Journal of Network Intelligence
Taiwan Ubiquitous Information
 ISSN 2414-8105
Volume **10**, Number **4**, November **2025**

Visual Communication Based on Multi-label Classification: Boosting Design Creativity with the Digital Tool Procreate

Ying Xu*

Zhuhai College of Science and Technology
Zhuhai 519041, P. R. China
Faculty of Creative Technology and Heritage
Universiti Malaysia Kelantan
Kuantan State 16300, Malaysia
xuying202211@163.com

*Corresponding author: Ying Xu
Received September 5, 2024, revised April 20, 2025, accepted September 6, 2025.

Abstract. *Aiming at the issue that existing visual communication methods have low correlation between image features and labels, which results in low efficiency of visual communication, this paper designs a visual communication approach relied on multi-label classification. Firstly, the low-brightness image was converted to Lab color mode by Procreate software, and the Curvelet transform was used to decompose the Procreate-converted image into high-frequency components and low-frequency components, which described the details and the overall brightness of the image, respectively. On this basis, the high-frequency and low-frequency components are enhanced by the channel-domain and spatial-attention mechanisms, respectively, so as to recover the detail information of the image and enhance the overall brightness, and then the hybrid-domain attention feature maps of the high and low-frequency components are obtained by the splicing and fusion method. A Graph Convolutional Network (GCN) is adopted to model the ground relationship between feature map labels and to build a global visual optimization model of the image to enhance the global details of the image. The experimental outcome indicate that the average classification accuracy and Peak Signal-to-Noise Ratio (PSNR) of the proposed method are 91.82% and 49.5 dB, respectively, which are significantly better than other current approaches, and not only improve the classification performance of the image, but also significantly improve the quality of the image.*
Keywords: multi-label classification; visual communication; Procreate software; graph convolutional network; attention mechanism

1. **Introduction.** Visual communication is a process in which a communicator uses visual symbols to convey specific information to the public and influence them in order to achieve a set goal or purpose [1]. Because of its functional and purposeful characteristics, visual communication combines the rationality of science and the sensibility of art. As the information technology rapidly growing, especially the emergence of artificial intelligence, visual communication design has also gradually broken through the traditional design mode, more designers began to get rid of the traditional paper media or other two-dimensional plane media, with the help of new media fusion of new technology to try more diversified design methods, and the participation of the computer in the design process is greatly increased [2, 3]. Image visual classification is a fundamental task in

visual communication, it is crucial to study image categorization well for the application of visual communication.

1.1. **Related work.** Li et al. [4] aimed to improve the visual quality of images by color space model transformation, combined with exposure interpolation method and multi-scale fusion strategy. Jung et al. [5] enhanced the wavelet decomposed low-frequency image and high- frequency image separately, and the Peak Signal-to-Noise Ratio (PSNR) was improved by 1. 2%∼6. 9%. Hu et al. [6] extracted visual features based on the information of colors, textures, and edges in the image, and SVM is used for classification, but the classification efficiency is not high. Zheng et al. [7] introduced probabilistic analysis to model the features of individual components in an image and combined it with global image features to classify the vision. Streeb et al. [8] performed image generation through digital tool Procreate and extracted its visual features and used decision tree for visual classification of images. Nithyananda and Ramachandra [9] proposed unsupervised features using K-mean clustering algorithm for visual enhancement. With the significant results of deep learning in the field of visual communication, Liu et al. [10] used CNN to extract texture feature information of an image for visual classification. Zhang et al. [11] suggested a two-row deep convolutional network framework to gain geometrically-sensitive and location-aware characteristic representations for visual categorization of images. Hui et al. [12] introduced an attention scheme to focus on the style of an image Significantly related feature regions. For the visual fine-grained analysis of this complex scene, single-label classification methods can hardly meet the needs of the industry. Multi-label image categorization aims to forecast multiple possible labels for a given image at the same time [13], which can interpret the rich semantic content of the image at a fine-grained level and provide basic image information for visual communication applications. Li et al. [14] used CNN for feature extraction and Recurrent Neural Network (RNN) to capture feature label correlation and optimize the visual communication based on the prediction of each label, but the optimized images were not clear. Zhou et al. [15] learned the attention feature maps of all labels to explore the potential relationship between labels and combined them with regularized classification to improve the visual communication performance. Lyu et al. [16] introduced the cyclic attention mechanism and proposed a method based on the adaptive search of target category to search for the region of attention and contextual correlation to enhance the visual communication performance of the image. In multi-label categorization problems, capturing label correlation usually requires network structures suitable for relational modelling [17], CNN and RNN models have difficulties in establishing connections between labels and cannot capture label correlation well.

Graph Neural Networks (GNN) are usually used for tasks such as graph analysis and generating node word embeddings, while the dependencies between labels in a multi-label classification task can be represented by nodes and edges in the graph structure. Based on the above idea, Cheng et al. [18] used Graph Convolutional Network (GCN) for multi-label image classification and visual optimization based on the results of the classification, with a PSNR of only 47dB. Qu et al. [19] used GCN to extract the semantic information of the image related to the labels for visual communication optimization.

1.2. **Contribution.** In summary, existing research has not fully combined the characteristics of the image itself for visual communication optimization, resulting in inefficient visual communication, for this reason, this paper designs a visual communication method based on multi-label classification. Firstly, to address the issue of poor application of RGB channel to image processing, Procreate software is used to convert the low-brightness image into lossless Lab color mode, and the Curvelet transform is introduced to decompose

the image in Lab color mode into high-frequency component and low-frequency component, which describes the details and overall brightness of the image, respectively. The two components are then enhanced using the channel-domain attention mechanism and the spatial attention mechanism, respectively, whereby the image details are recovered and the overall brightness is enhanced, and the enhanced results are spliced and fused to obtain a hybrid-domain attention feature map. Finally, the relationship between feature map labels is modelled using GCN to establish a global visual optimization model for images. The experimental outcome indicates that the classification accuracy of the proposed method is improved by 3.24%–11.1%, the PSNR is improved by 0.5dB–3dB, and the image classification performance is higher and the visual quality is better.

2. Theoretical analysis.

2.1. Design for visual communication.

Visual communication refers to the use of vision as the art form of communication and expression, and the use of graphics, text, and color as design elements, the need to communicate and valuable information for the image, intuitive conception and design, so as to improve the quality and effect of information communication [20].

(1) Graphics in visual communication design, i.e. illustrations, can be used to perceive the information conveyed by the graphic language through associations, and from the perspective of digital media, diagrams can be divided into two types, graphics and images, graphics are composed of external contour lines and fill color blocks, which is a scaled and undistorted image format. Compared with images, its shape is relatively simple and color variations are not rich enough [21].

(2) Graphic text is evolved from graphics; symbolic text is to give special meaning to a certain symbol. As there are many kinds of text, in visual communication design, graphic text is not separated, the text plays the role of illustration, is the interpretation and interpretation of the graphic meaning.

(3) Color elements in visual communication design are mainly used to enrich the textual and graphic contents conveyed, stimulate the visual nerves of the viewers and arouse their attention and interest.

2.2. Graph neural network.

GNN is a type of neural network used to process graph structure data and is classified into GCN and Graphical Attention Network (GAN) [22]. GCN uses convolutional computation for graph structure computation and all the parameters are shared, which not only reduces the amount of parameters that need to be solved but also reduces the complexity of the network as compared to GAN, so this paper utilizes GCN for multi-label classification of visual images. GNN consists of three elements: nodes, neighbors and edges, nodes contain label categories and corresponding feature sequences, and edge weights represent the correlation between labels. GNN treats each label as a node, and correlations between labels as edges, and mines and learns label correlations through graph propagation mechanism [23]. $G = (N, E)$ is denoted as a collection of nodes and edges, where node $n \in N$ denotes an object and edge $e \in E$ denotes a relationship. All nodes and edges get the node state x_n through the local transfer function f_w. f_w is used to update the current node state. The predicted label o_n of a node is obtained from the feature prediction function g_w based on the features and state of the node, shown below.

$$x_n = f_w\big(l_n, l_{\mathrm{co}[n]}, x_{\mathrm{pe}[n]}, l_{\mathrm{pe}[n]}\big) \tag{1}$$

$$o_n = g_w\big(x_n, l_n\big) \tag{2}$$

where $l_n, l_{\mathrm{co}[n]}, x_{\mathrm{pe}[n]}, l_{\mathrm{pe}[n]}$ are denoted as the features of node n, the features of the edges of its nodes, and the features and states of the nodes neighboring node n, respectively. An example of the computation of neighboring nodes in the graph structure is shown in Figure 1.

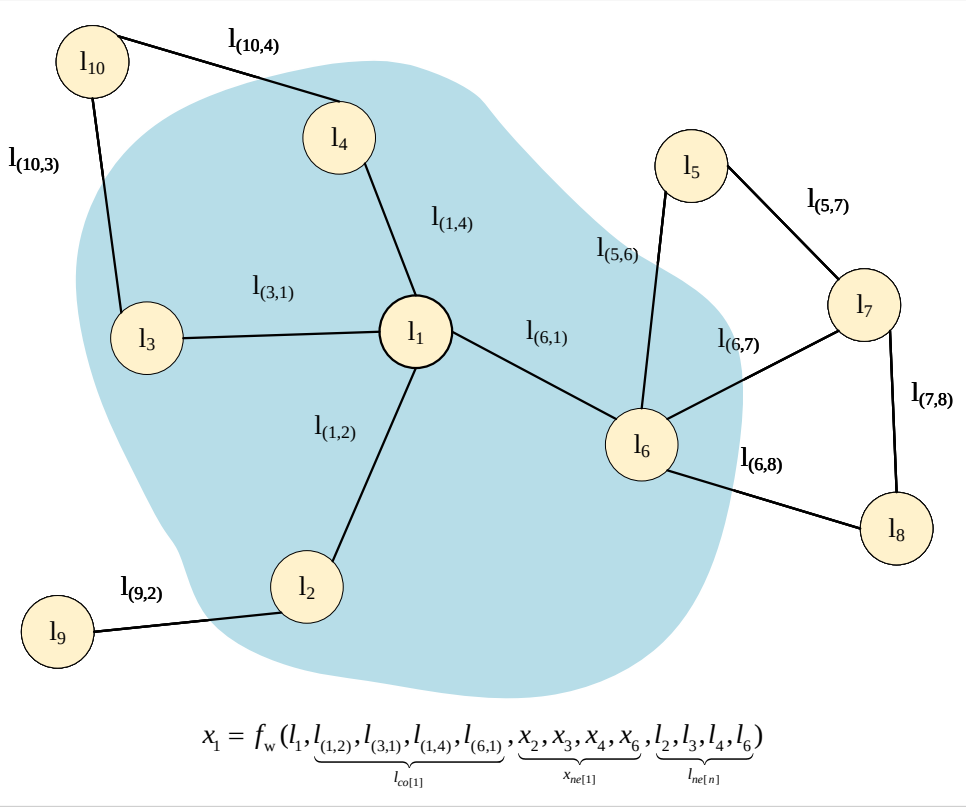

Figure 1. The computation of neighboring nodes in the graph structure

The GNN is trained by defining a loss function to iteratively optimize the parameters, and by recursively and iteratively transferring information between nodes according to the topology, thus updating the state of the hidden layer of each node. The node state update expression is shown below.

$$x_n(t+1) = f_w\big(l_n, l_{\mathrm{co}[n]}, x_{\mathrm{ne}[n]}(t), l_{\mathrm{ne}[n]}\big) \tag{3}$$

$$O_n(t) = g_w\big(x_n(t), l_n\big) \tag{4}$$

3. **Visual image pre-processing based on Procreate software and Curvelet transforms.**

3.1. **Color model conversion based on Procreate software.** To enhance the visual image quality, enhancement of image clarity and correction of image color are considered to enhance the visual communication of the image. The clarity of the picture is mirrored in the high- frequency portion of the picture, and the color of the picture is mirrored in the low-frequency portion of the picture, therefore, the high-frequency portion of the picture and the low- frequency portion of the picture are separated, and the clarity and the distortion of the image's color are enhanced together to achieve the enhancement of the image quality. Procreate is a powerful digital tool with powerful color conversions that can enhance creativity in painting [24]. Since Lab color mode is more compatible

with human physiological visual communication than RGB color mode, it covers the full range of colors visible to the human eye [25]. Therefore, the replacement of any color mode with Lab color mode by Procreate software will not cause any loss of color in the image, provided that the number of color bits is sufficient, the conversion principle is implied in Figure 2. Only luminance information exists in the luminance channel L, and only color information exists in the a and b channels, but there is no direct conversion between Lab and RGB color modes, so the mapping relationship in Equation (5) can be used to convert the RGB color mode of the image to XYZ color mode.

bl

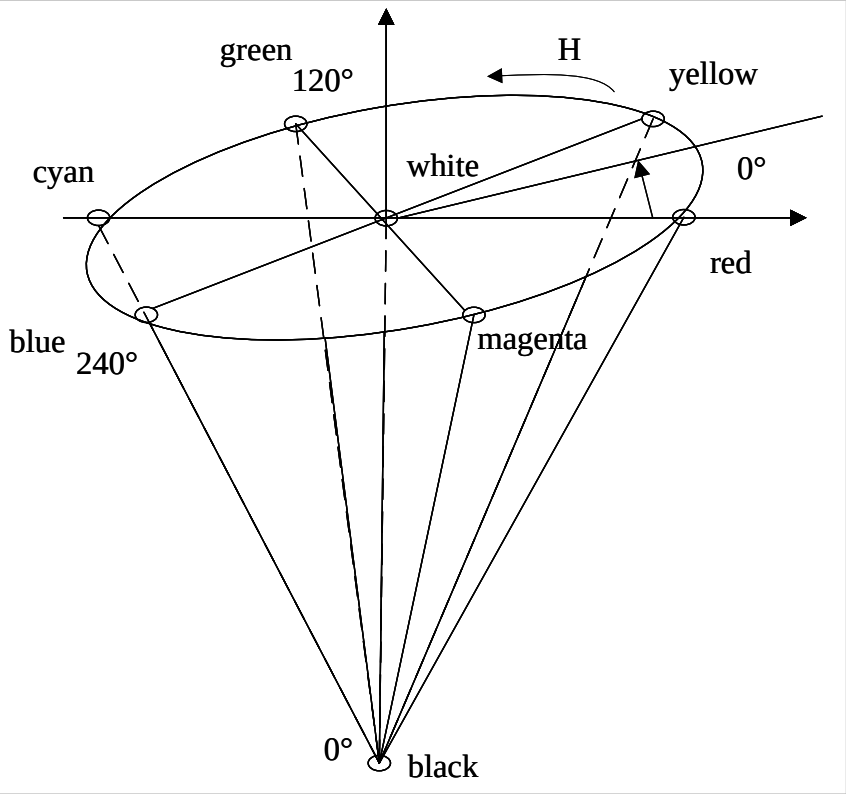

Figure 2. Lab Color Model

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4125 & 0.3576 & 0.1904 \\ 0.2127 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9503 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{5}$$

Convert the XYZ color mode to Lab color mode using the conversion relationship shown in Equation (6).

$$\begin{cases} L = 116 \times f\left(\dfrac{Y}{Y_\delta}\right) - 16 \\ a = 500 \times \left[f\left(\dfrac{X}{Y_\delta}\right) - f\left(\dfrac{Y}{Y_\delta}\right) \right] \\ b = 200 \times \left[f\left(\dfrac{Y}{Z_\delta}\right) - f\left(\dfrac{Z}{Z_\delta}\right) \right] \end{cases} \tag{6}$$

where X_δ, Y_δ, and Z_δ are parameters that guarantee that the mapping range of the RGB color mode is the same as that of the XYZ color mode. g denotes a function, as shown below.

$$g(t) = \begin{cases} t^3, & t > \left(\frac{6}{29}\right) \\ \frac{1}{3}\left(\frac{29}{6}\right)^2 t + \frac{4}{29}, & \text{otherwise} \end{cases} \tag{7}$$

where t denotes any single channel value of the XYZ color mode optimized by the mapping range.

3.2. **High and low frequency component division based on Curvelet transforms.** After converting the picture's color mode to Lab color mode, the picture of Lab color mode is decomposed into high-frequency and low-frequency constituents by using the Curvelet transform, which is based on the ridge wave theory and contains two components of the multi- scale ridge wave transform and filtering, and the edge information of the image can be described more sparsely by the Curvelet transform [26]. Curvelet transformations are as follows.

$$\alpha_0(\omega) + \sum_{s \geq 0} \lambda_s(\omega)^2 = 1 \tag{8}$$

where α_0 and λ_s $(s = 1, 2, \ldots, n)$ denote a low-pass filter and a band-pass filter respectively with bandwidths of $|\omega| \leq 2$. $f(x_1, x_2)$ denotes the two-dimensional product function of a low-luminance image in Lab color mode, and its Curvelet transform can be described as follows.

(1) Mapping $f(x_1, x_2)$ into different bandwidths through α_0 and λ_s, thus obtaining different frequency bands.

$$P_0 f = \alpha_0 * f, \quad \Delta_0 f = \varphi_0 * f, \ldots, \ \Delta_s f = \varphi_s * f \tag{9}$$

where $P_0 f, \Delta_0 f$ and $\Delta_s f$ denote different frequency bands and f denotes the bandpass component.

(2) The Curvelet transform coefficients $P_0 f$ are obtained by applying single-scale ridge wave transform to different bandpass components with A_μ fixed.

$$A_\mu = [\Delta_s f, \ \lambda_{Q,\alpha}] \tag{10}$$

where $Q \in Q_s, \ \Lambda \in \Gamma, \ \lambda_{Q,\alpha}$ and Q denote the multiscale ridge-wave dictionary and the binary square, respectively, Q_s and Γ denote the set of overall binary squares and the parameter space of ridge-wave transform under the scale s condition, respectively.

Based on the above description, the discrete Curvelet transform of a 2D low brightness image $f(x, y)$ in Lab color mode can be obtained. The J-level 2D wavelet decomposition is applied to $f(x, y)$ as follows.

$$f(x, y) = C_J(x, y) + \sum_{j=1}^{J} w_j(x, y) \tag{11}$$

where C_J and w_j denote the approximate or low-frequency component of $f(x, y)$ and the high-frequency component of $f(x, y)$, respectively. C_J is fixed and w_j is chunked. In order to reduce the edge effects generated in the chunking process, smoothing and overlapping chunking are used in the processing [27]. Equation (12) shows the smoothing window requirement.

$$\sum_{Q \in Q_s} w_Q^2(x_1, x_2) = 1 \tag{12}$$

The correlation between the scale of the chunks and the scale j can be described by Equation (13)

$$E = 2^{\lceil \frac{j}{2} \rceil} \times E_{\min} \tag{13}$$

The Curvelet transform coefficients of $f(x, y)$ are obtained by applying the ridge-wave transform to different sub-blocks.

4. Optimization of visual communication effects based on multi-label classification.

4.1. High and low frequency component enhancement based on mixed domain attention. To improve the efficiency of visual communication, this paper uses the multi-label classification to optimize the visual communication effect, and the overall framework is shown in Figure 3. The high and low frequency components are augmented using the channel domain attention mechanism and the spatial attention mechanism respectively, and the augmented outcome are spliced and fused to generate the hybrid domain attention feature map. GCN is used to model the ground dependencies between feature map labels to establish a global visual optimization model, thus improving the quality of visual communication.

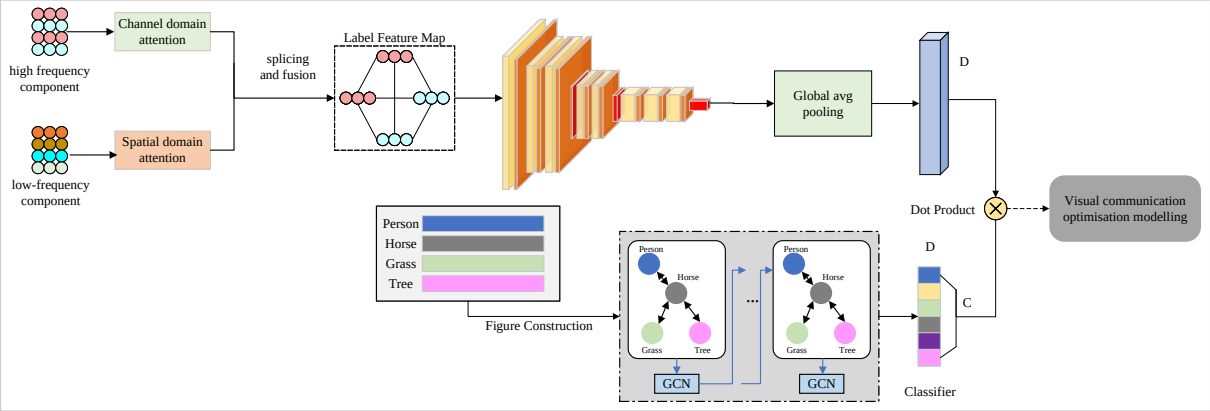

Figure 3. Structure of the visual communication method based on multi-label classification

After decomposing $f(x, y)$ into high-frequency and low-frequency components by the Curvelet transform, the two components are enhanced separately. For the high-frequency component, channel-domain attention [28] is used to enhance the details, and the details in $f(x, y)$ are restored to obtain the enhanced high-frequency component; for the low-frequency component, spatial-domain attention [29] is adopted to enhance the luminance of $f(x, y)$, and the enhanced low-frequency component is obtained.

(1) High-frequency component enhancement. Since the pixel values within the high-frequency component are usually small, the attention mechanism is very suitable for the high- frequency component training process of $f(x, y)$, which achieves effective processing and utilization of visual information by selectively focusing on parts of the input information. Each channel map of the high-frequency component of an image can be viewed as a category- specific response, and by exploiting the interdependencies between channels, the relevance of the tags of the picture are able to be mined to enhance visual communication.

Assuming that the characteristic picture $B \in \mathbb{R}^{C \times H \times W}$ of the high-frequency components, the dimensionality of B is compressed to obtain $B \in \mathbb{R}^{C \times M}$, where $M = H \times W$.

Finally, adopting Sigmoid function to obtain the channel domain attention characteristic picture $Y \in \mathbb{R}^{C\times C}$ as bellow.

$$y_{ji} = \delta\big(B_i \cdot B_j\big) \tag{14}$$

where y_{ji} stands for the effect of the i-th channel on the j-th channel and δ is the ReLU function. Then the transpose of Y and B are matrix multiplied to get the output $\in \mathbb{R}^{C\times H\times W}$. Finally, it is multiplied element by element with B to get the final output $X \in \mathbb{R}^{C\times H\times W}$.

$$X = \sum_{i=1}^{C} \big(y_{ji}\, B_i\big) + B_i \tag{15}$$

The above operations enable the modelling of long-range dependencies between channels and the enhancement of the correlation between high-frequency component labels.

(2) Low frequency component enhancement. The colors, contours, shapes etc. of the image are mainly in the low-frequency part of the image. In this section, the spatial-domain attention module is used to enhance the reciprocity among the picture and the tags of the low- frequency components. Assuming a feature map $A \in \mathbb{R}^{C\times H\times W}$ of the low-frequency components, A is passed through a convolutional level with a convolutional kernel of 1×1, a BN layer with a ReLU excitation operation to obtain a characteristic picture $t \in \mathbb{R}^{C_d\times H\times W}$ based on the categories of the dataset, where C_d is the amount of classes. The characteristic pictures are then aggregated to produce the characteristic pictures of $1 \times C \times H$ using a convolutional layer with convolutional kernel 3×3. Finally, the characteristic picture is normalized to gain the enhanced characteristic picture X_s.

$$t = \tau\big(\mathrm{BN}\big(f_{1\times1}(A)\big)\big) \tag{16}$$

$$X_s = \delta\big(\mathrm{BN}\big(f_{3\times3}(t)\big)\big) \tag{17}$$

(3) Pool X_s and X_c respectively, and then splice them along the channel dimension direction to obtain the fused high and low frequency hybrid domain attention feature maps as follows.

$$X = \mathrm{concat}\big(\,\mathrm{avgpool}(X_s),\ \mathrm{avgpool}(X_c)\big) \tag{18}$$

where avgpool denotes the average pooling layer and concat denotes the splicing operation.

4.2. **Optimization of visual communication effects based on multi-label classification.** To capture the correlation of different feature labels in the high and low frequency components for subsequent optimization of visual communication, this paper uses GCN to effectively model the relationship between the labels of attention feature maps in the high and low frequency mixed domains on improving the performance of visual communication. The input to the classifier consists of nodes as well as a label correlation matrix, where the nodes are represented by the feature vectors of the labels and the label relationship matrix is constructed from the co- occurrence matrix of the labels.

If the dependency of labels is modelled using conditional probability $P(W_j \mid W_i)$, but $P(W_j \mid W_i) \neq P(W_i \mid W_j)$, it means that the label co-occurrence relationship matrix is asymmetric. To build the relationship matrix, this paper first counts the number of occurrences of each tag pair and then builds the relationship matrix. This is shown as follows.

$$\mathrm{cooc}(W_i, W_j) = \begin{cases} 1, & \mathrm{Label}(W_i) = 1 \wedge \mathrm{Label}(W_j) = 1 \\ 0, & \text{others} \end{cases} \tag{19}$$

$$N_{ij} = \sum_{\Omega} \mathrm{cooc}(W_i, W_j) \tag{20}$$

where W_i denotes the occurrence of label i in the image, $\text{cooc}(W_i, W_j)$ denotes whether W_i and W_j co-occur. 1 denotes that W_i and W_j co-occur, 0 denotes that W_i and W_j do not co- occur, and N_{ij} denotes the number of times W_i and W_j co-occur in the image. The conditional probability matrix using the label co-occurrence matrix is $P = N_i/T_i$, where T_i denotes the total number of occurrences of W_i and $P_{ij} = P(W_j \mid W_i)$ denotes the probability of W_i occurring when W_j occurs. Since the label distributions of the test and training datasets may be inconsistent, noise exists. Therefore, the relationship matrix P with threshold ρ is binarized to filter the noise as follows

$$R_{ij} = \begin{cases} P_{ij}, & P_{ij} \geq \rho \\ 0, & P_{ij} < \tau \end{cases} \tag{21}$$

where R_{ij} is the binary matrix of label correlations.

Label dependency is implicitly modelled through labels by learning the information propagated between nodes through GCN. The mapping function of the GCN is learnt by learning from the label representation, i.e., $W = \{w\}^C$ (where C denotes the label species). Using stacked GCN layers, each layer l takes as input the node representation X^l of the previous layer and generates an updated node representation X^{l+1} as output. The first layer takes the label embedding as input and the last layer outputs $W \in \mathbb{R}^{C \times D}$, where D is the dimension of the label feature representation. The predicted category scores can be obtained by applying the learned classifier to the image representation as follows.

$$\hat{y} = Wx \tag{22}$$

where x represents the image representation.

For the goal of further optimizing the visual communication, a global visual optimization model of the image needs to be built using the different feature labels of the image's high and low frequency hybrid domain attention feature maps as follows.

$$\theta_{\text{Global}}(A, B) = \frac{W}{\varepsilon(A, B)} \tag{23}$$

where the visual optimization output is $\theta_{\text{Global}}(A, B)$, the neighbourhood standard deviation is $\varepsilon(A, B)$, and the global feature labels are W.

To meet the adaptive requirements, according to the model established above to develop the best objective function, so as to better complete the image detail enhancement, to achieve the optimization of the image visual information communication effect, the objective function is shown in the following equation.

$$\vartheta = \log\left(\frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\frac{(G-1)^2}{\left\|f(A,B) - f(A,B)\right\|^2}\right) \tag{24}$$

where the established objective function is ϑ, the image gray level is G, i and j represent the labels corresponding to the high and low frequency components of the image respectively, and the expression of function $f(\cdot)$ is as follows.

$$H^{l+1} = h\big(\hat{D}\, H^l W^l\big) \tag{25}$$

where $h(\cdot)$ denotes the convolution operation, $\hat{D}$ is the normalized form of the relation matrix D, and W^l is the weight matrix parameter to be learned.

Based on the above multi-label classification of images, the completeness of the visual information conveyed in images is improved, and the global details of images are enhanced, so as to optimize the effect of visual information conveyed in images.

5. **Experiments and analysis of results.**

5.1. **Classification performance analysis.** The DLRSD visual dataset [30] is used in the experiments, which contains 2109 images with 10 categories of labels (buildings, cars, bushes, flowers, fields, lawns, pavements, beaches, trees, animals). In this paper, the color models of these images are transformed by the Procreate software, where 80% of these images are used as the training set, and the rest are used as the test set. The hardware environment is an AMD 2600X Six-Core processor at 3.8 GHz, and MATLAB is used as the development tool for the multi-label visual optimization method. To facilitate the analysis, the proposed method is compared and analysed with other methods, the method in the literature [16] is denoted as VA-RNN, the method in the literature [18] is denoted as ML-GCN, the method in the literature [19] is denoted as SR-GSN, and the method in this paper is denoted as VC-GCN, and the evaluation metrics of the classification performance are chosen as the accuracy of the individual category labels (AP), the average classification accuracy (mAP) , average category precision (CP), average category recall (CR) and average F1 (CF1). The classification accuracy of the proposed method on various types of images is shown in Figure 4, the recognition accuracy of buildings, cars, flowers, fields, pavements, beaches and animals is above 90%, but the classification accuracy of bushes, lawns and trees is relatively low. Due to the more similar colors and shapes in these three types of pictures, the classification accuracy is not as good as that of the other seven, but the overall classification accuracy is above 90%, which verifies the effectiveness of the designed approach. A comparison of the classification performance of the proposed method with the other three methods is shown in Table 1. VC-GCN has the best mAP, CP, CR, and CF1 metrics among all the methods, with mAP and CF1 of 91.82% and 91.96%, respectively, which are 11.1% and 13% higher than VA-RNN, 6.33% and 8.44% higher than ML-GCN, and 3.24% and 4.4% higher than SR-GCN, respectively. and 3.24% and 4.2% respectively compared to SR-GCN. VA-RNN has the worst classification performance, which is based on image classification through the combination of RNN and attention mechanism, but it does not take into account the relevance of internal feature labels, which results in less efficient classification than ML- GCN. ML-GCN uses GCN for multi-label classification of images, but it does not enhance the high and low-frequency components in the image, and the classification effect is average.SR- GCN extracts semantic information related to the labels by using CNNs, but the distribution of the labels is noisy and it does not highlight the key parts of the image, which results in a worse classification performance than that of VC-GCN. Therefore, VC-GCN has a significant superiority in image classification.

Table 1. Image categorization performance comparison

Method	VA-RNN	ML-GCN	SR-GCN	VC-GCN
mAP (%)	80.72	85.49	88.58	91.82
CP (%)	77.65	82.94	86.41	93.51
CR (%)	80.31	84.11	89.17	90.46
CF1 (%)	78.96	83.52	87.76	91.96

5.2. **Visual communication effectiveness analysis.** To test the visual quality of images after performing multi-label classification, this paper sets image structural similarity (SIMM) and PSNR [31] as the quantitative test metrics for visual communication effects. The SIMM comparisons of the four methods under different number of images are shown in Table 2. SSIM is an important index to measure the visual quality of images, and the larger the value of SSIM, the better the visual communication effect. In terms of the data

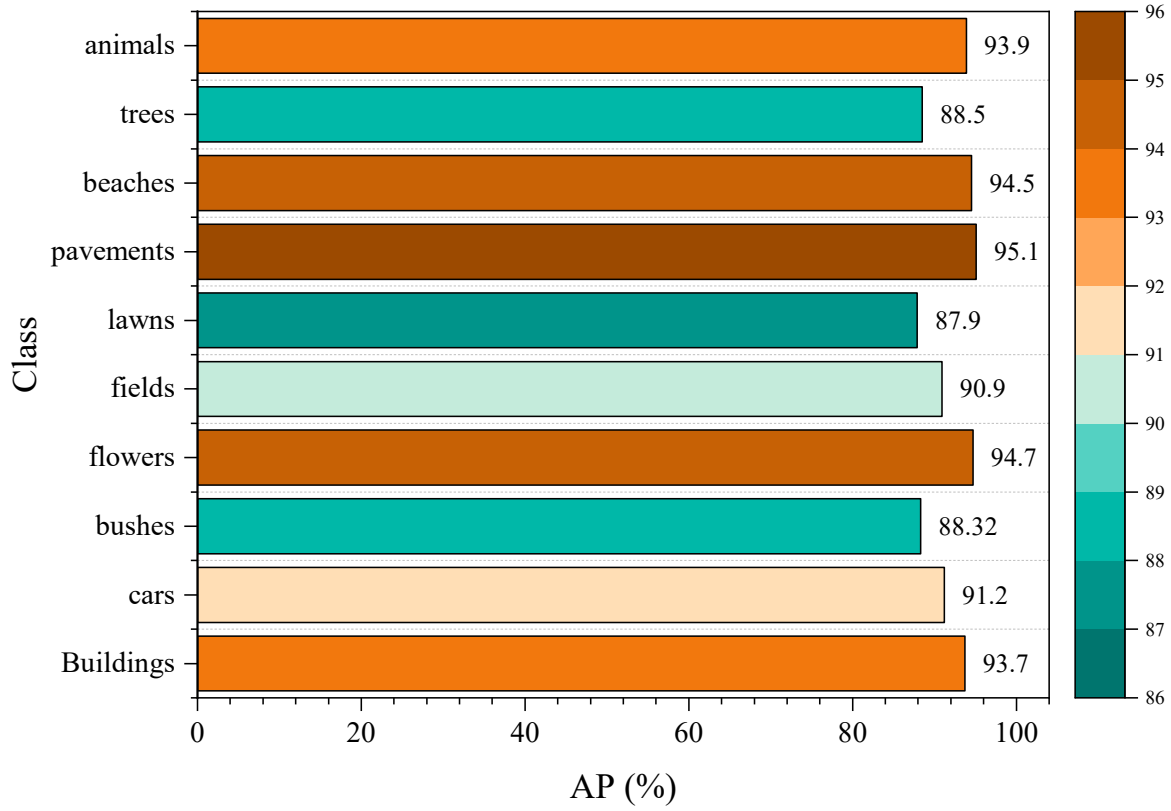

Figure 4. Classification accuracy of various types of images

in Table 1, it can be seen that the SIMM of VC-GCN is higher than that of the comparison method, which has a SIMM of 0.986 after classifying 100 images, closest to 1. At this time, the SSIM of the comparison method is only 0.914, 0.922 and 0.935. Compared with the four methods, the structural similarity of the proposed method is improved by more than 0.05, which fully demonstrates that the enhancement and classification of the high and low frequency components of images by using the attention mechanism and GCN can effectively optimize the visual communication of images.

Table 2. SIMM comparison of four methods

Number of images	VA-RNN	ML-GCN	SR-GCN	VC-GCN
20	0.913	0.926	0.919	0.979
40	0.921	0.935	0.943	0.981
60	0.916	0.918	0.951	0.992
80	0.913	0.921	0.939	0.987
100	0.914	0.922	0.935	0.986

The image PSNR test results of various methods are shown in Figure 5. As can be seen from Figure 5, compared with VA-RNN, ML-GCN and SR-GCN, the PSNR curve obtained by VC-GCN is higher than that obtained by the comparison method on the whole, with the highest value reaching 49.5dB, although the PSNR of SR-GCN also reaches 49.0dB. However, the PSNR optimized by this method is not stable enough and fluctuates greatly. The PSNR of VA- RNN is about 46.5dB, which is the lowest, but relatively stable. The PSNR of ML-GCN has a certain fluctuation, and the fluctuation range is between 45.4 dB-48.1dB. Therefore, it can be seen that the proposed method has

the highest peak signal-to-noise ratio and is relatively stable, effectively improving the PSNR of images.

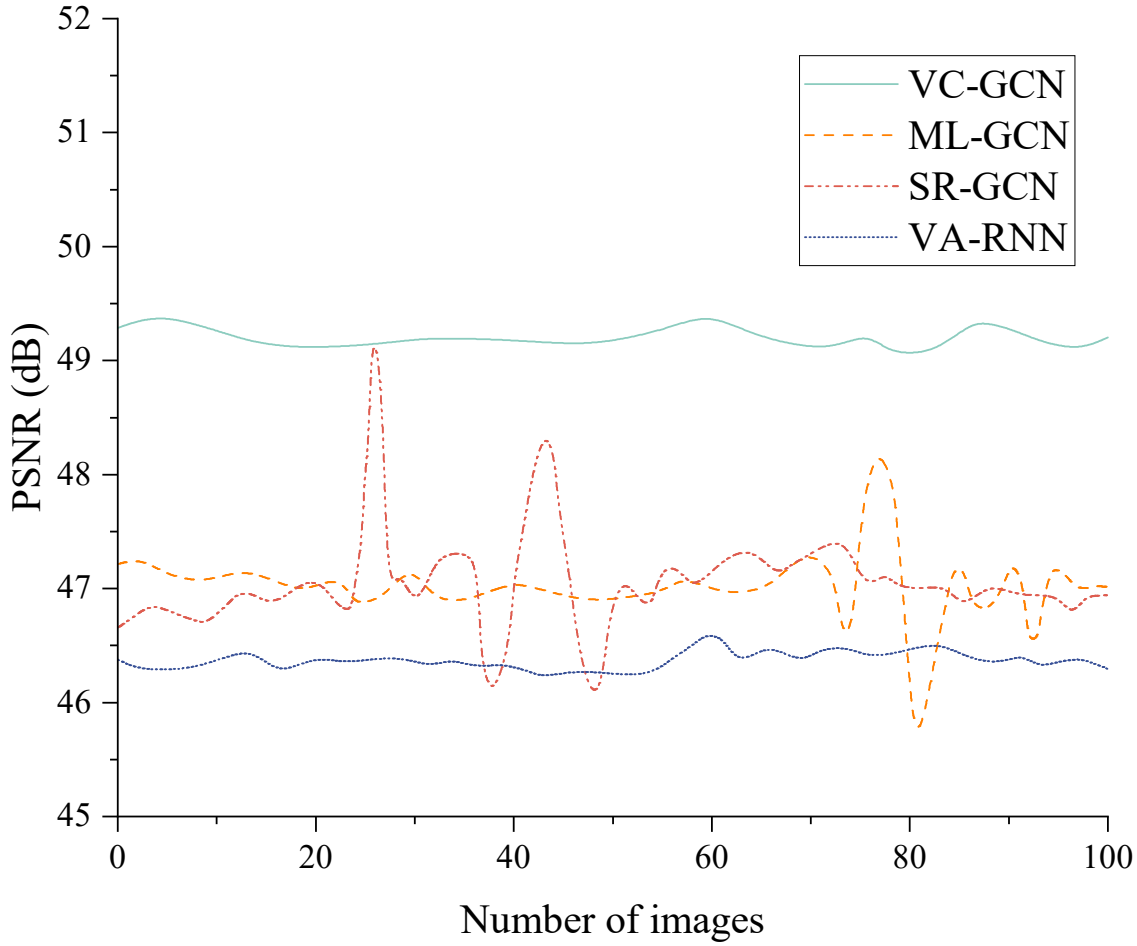

Figure 5. Comparison of PSNR for different visual communication methods

6. **Conclusion.** Current visual communication methods do not combine the characteristics of the image itself to establish the relationship between features and labels, resulting in poor visual communication, for this reason, this paper designs a visual communication method based on multi-label classification. Firstly, the Procreate software was used to convert the low-brightness image into lossless Lab color mode, and the Curvelet transform was introduced to decompose the converted image into high-frequency and low-frequency components, which describe the details and the overall brightness of the laser image, respectively. On this basis, the two components are augmented using the channel domain attention mechanism and the spatial attention mechanism, respectively, and spliced and fused to gain the hybrid domain attention feature map, and the dependencies between the feature map labels are modelled using GCN to establish a global visual optimization model of the image. The experimental outcome implies that the suggested method significantly improves the classification performance and visual communication efficiency of images, and has higher practical application value. In the subsequent optimization process, this paper will mainly focus on the expansion of the proposed method, in order to expand the application areas of the suggested method and reduce the dependence of the suggested method on the actual scene images.

Acknowledgment. This work is supported by the 2023 Zhuhai College of Science and Technology Teaching Quality Project "Design Basic Course Teaching and Research Department" (No. ZLGC20230504) and the "Zhuhai College of Science and Technology 'Three Levels' Talent Construction Project Funding" project.

REFERENCES

[1] U. Russmann, and J. Svensson, "Introduction to visual communication in the age of social media: Conceptual, theoretical and methodological challenges," Media and Communication, vol. 5, no. 4, pp. 1-5, 2017.
[2] H. Huo, and F. Wang, "A Study of Artificial Intelligence-Based Poster Layout Design in Visual Communication," Scientific Programming, vol. 2022, no. 1, 1191073, 2022.
[3] S. Xiang, "The application of big data and Artificial intelligence in visual communication design," Journal of Computational Methods in Sciences and Engineering, vol. 23, no. 2, pp. 639-649, 2023.
[4] Z. Li, Z. Wei, C. Wen, and J. Zheng, "Detail-enhanced multi-scale exposure fusion," IEEE Transactions on Image processing, vol. 26, no. 3, pp. 1243-1252, 2017.
[5] C. Jung, Q. Yang, T. Sun, Q. Fu, and H. Song, "Low light image enhancement with dual-tree complex wavelet transform," Journal of Visual Communication and Image Representation, vol. 42, pp. 28-36, 2017.
[6] J. Hu, D. Li, Q. Duan, Y. Han, G. Chen, and X. Si, "Fish species classification by color, texture and multi-class support vector machine using computer vision," Computers and Eectronics in Agriculture, vol. 88, pp. 133-140, 2012.
[7] Y.-Y. Zheng, J.-L. Kong, X.-B. Jin, X.-Y. Wang, T.-L. Su, and J.-L. Wang, "Probability fusion decision framework of multiple deep neural networks for fine-grained visual classification," IEEE Access, vol. 7, pp. 122740-122757, 2019.
[8] D. Streeb, Y. Metz, U. Schlegel, B. Schneider, M. El-Assady, H. Neth, M. Chen, and D. A. Keim, "Task-based visual interactive modeling: Decision trees and rule-based classifiers," IEEE Transactions on Visualization and Computer Graphics, vol. 28, no. 9, pp. 3307-3323, 2021.
[9] C. Nithyananda, and A. Ramachandra, "Adaptive Image Enhancement Using Image Properties and Clustering," International Journal of Image, Graphics and Signal Processing, vol. 8, no. 8, pp. 9, 2016.
[10] X. Liu, R. Zhang, Z. Meng, R. Hong, and G. Liu, "On fusing the latent deep CNN feature for image classification," World Wide Web, vol. 22, pp. 423-436, 2019.
[11] X. Zhang, L. Tang, H. Luo, S. Zhong, Z. Guan, L. Chen, C. Zhao, J. Peng, and J. Fan, "Hierarchical bilinear convolutional neural network for image classification," IET Computer Vision, vol. 15, no. 3, pp. 197-207, 2021.
[12] Y. Hui, J. Wang, Y. Shi, and B. Li, "Low light image enhancement algorithm based on detail prediction and attention mechanism," Entropy, vol. 24, no. 6, 815, 2022.
[13] J. Wu, V. S. Sheng, J. Zhang, H. Li, T. Dadakova, C. L. Swisher, Z. Cui, and P. Zhao, "Multi-label active learning algorithms for image classification: Overview and future promise," ACM Computing Surveys (CSUR), vol. 53, no. 2, pp. 1-35, 2020.
[14] L. Li, X. Zhu, Y. Hao, S. Wang, X. Gao, and Q. Huang, "A hierarchical CNN-RNN approach for visual emotion classification," ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), vol. 15, no. 3s, pp. 1-17, 2019.
[15] W. Zhou, Z. Xia, P. Dou, T. Su, and H. Hu, "Double attention based on graph attention network for image multi-label classification," ACM Transactions on Multimedia Computing, Communications and Applications, vol. 19, no. 1, pp. 1-23, 2023.
[16] F. Lyu, Q. Wu, F. Hu, Q. Wu, and M. Tan, "Attend and imagine: Multi-label image classification with visual attention and recurrent neural networks," IEEE Transactions on Multimedia, vol. 21, no. 8, pp. 1971-1981, 2019.
[17] M.-L. Zhang, and Z.-H. Zhou, "A review on multi-label learning algorithms," IEEE Transactions on Knowledge and Data Engineering, vol. 26, no. 8, pp. 1819-1837, 2013.
[18] Y. Cheng, M. Ma, X. Li, and Y. Zhou, "Multi-label classification of fundus images based on graph convolutional network," BMC Medical Informatics and Decision Making, vol. 21, pp. 1-9, 2021.
[19] X. Qu, H. Che, J. Huang, L. Xu, and X. Zheng, "Multi-layered semantic representation network for multi-label image classification," International Journal of Machine Learning and Cybernetics, vol. 14, no. 10, pp. 3427-3435, 2023.

[20] M. Agrawala, W. Li, and F. Berthouzoz, "Design principles for visual communication," Communications of the ACM, vol. 54, no. 4, pp. 60-69, 2011.
[21] M. Fan, and Y. Li, "The application of computer graphics processing in visual communication design," Journal of Intelligent & Fuzzy Systems, vol. 39, no. 4, pp. 5183-5191, 2020.
[22] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and S. Y. Philip, "A comprehensive survey on graph neural networks," IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 1, pp. 4-24, 2020.
[23] D. Lin, J. Lin, L. Zhao, Z. J. Wang, and Z. Chen, "Multilabel aerial image classification with a concept attention graph neural network," IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-12, 2021.
[24] E. A. Lee, "Is software the result of top-down intelligent design or evolution?," Communications of the ACM, vol. 61, no. 9, pp. 34-36, 2018.
[25] Y. Cheng, C. Jung, and K. Taeyong, "Projection-Mapping Art Using Opera Facial Makeup Patterns in Projection: Based on the Artwork "Mask"," TECHART: Journal of Arts and Imaging Science, vol. 8, no. 3, pp. 10-14, 2021.
[26] J. Ma, and G. Plonka, "The curvelet transform," IEEE Signal Processing Magazine, vol. 27, no. 2, pp. 118-133, 2010.
[27] J. M. Fadili, and J.-L. Starck, "Curvelets and ridgelets," Encyclopedia of Complexity and Systems Science, vol. 14, pp. 1718-1738, 2009.
[28] Y. Yu, Y. Zhang, Z. Song, and C.-K. Tang, "LMA: lightweight mixed-domain attention for efficient network design," Applied Intelligence, vol. 53, no. 11, pp. 13432-13451, 2023.
[29] K. Dong, and S. Zhang, "Deciphering spatial domains from spatially resolved transcriptomics with an adaptive graph attention auto-encoder," Nature Communications, vol. 13, no. 1, 1739, 2022.
[30] G. Sumbul, and B. Demir, "Generative reasoning integrated label noise robust deep image representation learning," IEEE Transactions on Image Processing, vol. 32, pp. 4529-4542, 2023.
[31] A. Horé, and D. Ziou, "Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure?," IET Image Processing, vol. 7, no. 1, pp. 12-24, 2013.