of architectural genes of ancient villages

III. A. Architectural gene extraction

Based on the ancient village building identification model, the 24 most representative buildings (No. V1~V24) of the ancient village buildings in Henan Province were selected as cases to extract the architectural genes and construct the gene maps.

Research scholars in design, architecture and other related fields, as well as teachers and students from universities were invited to evaluate and score the buildings. Questionnaire research was conducted to assess and score the five types of architectural genes of the 24 representative buildings selected, and a five-level scale was adopted to analyze the architectural genes that best reflect the spirit of the times and national memories in each building from the user perception level. Questionnaires were designed according to 5 types of architectural genes: façade style, roof shape, decorative patterns, color matching and building materials, and respondents were invited to score the questionnaires. After finishing and recovering 100 valid questionnaires, the obtained data were sorted out by SPSS and analyzed through reliability analysis and regression analysis to carry out rigorous and scientific data analysis.

After the reliability analysis, the reliability coefficient of the questionnaire research of 24 buildings exceeded the acceptable value of 0.821, indicating that the results of the questionnaire have a high reliability. And the five architectural genes were defined as independent variables, and whether they embodied the architectural epochal characteristics of the ancient villages in Henan Province was defined as the dependent variable, and regression analysis was conducted to explore the independent variables that best embodied the epochal spirit and national memories in the architecture of the ancient villages in Henan Province from the perspective of the interviewees.

Table 1: Regression analysis of V1~V8

Item			V1	V2	V3	V4	V5	V6	V7	V8
	Regression parameter	R	0.754a	0.691a	0.745a	0.774a	0.661a	0.598a	0.655a	0.669a
		Rsq	0.454	0.418	0.492	0.334	0.506	0.432	0.473	0.466
		F	15.724	14.104	12.221	18.933	11.611	20.594	17.152	12.574
		Sigmf F	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Facade style	RC		0.203	0.294	0.195*	-0.041	0.136	0.244*	0.189	0.045
	T		1.865	2.541	2.458	-0.574	1.415	2.485	2.035	0.489
	sig		0.096	0.046	0.058	0.895	0.458	0.041	0.063	0.794
Roof style	RC		0.157	0.168	0.286**	0.243**	0.099	0.192	0.278**	0.056
	T		1.426	1.586	2.814	2.236	0.815	1.862	3.075	0.426
	sig		0.326	0.175	0.008	0.049	0.523	0.117	0.006	0.725
Decorative pattern	RC		-0.043	0.046	0.326	0.154	0.306	0.022	-0.032	0.245
	T		-0.356	0.268	2.103	1.245	2.351	0.163	-0.228	2.147
	sig		0.785	0.825	0.064	0.342	0.046	0.896	0.879	0.068
Color collocation	RC		0.182	0.325**	-0.084	0.224	0.048*	0.263*	0.348**	0.164
	T		1.528	3.106	-0.825	1.658	0.695	2.846	3.748	0.997
	sig		0.178	0.005	0.596	0.149	0.042	0.022	0.003	0.462
Building material	RC		0.428**	-0.053	0.178	0.326**	0.163	0.224*	0.194	0.297*
	T		2.815	-0.422	1.632	2.968	0.996	2.211	1.842	2.352
	sig		0.006	0.795	0.204	0.003	0.368	0.041	0.128	0.046

The results of regression analysis are shown in Tables 1 to 3, where RC is the regression coefficient. From the results of regression analysis, it can be seen that there is a close correlation between the façade style gene, roof shape gene, decorative pattern gene, color matching gene, and building material gene and the era characteristics reflected by different style types of buildings, and the degree of close correlation between the five independent variables and the dependent variable varies from building to building, which indicates that the degree of emphasis on the characteristics of the era of the intersection of the era embodied in the buildings of the different styles of the types of ancient villages in Henan Province is different. The values marked with “*” in the regression analysis table show that the architectural genes are closely related to the artistic style and period characteristics of the buildings, so the buildings with “*” are selected to extract the architectural genes to construct the gene maps.

Table 2: Regression analysis of V9~V16

Item	Regression parameter		V9	V10	V11	V12	V13	V14	V15	V16
		R	0.725a	0.734a	0.661a	0.739a	0.663a	0.764a	0.569a	0.653a
		Rsq	0.382	0.393	0.402	0.358	0.467	0.428	0.571	0.389
		F	19.224	21.199	26.121	17.803	19.829	20.532	19.773	15.561
		Sigmf F	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Facade style	RC		0.189	0.125	0.014	-0.025	0.222*	0.064	0.016	0.225*
	T		0.506	1.753	0.858	2.269	1.419	1.535	0.381	2.252
	sig		0.624	0.004	0.037	0.711	0.159	0.052	0.608	0.373
Roof style	RC		0.139	0.216*	0.255*	0.202**	0.123	0.253**	0.248**	0.153
	T		1.243	0.725	1.715	0.696	1.254	0.733	1.489	0.593
	sig		0.265	0.263	0.488	0.633	0.655	0.005	0.637	0.085
Decorative pattern	RC		0.358***	0.146	0.349**	0.153	0.186*	0.276	0.318**	0.385***
	T		1.154	2.086	1.364	1.963	1.804	0.854	2.371	1.553
	sig		0.584	0.125	0.627	0.532	0.411	0.556	0.153	0.678
Color collocation	RC		-0.158	0.301*	0.354***	0.118	0.289*	-0.176	0.117	0.286
	T		1.119	2.102	1.613	1.094	1.131	2.367	1.033	1.693
	sig		0.629	0.024	0.546	0.376	0.262	0.565	0.453	0.009
Building material	RC		0.265**	0.202	0.132	0.183	0.316***	0.142	0.149	0.285**
	T		0.982	0.727	1.033	0.762	1.092	1.498	1.279	1.465
	sig		0.169	0.499	0.328	0.379	0.013	0.115	0.007	0.092

Table 3: Regression analysis of V17~V24

Item	Regression parameter		V17	V18	V19	V20	V21	V22	V23	V24
		R	0.705a	0.656a	0.696	0.545a	0.566a	0.697a	0.667a	0.701a
		Rsq	0.451	0.514	0.562	0.445	0.591	0.549	0.434	0.577
		F	20.499	26.616	30.571	15.455	26.084	18.831	23.076	20.227
		Sigmf F	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Facade style	RC		0.249*	0.065	-0.022	0.248*	0.022	0.173	0.071	0.085
	T		0.561	-0.187	-0.203	1.333	1.103	1.114	0.881	0.419
	sig		0.599	0.346	0.003	0.478	0.223	0.234	0.199	0.781
Roof style	RC		0.279	0.345**	0.286	0.012	0.358**	0.196	0.345**	0.157
	T		1.525	1.222	1.618	1.181	0.993	0.454	1.628	2.129
	sig		0.133	0.041	0.096	0.461	0.717	0.337	0.744	0.026
Decorative pattern	RC		0.315**	0.394**	0.195	0.116	0.108	0.073	0.178	0.283*
	T		1.343	1.153	1.522	0.966	1.223	2.294	1.207	1.496
	sig		0.566	0.321	0.017	0.568	0.603	0.595	0.023	0.545
Color collocation	RC		0.262*	0.099	0.274*	0.069	0.049	0.298*	0.186	0.157
	T		1.747	0.551	2.206	1.368	2.301	0.241	1.179	1.055
	sig		0.447	0.089	0.072	0.627	0.642	0.205	0.116	0.431
Building material	RC		0.171	0.046	0.142	0.131	0.155	0.034	0.216	0.111
	T		0.389	0.925	0.761	2.046	0.645	1.542	0.567	1.517
	sig		0.335	0.035	0.436	0.789	0.342	0.048	0.419	0.726

III. B. Quantitative building analysis

The main monolithic buildings of ancient villages in Henan Province include three basic types of main houses, compartments and inverted seats, and their plan layouts are characterized by their own features. The results of the quantitative analysis of the architectural layout are shown in Figure 5.

The research data show that the main house, as the living space of the owner of the compound, has a depth of 1 room and a width of 3 to 7 rooms, with 6 rooms being the most. According to the statistics of this study on ancient villages in Henan Province, the 6-room zigzag and three-liter, two-dark, two-sleeve layout constitutes the prototype of the landscape genes of the main house of the houses, and thus generates a series of variants, enriches the type of the main house building plan, and highlights the central position of the main house in the compound. The main room in the plan of the main house in the shape of a zigzag is centered in the bright room of the hall, which serves as a place for family deliberation and enshrinement of ancestral tablets, and bedrooms are often formed by partitions on both sides, with kangs set up at the end of the room, and the main plan variations include the increase or decrease in the number of houses horizontally, the installation of a front-eaved platform or a holding building vertically, and the formation of a two-storey pavilion vertically, with the changes in the number of houses dominated by the 3- and 7-room rooms, and the number of variations of two-storey pavilions is one and a half times as many as that of the original model, which is far more than the rest of the variants. Because of the strong adaptability of the zigzag form and its variants, it is mostly used in compounds that are more restricted by the terrain and the residence base. The two-sleeve plan is based on the zigzag pattern, with a side room protruding from the end room on each side, which is called a “sleeve”, forming a “concave” plan, with the main hall located in the bright and secondary rooms, and the bedrooms in the end room, and the kang located in the sleeve, with the main variations of the plan being the horizontal increase and decrease of the number of houses, and the vertical increase and decrease of the number of sleeves. The main plane variants are horizontal increase or decrease in the number of houses, vertical increase in the number of sleeve flinging rooms with the front gable to hold the building, vertical formation of two-story pavilion and other different forms, of which the prototype and the latter two variants of the form of the proportion is similar, and less change in the number of houses. This sleeve layout of the internal separation of movement and static and privacy, external protection against wind and sand, cold and warmth of the role, commonly used in the less restricted by the terrain, better economic conditions, higher social status of the large mansions.

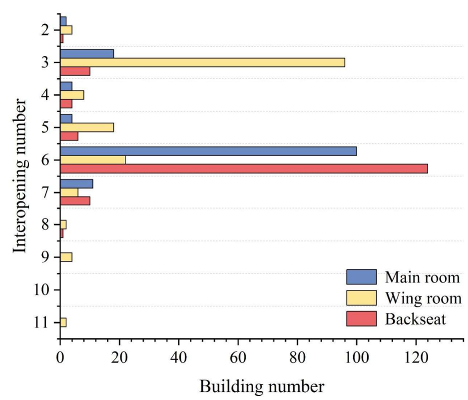

Figure 5: Quantitative analysis of building layout type

By synthesizing the above gene recognition contents of courtyard pattern and building layout, the building plan atlas of ancient villages in Henan Province is established as shown in Fig. 6 and Fig. 7. By analyzing the architectural layout of the ancient villages in Henan Province, it can be seen that the ancient villages in Henan Province used the main room as the living room, the compartments as the residence of the children of the courtyard owners, and the inverted seat as the auxiliary rooms such as storage and toilet, etc. The depth of the three types of rooms was 1 room, and the width of the rooms was 3~7 rooms. Among them, the prototype of the main room is 6-room zigzag with three bright and two dark two-sleeve layout, the compartment room is 3-room zigzag layout as the landscape gene prototype, and the prototype of the inverted seat is mainly 5-room zigzag layout.

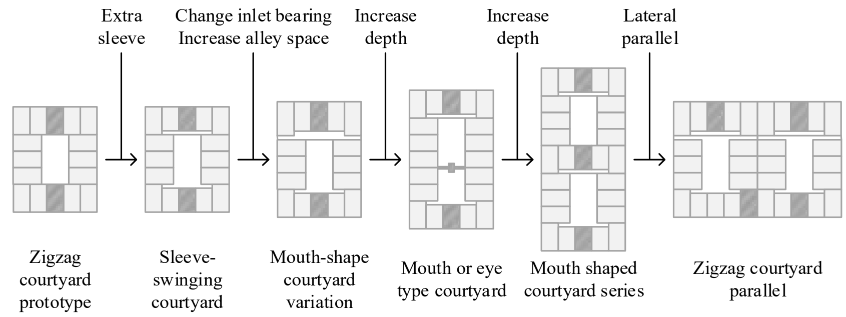

Figure 6: Compound type

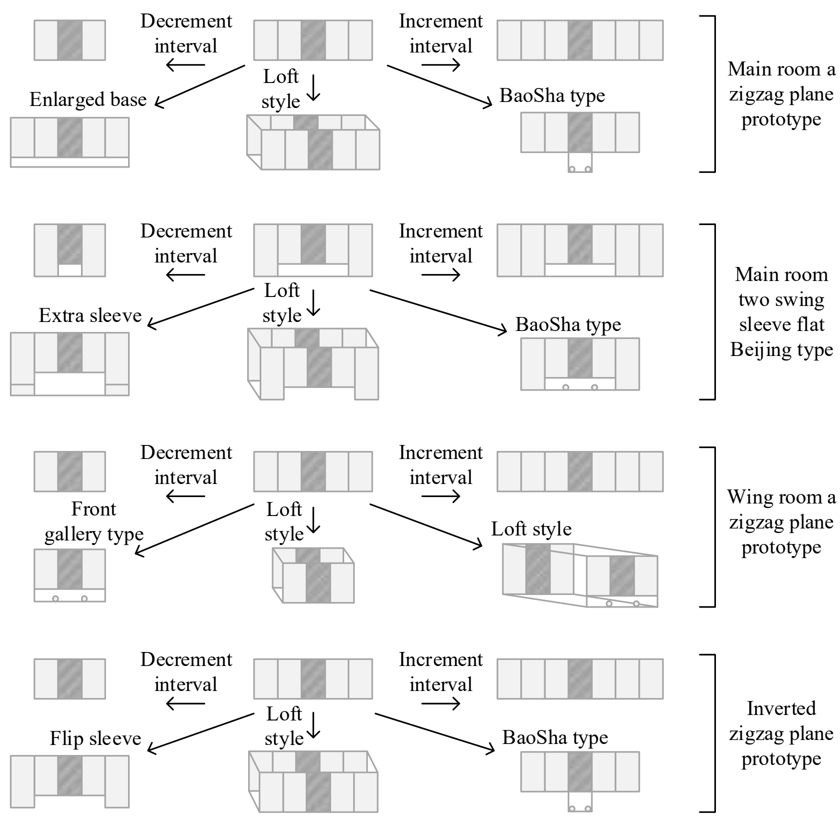

Figure 7: Flat layout

IV. Conclusion

Out of the traditional methods of field research and information collection, the article identifies the overall architecture of ancient villages in Henan Province through the improved U²-Net model, and then deeply analyzes the architectural colors using deep learning algorithms, in order to realize the accurate extraction of the overall and detailed data of the architecture of ancient villages. By extracting the architectural genes of the ancient villages in Henan Province, the architectural genes are quantified and an architectural gene map is constructed. From the perspective of five aspects, namely, façade style, roof shape, decorative patterns, color matching and building materials, there is a close correlation between the ancient village buildings in Henan Province and the characteristics of the times reflected in the typical buildings extracted in this paper. The architectural layout of the ancient villages in Henan Province is based on the main room as the living space, and most of the rooms are 6 rooms in width. Most of the architectural prototypes are in the shape of a 6-room zigzag with three bright, two dark, and two fluttering sleeves, and the number of two-story attic variants is 1.5 times as many.

References

[1] Dong, X. (2020). Researches and Protection of Bang-tang Ancient Village Architecture Decoration in Leizhou. Acad. J. Humanit. Soc. Sci, 3, 134-147.

[2] Xu, L., Sang, K., Li, G., Lin, G., Luo, Q., & Giordano, A. (2023). Heritage evaluation and analysis based on entropy weight method: the study of Wengji ancient village in China. Journal of Housing and the Built Environment, 38(3), 1843-1868.
[3] ZHONG, H. Y., & LI, Q. Y. (2017). The Study of Village Architecture and the Development of Ancient Villages: Taking Pazhou Ancient Village in Guangzhou as Research Object. Journal of Literature and Art Studies, 7(9), 1223-1228.
[4] Tong, F. (2018, February). Study on the Strategy of Ancient Village Development in the New Era. In 6th International Conference on Social Science, Education and Humanities Research (SSEHR 2017) (pp. 61-64). Atlantis Press.
[5] Beibei, W. A. N. G. (2024). THE CULTURAL LANDSCAPE OF HONGSHIYA VILLAGE, HENAN PROVINCE, CHINA. Procedia of Multidisciplinary Research, 2(10), 6-6.
[6] Cao, K., Liu, Y., Cao, Y., Wang, J., & Tian, Y. (2024). Construction and characteristic analysis of landscape gene maps of traditional villages along ancient Qin-Shu roads, Western China. Heritage Science, 12(1), 37.
[7] Hui, C. H. E. N. (2020). Analysis of landscape gene identification and its characteristics of traditional villages: A case study of Zhuge Bagua village. Journal of Landscape Research, 12(3), 101-107.
[8] Wang, L., Sun, C., Wang, M., & Xiao, X. (2024). Construction and Characterization of Traditional Village Landscape Cultural Genome Atlases: A Case Study in Xupu County, Hunan, China. Sustainability, 16(21), 9524.
[9] Zhang, W., & Yang, H. (2024). Quantitative Research of Traditional Village Morphology Based on Spatial Genes: A Case Study of Shaanxi Province, China. Sustainability, 16(20), 9003.
[10] Wei, K., He, Y., Wang, M., Zhu, R., & Wang, Z. (2025). Identification, inheritance and restoration of traditional village landscape gene: a case study of Lidipo Village in Tongchuan, Shaanxi Province. npj Heritage Science, 13(1), 18.
[11] Wu, H., Liang, T., & Shen, T. (2023). The Spatial Characteristics of Traditional Villages and Their Heritage Protection based on Landscape Genes. WSEAS Trans. Environ. Dev, 19, 320-328.
[12] Wang, W., Shi, Q., & Wang, G. (2024). Analysis of performance and genetic characteristics of cultural landscapes in traditional villages along the Jinzhong section of the Wanli Tea Road from a landscape gene information chain perspective: A case study of Xiamen Village. Sustainability, 16(18), 8131.
[13] Zhang, Y., & Huang, Z. (2025). Genetic Characteristics of Spatial Network Structures in Traditional Bouyei Village Architecture in Central Guizhou. Sustainability, 17(4), 1435.
[14] Wang, W., Shi, Q., & Wang, G. (2024). Exploration of the landscape gene characteristics of traditional villages along the Jinzhong section of the Wanli Tea Road from the perspective of the village temple system. Land, 13(10), 1602.
[15] Li, G., Chen, B., Zhu, J., & Sun, L. (2024). Traditional Village research based on culture-landscape genes: a Case of Tujia traditional villages in Shizhu, Chongqing, China. Journal of Asian Architecture and Building Engineering, 23(1), 325-343.
[16] Yu, C., Zhou, Z., Gao, J., Zhang, X., Zheng, Q., Liu, Z., ... & Wen, S. (2025). Multi-scale comparison of the formation mechanisms in landscape genes of traditional villages. Scientific Reports, 15(1), 1-16.
[17] Xia, H., & Chen, F. (2025). Recognition and optimization of landscape genes in traditional settlements: a case of Meishan area. Journal of Environmental Engineering and Landscape Management, 33(1), 85-96.
[18] Feng, Y., Wei, H., Huang, Y., Li, J., Mu, Z., & Kong, D. (2023). Spatiotemporal evolution characteristics and influencing factors of traditional villages: the Yellow River Basin in Henan Province, China. Heritage Science, 11(1), 97.
[19] Liu, W., Xue, Y., & Shang, C. (2023). Spatial distribution analysis and driving factors of traditional villages in Henan province: a comprehensive approach via geospatial techniques and statistical models. Heritage Science, 11(1), 185.
[20] Wang, W., Shi, Q., & Wang, G. (2024). Analysis of performance and genetic characteristics of cultural landscapes in traditional villages along the Jinzhong section of the Wanli Tea Road from a landscape gene information chain perspective: A case study of Xiamen Village. Sustainability, 16(18), 8131.
[21] Dang, A., Zhao, D., Chen, Y., & Wang, C. (2020). Conservation of cave-dwelling village using Cultural Landscape Gene Theory. Spatial synthesis: Computational social science and humanities, 97-105.
[22] Yadav Lav K. & Wilde H. Dayton. (2022). Identification and bioinformatic characterization of rare variants of Rhododendron canescens architecture genes. Euphytica,218(6).
[23] Hongtao Xiao,Yingfang Zhu,Yurong Sun,Gui Zhang & Zhiwei Gong. (2024). Wildfire Spread Prediction Using Attention Mechanisms in U^2-NET. Forests,15(10),1711-1711.
[24] Qingan Yao,Han Xu,Yuncong Feng,Xuexiao Wang & Congmin Zhang. (2025). YOLO-MFX: lightweight YOLO with improved flame detection for small targets. Journal of Real-Time Image Processing,22(2),92-92.