

Intelligent Layout of Graphic Elements Based on Gabor Wavelet and Deep Neural Network

Lin Ma*

School of Science and Technology
Beijing Open University
Beijing 100081, P.R. China
colin4119@163.com

Cawlton Evans

Faculty of Engineering
Dalhousie University
Halifax B3H 4R2, Canada
ef2528@163.com

*Corresponding author: Lin Ma

Received May 26, 2023, revised July 30, 2023, accepted September 04, 2023.

ABSTRACT. *Personalized photo collages and personalized puzzle posters often appear in various social apps. This personalized visual communication effect has been loved by the public. However, the existing image layout technology has the following problems: the appearance shape is relatively simple and the content layout is too simple. In order to solve the above two problems, this paper designs an artificial intelligence layout method based on improved Gabor wavelet and deep learning. Firstly, the input planar graphic elements are preprocessed by normalizing graphics and defining graphic blocks respectively. Then, an improved Gabor wavelet feature extraction algorithm is proposed to effectively reduce the feature dimension in multi-scale analysis. Secondly, Lloyd algorithm is used to divide all kinds of complex contour shapes so that graphic elements are evenly distributed in the whole area. Finally, according to the attention picture selected by the user, the reduced-dimension matrix is identified and translated by using the depth neural network, so as to obtain the matching result between the set of planar graphic elements and the spatial position. The experimental results verify the effectiveness of the proposed intelligent layout method. Compared with the random layout algorithm, the proposed intelligent layout method can gather similar graphic elements together and place the user's attention image in the center of the outline.*

Keywords: Artificial intelligence; Jigsaw posters; Personalisation; Gabor wavelets; Deep neural networks; Visual communication

1. Introduction. There are many scenarios in visual communication where intelligent image layout is required. For example, many people like to share their personalised photo collages on social apps [1,2]. Travel companies often create jigsaw posters for promotional purposes [3,4]. Advertising images that use a personalised layout will gain more attention. People have a basic demand, that is, how to quickly generate personalized layout pictures. Therefore, the study of intelligent image layout technology is worthy of attention and has some significance [5,6]. Intelligent image layout technology refers to the use of a collection of flat graphic elements as input to generate personalised image layouts in different styles using image layout related technology in the computer.

A typical application of artificial intelligence image layout technology is photo collage [7,8]. Photo collage refers to a series of scaling and rotation of multiple photos according to the layout of a preset template, thus fusing multiple images into one information-rich and well-laid out image. In addition to photo collage [9,10], typical applications of intelligent image layout technology include personalised jigsaw posters and advertising image layouts. The input to all three typical applications can be considered as flat graphic elements. Intelligent image layout technology can organise flat graphic elements in a reasonable way and get good visual communication effects [11,12]. The main process of intelligent image layout is divided into two steps, which are (1) selecting the layout in appearance (choosing a template for the layout) and (2) selecting the placement position according to the content of the graphic elements. The main technical processes of intelligent image layout include layout in appearance, selection of templates, selection of images and layout in content [13,14]. Of these, the selection of templates and the selection of images are done interactively by the user. Therefore, appearance layout and content layout are the key technologies and difficulties of intelligent image layout.

The appearance layout classification includes [15,16]: stitching layout, regular layout and free layout, as shown in Figure 1. The stitching layout is a layout effect where images are joined from top to bottom to form one long image. The stitching layout is characterised by its simplicity and ease of use. A regular layout is a spatial division of the canvas into a more regular grid. Regular layouts are characterised by their neat and compact proportions. Compared to stitching and regular layout [17], free layout is a very special form of layout, which has a random and free character and gives a new and beautiful visual communication effect. Free layout is currently the most common method of personalising photo collages. Therefore, the intelligent image layout technique studied in this paper uses free layout. Content layout is a technique for distributing

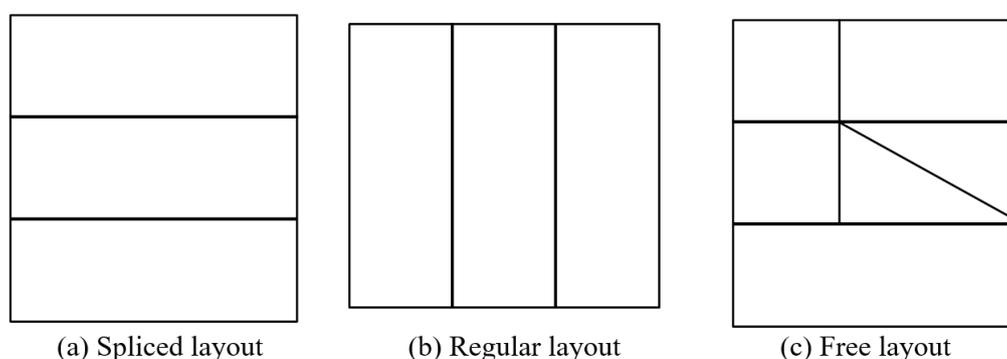


Figure 1. Classification of exterior layouts

images in different spatial positions according to their content on the basis of the exterior layout. Content layout is mainly concerned with the mapping relationship between image content, image position and the matching relationship between image positions. There are two types of content layouts available [18,19]: tiled layouts and random layouts. Tiled layouts are the simplest form of content layout. A tiled layout distributes the images in the order they are entered, from left to right or top to bottom, across the appearance of the layout. A random layout maps the set of images to different positions in a random way [20]. The characteristic of random layout is that even if the same set of images is entered, each layout may produce a different layout result. There are a number of photo collage apps available, such as Meituxiu [21], Baidu Magic Image app, and Collagelt Pro [22]. However, existing image layout techniques suffer from the following problems.

(1) The appearance shape is relatively single. The free layouts in existing photo collage software are less innovative, although they give a sense of freedom and discretion. This is because free layouts do not actively adapt to different contour shapes. Therefore, the existing layout techniques do not meet the individual requirements of the user.

(2) Overly simple content layout. Traditional content layout techniques only use simple tiling and randomisation, and therefore do not allow images to be laid out appropriately based on similarities between content. Traditional content layout techniques do not allow images that are of interest to users to appear in the most prominent positions, resulting in a low level of personalisation.

Therefore, the purpose of this study is to design an intelligent image layout technique that conforms to the visual communication effect, so as to solve the user personalised display problem in two aspects such as variable appearance and content relevance.

1.1. Related Work. With the introduction of deep learning theory and the improvement of computing performance of hardware devices, Deep Neural Network (DNN) technology has been very successful in the fields of computer vision and natural language processing with high recognition rates [23,24]. Therefore, this paper attempts to introduce DNNs into intelligent image layout in order to achieve personalised image layout requirements. In addition, the input of DNN is the features of planar graphic elements [25,26]. Therefore, the feature extraction of planar graphic elements is very critical. With the development of wavelet analysis theory, Harr wavelets [27], Spline wavelets [28] and Gabor wavelets [29,30] have also been widely used. The method based on Gabor wavelet extracted features has obvious frequency characteristics and direction selectivity, can capture edge-sensitive local structure information, and has better robustness to illumination changes and contrast changes. Therefore, an intelligent image layout technique based on Gabor wavelet feature extraction and deep neural networks is proposed in this paper. The experimental results verify the effectiveness of the proposed algorithm.

There are two types of image layout methods: manual and automatic. Manual image layout is undoubtedly a time-consuming and labour-intensive method, and the collage effect varies from person to person [31], resulting in a very unstable visual communication effect. Sweeney [32] used a reasonably chosen layout criterion to define the energy function and minimised the energy-linked function by subjecting the image to a series of transformations (scaling, rotation) to achieve the image collage. Young [33] used the clustering centre of the image as the centre of the layout and extracted the important regions from the image. Then, the images are sorted to obtain the collage. However, the above two methods may produce the problem of loss of key information of the image when the image is cropped.

Content image layouts allow for the logical arrangement of a given collection of flat graphic elements on a canvas to create content-related visual communication effects. Common content image layouts generally need to consider characteristics such as the importance of the images, similarity between images, and therefore need to extract important features of the images, such as the colour and texture of the images. Gan et al. [34] display important images in a larger area by defining the importance of the images. In this case, the importance is calculated based on the complexity and variance of the images. Ultimately, image layouts related to the importance of the images were generated. Yang et al. [35] learned the importance of the images by means of user interaction. In conclusion, content layout is a deeper form of layout on top of its foundation as opposed to spatial layout.

1.2. Motivation and contribution. As a more advanced machine learning method, deep learning models have shown very good performance in the field of image processing,

with powerful classification and identification capabilities. Therefore, this paper attempts to introduce DNN into artificial intelligence image layout in order to achieve personalised image layout requirements. In addition, feature extraction of flat graphic elements is crucial. Therefore, the Gabor wavelet technique is used to extract the features of the planar graphic elements (the input to the DNN).

The main innovations and contributions of this paper include:

(1) For the first time, DNN is applied to intelligent image layout, thus allowing personalised layout of a given collection of flat graphic elements based on similarity relationships between canvas contours of different shapes and picture content.

(2) The performance of DNN is very dependent on the input of the model, which determines the similarity between planar graphic elements by the difference of the input image features. In order to improve the accuracy of similarity judgment, this paper proposes an improved Gabor wavelet feature extraction method, which makes the Gabor filtered image with more detailed texture information.

2. Pre-processing of flat graphic elements.

2.1. Scale normalisation. Intelligent image layout techniques are divided into processes such as image pre-processing, selection of graphic blocks and arrangement of graphic blocks. The pre-processing of flat graphic elements is an essential operation in the intelligent layout process, but the input flat graphic elements have various problems, such as shadows, highlights and low contrast. Therefore, extraneous factors such as lighting and background need to be eliminated in order to provide the best conditions for subsequent feature extraction.

The input target planar graphic elements are first scale normalised in order to ensure that all samples have the same size. Calculate the covariance matrix \mathbf{M} of the image.

$$\mathbf{M} = \begin{bmatrix} \mu_{20} & \mu_{11} \\ \mu_{11} & \mu_{02} \end{bmatrix} \quad (1)$$

where μ represents the central moments. The coordinate system is rotated using the eigenvectors of \mathbf{M} . Let the eigenvalues of \mathbf{M} be λ_1 and λ_2 , then the eigenvectors are calculated as follow.

$$\mathbf{e}_i = \begin{bmatrix} e_{ix} \\ e_{iy} \end{bmatrix} = \begin{bmatrix} \frac{\mu_{11}}{\sqrt{(\lambda_i - \mu_{20})^2 + \mu_{11}^2}} \\ \frac{\lambda_i - \mu_{20}}{\sqrt{(\lambda_i - \mu_{20})^2 + \mu_{11}^2}} \end{bmatrix}, i = 1, 2 \quad (2)$$

where $\mathbf{e}_1 = [e_{1x}, e_{1y}]^T$ and $\mathbf{e}_2 = [e_{2x}, e_{2y}]^T$ denote the eigenvectors corresponding to λ_1 and λ_2 , respectively. \mathbf{E} denotes the resulting rotation matrix.

$$\mathbf{E} = \begin{bmatrix} e_{1x} & e_{1y} \\ e_{2x} & e_{2y} \end{bmatrix} \quad (3)$$

The eigenvectors are assumed to be orthogonal to each other ($e_{1x}e_{1y} + e_{2x}e_{2y} = 0$), thus ensuring that \mathbf{M} is a symmetric matrix. The coordinates of the rotated transformed image are shown as follow.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \mathbf{E} \begin{bmatrix} x - \bar{x} \\ y - \bar{y} \end{bmatrix} \quad (4)$$

where (\bar{x}, \bar{y}) represents the centre of gravity coordinates of the image. The coordinate system is normalised according to the eigenvalues of \mathbf{M} .

$$\mathbf{W} = \begin{bmatrix} \frac{(\lambda_1 \cdot \lambda_2)^{1/4}}{\sqrt{\lambda_1}} & 0 \\ 0 & \frac{(\lambda_1 \cdot \lambda_2)^{1/4}}{\sqrt{\lambda_2}} \end{bmatrix}, \frac{(\lambda_1 \cdot \lambda_2)^{1/4}}{\sqrt{\lambda_1}} \cdot \frac{(\lambda_1 \cdot \lambda_2)^{1/4}}{\sqrt{\lambda_2}} = 1 \quad (5)$$

The final result is a normalised image coordinate that is independent of the coordinates.

$$\begin{bmatrix} x'' \\ y'' \end{bmatrix} = \mathbf{WE} \begin{bmatrix} x - \bar{x} \\ y - \bar{y} \end{bmatrix} \tag{6}$$

2.2. Grayscale normalization. The histogram equalisation is then applied to the normalised image to achieve grey-scale normalisation. Grey-scale normalisation enhances the contrast between images, increases the dynamic range of the image and improves the quality of the image. Assuming that the sum of pixels in a digital image is and the total number of grey levels is L , calculate the probability of occurrence of pixels of different grey levels in the image.

$$P_r(r_k) = \frac{n_k}{N}, 0 \leq r_k \leq 1, k = 0, 1, 2, \dots, L - 1 \tag{7}$$

$$s = EN(r), 0 \leq r \leq 1 \tag{8}$$

where s is the individual grey scale of the input image and r is the individual grey scale of the transformed image. $EN(r)$ is monotonically increasing in the value interval, so the equalised transformation function [36] is shown as follow.

$$s_k = T(r_k) = \sum_{i=0}^k \frac{n_j}{N} = \sum_{i=0}^k p_r(r_j), 0 \leq r_k \leq 1, k = 0, 1, 2, \dots, L - 1 \tag{9}$$

A comparison of the histograms before and after equalisation is shown in Figure 2.

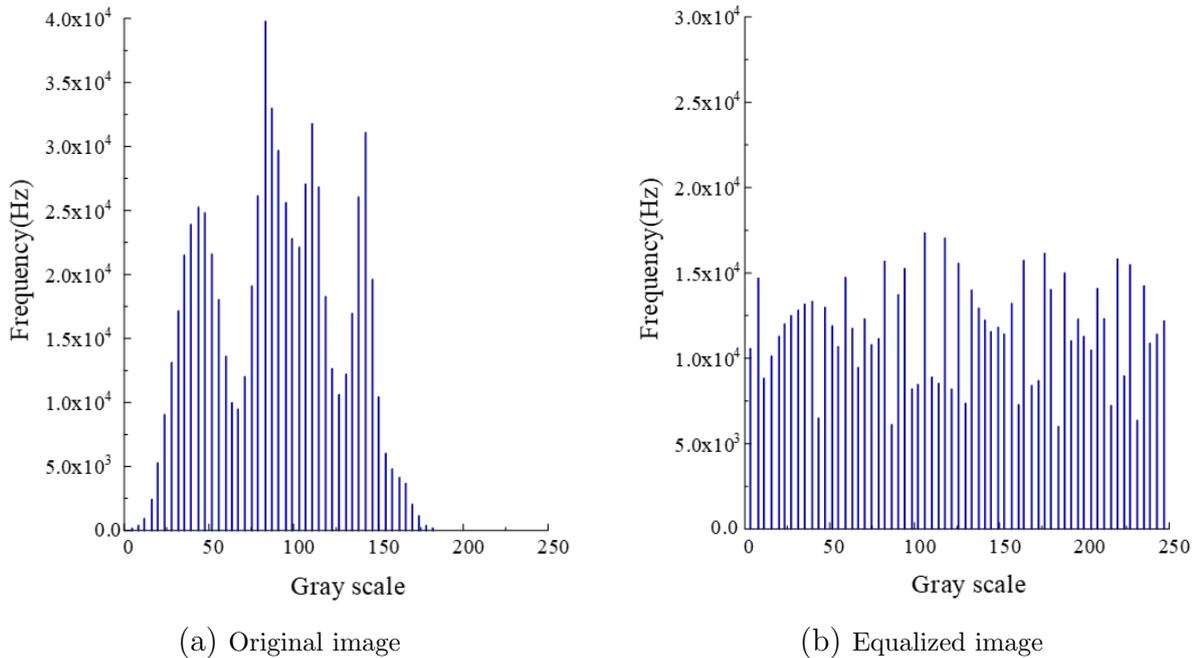


Figure 2. Histogram comparison before and after equalisation

2.3. Graphic block definition. In order to meet the requirement that important areas are not covered, the important areas of the image need to be obtained first. After the image has been cropped, only the content of the main areas of the image is retained. In this paper, the cropped image is seen as graphical blocks of different shapes. The set of these blocks will be described to suit the requirements of the image layout in terms of appearance: (1) the shape of the block and the shape of the cropped image correspond; (2) the centre of gravity of the block corresponds to the centre of the cropped image and the coordinates of the vertices of the block correspond to the coordinates of the vertices of

the cropped image; (3) the size of the block corresponds to the size of the cropped image; and (4) the orientation of the block needs to be appropriate for the user to view.

The geometric characteristics of the graphic block must be consistent with the cropped image, so that the shape of the graphic block is a collection of rectangles with many different aspect ratios. In addition, the image must not appear to have its contents reversed. Therefore, the orientation of the blocks must be restricted. The angle at which the blocks are placed must match the user's viewing habits in order to satisfy the visual communication effect. Therefore, based on experience, the angle is controlled to a range of 45 degrees from left to right, as shown in Figure 3.

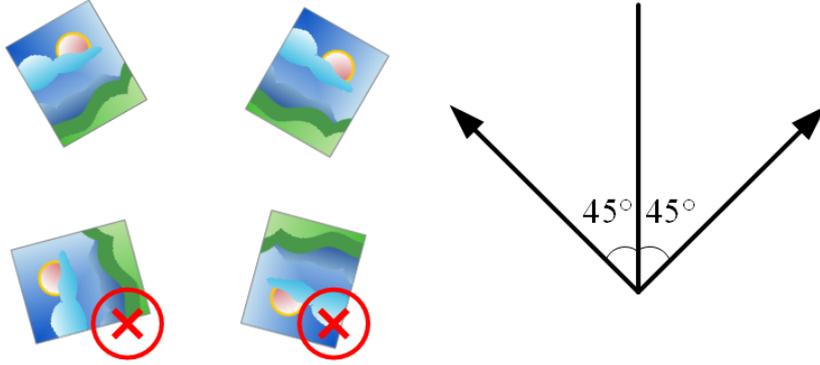


Figure 3. Graphic block orientation control

3. Feature extraction of planar graphic elements.

3.1. Question formulation. The objective of Gabor wavelet feature extraction is the analysis of texture properties. In order to analyse texture properties, we need to analyse the image in one dimension, both in terms of frequency and direction.

$$S(r) = \sum_{\theta=0}^{\pi} S_{\theta}(r) \quad (10)$$

$$S(\theta) = \sum_{r=0}^R S_r(\theta) \quad (11)$$

Where r is the frequency, θ is the direction, $S(r, \theta)$ is the 2D spectral function and R is the radius of the circle (circle centred on the origin).

A widely used method at this stage is to extract the texture spectrum energy information from the image by means of a multi-channel filter.

3.2. Composition of a two-dimensional Gabor wavelet filter. The two-dimensional Gabor wavelet is a typical method for multi-resolution analysis of images. Two-dimensional Gabor wavelets are able to achieve optimal resolution in both the null and frequency domains. The definition of the two-dimensional Gabor wavelet kernel function is shown below.

$$\psi_{u,v}(z) = \frac{\|k_{u,v}\|^2}{\sigma^2} \exp\left(-\frac{(k_{u,v} \cdot z)^2}{2\sigma^2}\right) \cdot \left[\exp(ik_{u,v}z) - \exp\left(-\frac{\sigma^2}{2}\right)\right] \quad (12)$$

where $k_{u,v}$ is the centre frequency of the filter, u is the direction factor, σ is a constant related to the frequency bandwidth of the wavelet (σ is usually set to 2π), v is the scale factor and z are the coordinates of the given position.

It can be seen that the bandwidth of the Gabor filter determines the ratio of the Gaussian window width to the wavelength.

$$\sigma = \sqrt{2 \ln 2} \left(\frac{2^\varphi + 1}{2^\varphi - 1} \right) \quad (13)$$

where φ is the bandwidth factor.

The Gabor wavelet filter can be divided into a real part Re and an imaginary part Im .

$$\psi_{u,v}(z) = \text{Re}(\psi_{u,v}(z)) + i \text{Im}(\psi_{u,v}(z)) \quad (14)$$

$$\text{Re}(\psi_{u,v}(z)) = \frac{\|k_{u,v}\|^2}{\sigma^2} \cdot \exp\left(-\frac{(k_{u,v} \cdot z)^2}{2\sigma^2}\right) \cdot \left[\cos(k_{u,v} \cdot z) - \exp\left(-\frac{\sigma^2}{2}\right) \right] \quad (15)$$

$$\text{Im}(\psi_{u,v}(z)) = \frac{\|k_{u,v}\|^2}{\sigma^2} \cdot \exp\left(-\frac{(k_{u,v} \cdot z)^2}{2\sigma^2}\right) \cdot [\sin(k_{u,v} \cdot z)] \quad (16)$$

The amplitude of the Gabor wavelet filter is $|G_{u,v}(z)|$.

$$|G_{u,v}(z)| = \sqrt{\text{Re}(\psi_{u,v}(z))^2 + \text{Im}(\psi_{u,v}(z))^2} \quad (17)$$

The centre frequency of the filter $k_{u,v}$ controls the wavelength and direction of the oscillating section as well as the width of the Gaussian window.

$$k_{u,v} = k_v (\cos \theta_u, \sin \theta_u)^T \quad (18)$$

$$k_v = \frac{k_{\max}}{f^v} \quad (19)$$

where k_{\max} is the maximum centre frequency of the Gabor filter (usually taken as $\pi/2$), f^v is the spatial factor of the filter in the frequency domain and θ_u is the directional selectivity of the Gabor filter.

3.3. Improved Gabor wavelet feature extraction algorithm. The parameters k_v and θ_u have a large influence on the performance of the 2D Gabor wavelet. By choosing different parameters, we can obtain multiple different Gabor filters. This means that we can generate multiple filter banks by scaling and rotating. However, when too many scales and orientations are selected, this results in a large amount of redundancy and leads to a dramatic increase in complexity. For example, on the sampling interval $\theta_u = [0, \pi]$, we can select 5 scales and 8 orientations to generate a Gabor filter bank.

$$\psi_{u,v}(z) \quad u = 0, 1, \dots, 7; v = 0, 1, \dots, 4 \quad (20)$$

The parameter k_v and the parameter θ_u are calculated as follow.

$$k_v = \frac{k_{\max}}{f^v} = 2^{-\frac{v+2}{2}} \pi, \theta_u = \frac{u\pi}{8} \quad (21)$$

Each Gabor filter in a Gabor filter bank can reflect local features in different frequency ranges, at different scales and orientations.

After the Gabor wavelet transform, planar graphical elements on the same scale in different directions contain more redundant information, leading to a rapid increase in the number of feature dimensions, creating a very serious 'data catastrophe'. This problem increases the computational complexity and also affects the recognition rate of subsequent DNN models. Therefore, this paper proposes to encode and fuse the Gabor wavelet feature amplitudes of different orientations at the same scale. The original Gabor features of the

planar graphic elements are assumed to be represented as $\{G_{u,v}(z)|u = 0, 1, \dots, 7, v = 0, 1, \dots, 4\}$, and the encoded fused image is represented as $\{R_v(z)|v = 0, 1, \dots, 4\}$. First, the mean value of the Gabor eigenmagnitude was calculated for eight different directions at the same scale avg_v .

$$avg_v = \frac{(G_{0,v} + G_{1,v} + \dots + G_{7,v})}{8} \quad (22)$$

Each Gabor feature amplitude is then binary transformed using the mean value avg_v as the threshold. Each binary coded value is also given a weight 2^p , resulting in a decimal coded value $R_v(z) \in [0, 255]$ representing the fused features.

$$s(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (23)$$

$$R_v(z) = \sum_{p=0}^7 s(G_{p,v} - avg_v) \cdot 2^p \quad (24)$$

Finally, the Local binary pattern (LBP) method was used for mapping. Combined with histogram statistics in chunks, this effectively reduces the number of feature dimensions when performing multi-scale analysis without reducing the texture feature information.

4. Intelligent layout method based on improved Gabor wavelets and DNN..

4.1. Contour division. There is a problem of oversimplification in existing personalised photo collages. The existing method is simply to lay out the graphic elements in a tiled or random way according to the input order of the graphic elements. However, if similar photos could be laid out together, it would be easier for users to navigate and find photos. At the same time, personalised puzzle posters need to have graphic elements with similar colours and textures laid out together. The layout of personalised advertising images requires that the images that are of most interest to the user are placed most prominently in the contour. Therefore, in order to meet the needs of different personalised scenarios in terms of content layout, this paper attempts to generate a personalised image layout style by putting similar images together and placing the graphic elements of interest to the user at the most prominent position in the canvas contour.

In order to evenly distribute the planar graphic elements throughout the area, the user-selected canvas contour needs to be evenly divided. Note that the number of meshes to be divided must be the same as the number of image sets entered by the user. The division of the contours in this paper consists of two main parts. The first part is the extraction of the canvas contours. The second part is the specific contour division algorithm. In this paper, Lloyd algorithm [37] is used to divide the contour area. The whole method flow of division is shown in Figure 4.

In this paper, the Lloyd algorithm is used to partition the extracted contour shapes. First, a corresponding number of scatter points are randomly generated inside the canvas contour shape based on the number of elements in the set of planar graphic elements. A Voronoi diagram is calculated and generated based on the scatter points inside the contour shape. In this paper, the Bowyer-Watson algorithm [38] is used to compute the Voronoi diagram. The Bowyer-Watson algorithm is characterised by its simplicity, small memory footprint and short computing time. the Bowyer-Watson algorithm has been widely adopted in the field of image contour extraction. Before the Bowyer-Watson algorithm can be executed, the contour shapes need to be segmented in the background image. After the histogram equalisation, the contour shape is segmented from the background image using the background difference method.

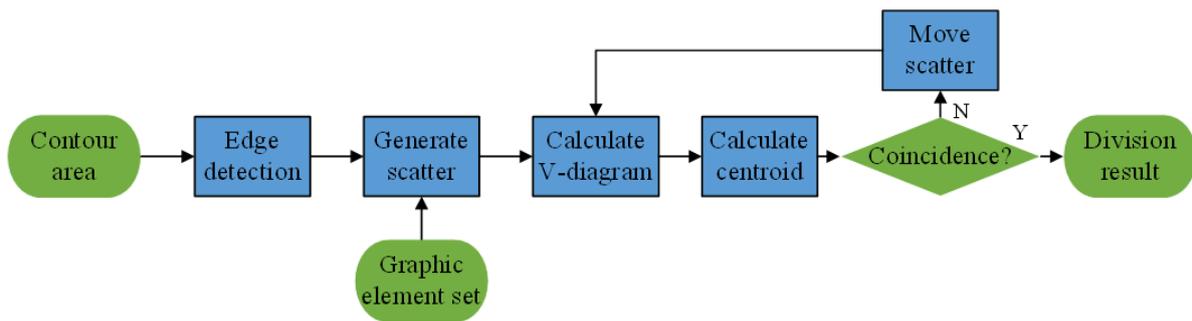


Figure 4. Flow of contour division

Then, determine if the centre of mass and the scatter point overlap. If the centre of mass and the scatter point are coincident, then the division is over. Otherwise, the scatter is moved to the centre of mass and the division of the contour shape continues. If the scatter and centre of mass coincide, or if the number of iterations reaches a certain upper limit, then the process ends. The final result of the Voronoi diagram is obtained by the Lloyd algorithm. The results of the contour division are shown in Figure 5.

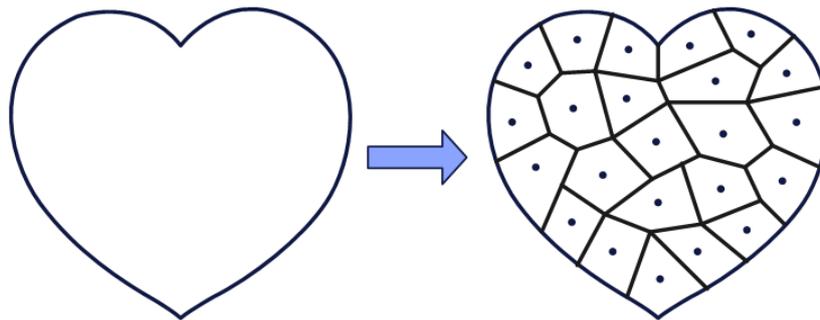


Figure 5. Effect of contour division

4.2. User interaction. Within the input collection of flat graphic elements, the user selects the graphic element of their interest through an interactive interface. In this paper, the position of the center of gravity of the contour is taken as the most significant position. Therefore, the concerned graphic element should be placed at the point of spatial location closest to the centre of gravity of the contour. In addition, other graphic elements similar to the concerned graphic element should be placed around the vicinity of the centre of gravity of the contour, as shown in Figure 6.

4.3. Matrix dimensionality reduction. After feature extraction of all graphic elements using a modified Gabor wavelet feature extraction algorithm, the feature distances between all graphic elements need to be calculated and a feature matrix constructed in order to be used as input to the DNN. The distance between each coordinate point reflects the feature distance between images. However, the feature distance matrix is high-dimensional data, resulting in images that cannot be mapped to two-dimensional coordinate points. Therefore, the feature distance matrix needs to be mapped to a two-dimensional space using a dimensionality reduction method.

Data dimensionality reduction is a process of mapping data from a high-dimensional space to a low-dimensional space by means of a transformation mapping method. There

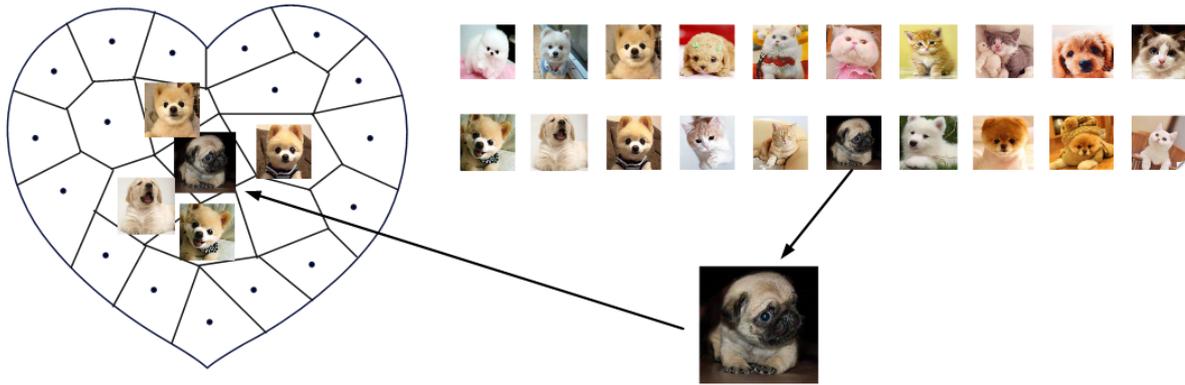


Figure 6. The layout of the concerned graphic elements

are many methods of dimensionality reduction, which can be classified as linear or non-linear according to different classification bases. Linear dimensionality reduction ensures linear relationships between datasets [39], for example methods such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). Although linear methods have the advantage of being fast and simple. However, for more complex data, the results are not satisfactory. Therefore, this paper adopts the typical IsoMap algorithm in nonlinear dimension reduction. The main idea of IsoMap algorithm is to use geodesic distance instead of Euclidean distance. IsoMap algorithm can discover the coordinate information embedded in the high-dimensional space. IsoMap algorithm can use the distance matrix as the input data set, and can keep the distance between images before and after dimensionality reduction will not. The IsoMap algorithm can use the distance matrix as the input dataset and keep the distance between images before and after dimensionality reduction unchanged.

4.4. DNN construction and training. In this paper, the first graphical element of interest to the user is first selected by means of user mutuality. Then, a modified Gabor wavelet feature extraction algorithm is used to extract features from all graphic elements. The feature distances between all graphical elements are calculated and a feature matrix is constructed.

The IsoMap algorithm is used in this paper to reduce the feature matrix to a two-dimensional space. Finally, DNN is used to identify and translate the dimension-reduced matrix. Specifically, bi-directional matching between coordinate points is performed by means of a customised objective function. By making one-to-one correspondence between graphic elements and spatial locations, the final image layout result is obtained. The objective function serves to ensure that the feature matrix is in the same scale range as the spatially located coordinate points that have been obtained. At the same time, the two-dimensional coordinate points are translated so that the coordinate points corresponding to the concerned graphical elements coincide with the most significantly located spatial coordinate points. The flow of the proposed intelligent layout method is shown in Figure 7. The DNN structure constructed in this paper is mainly composed of an input layer, a convolutional layer, an activation function, a pooling layer, a fully connected layer and an output layer, as shown in Figure 8. Let the input 2D feature matrix be $\mathbf{I}_{(i,j)}$. The 2D feature matrix $\mathbf{I}_{(i,j)}$ can be considered as an image. Therefore, we extract a block of images of size $N_{patch} \times N_{patch}$ with pixel point (i, j) as the centre. The trainable parameters in the DNN are represented as $(\mathbf{W}_{cov}^k, b_{cov}^k)$.

$$\mathbf{X}^k = pool(f(\mathbf{W}_{cov}^k * \mathbf{X}^{k-1} + b_{cov}^k)) \quad (25)$$

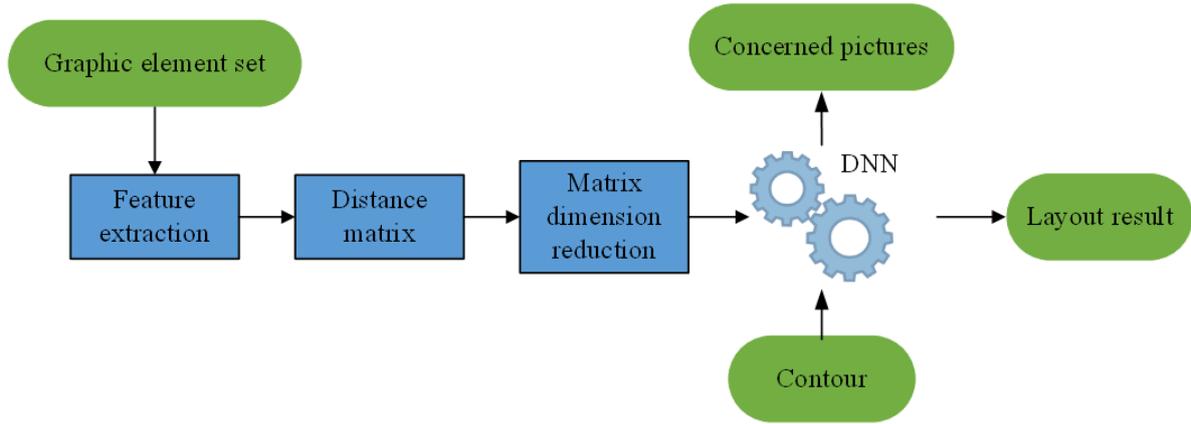


Figure 7. Flow of the intelligent layout method

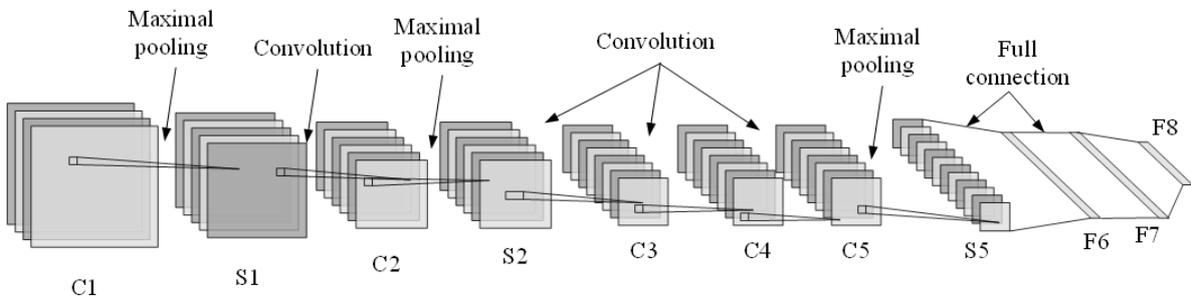


Figure 8. DNN structure

where $\mathbf{X}_{(i,j)}$ is the original input to the DNN, k indicates the number of layers in the DNN, \mathbf{W}_{cov}^k indicates the weight matrix trained in the convolutional layer k , f indicates the excitation function, $pool$ indicates the pooling function, and b_{cov}^k indicates the trainable bias in the convolutional layer k .

The DNN is trained using the BP algorithm and therefore the partial derivatives of the weights in the output layer L need to be calculated.

$$\frac{\partial L}{\partial W_L} = -(e(y) - f(x)) \cdot f'(x) \tag{26}$$

where $e(y)$ is the category label and $f'(\cdot)$ is the derivative of the excitation function.

Finally, the partial derivative of the bias in the output layer L is calculated.

$$\frac{\partial L}{\partial b_L} = -(e(y) - f(x)) \tag{27}$$

5. Experimental results and analysis.

5.1. Experimental design and parameter setting. In this paper, four experiments were designed to verify the effectiveness and adaptability of the proposed intelligent layout method from both qualitative and quantitative perspectives. Experiment 1: qualitatively verifies whether the proposed method can arrange flat graphic elements with similar content together, Experiment 2: qualitatively verifies whether the proposed method can place images of interest to users in a prominent position in the contour. Experiment 3: quantitatively verifies whether the proposed approach facilitates the user’s navigation of the

images. Experiment 4: quantitative verifies whether the proposed method improves user satisfaction with visual representations.

The experimental hardware environment is: Windows 7 operating system, Intel(R) Core(TM) i5 CPU, 4GB RAM, 500G hard disk. The experimental software environment is: MATLAB R2016a. The size of the input image is 480*560. The size of the graphics block N_{patch} is set to 5 and the number of nodes of the convolution layer n is set to 10.

5.2. **Experiment I.** The final results of the random layout and the proposed method were compared for the same planar graphic element input and contour, respectively, as shown in Figure 9. It can be seen that there is no connection between similar planar

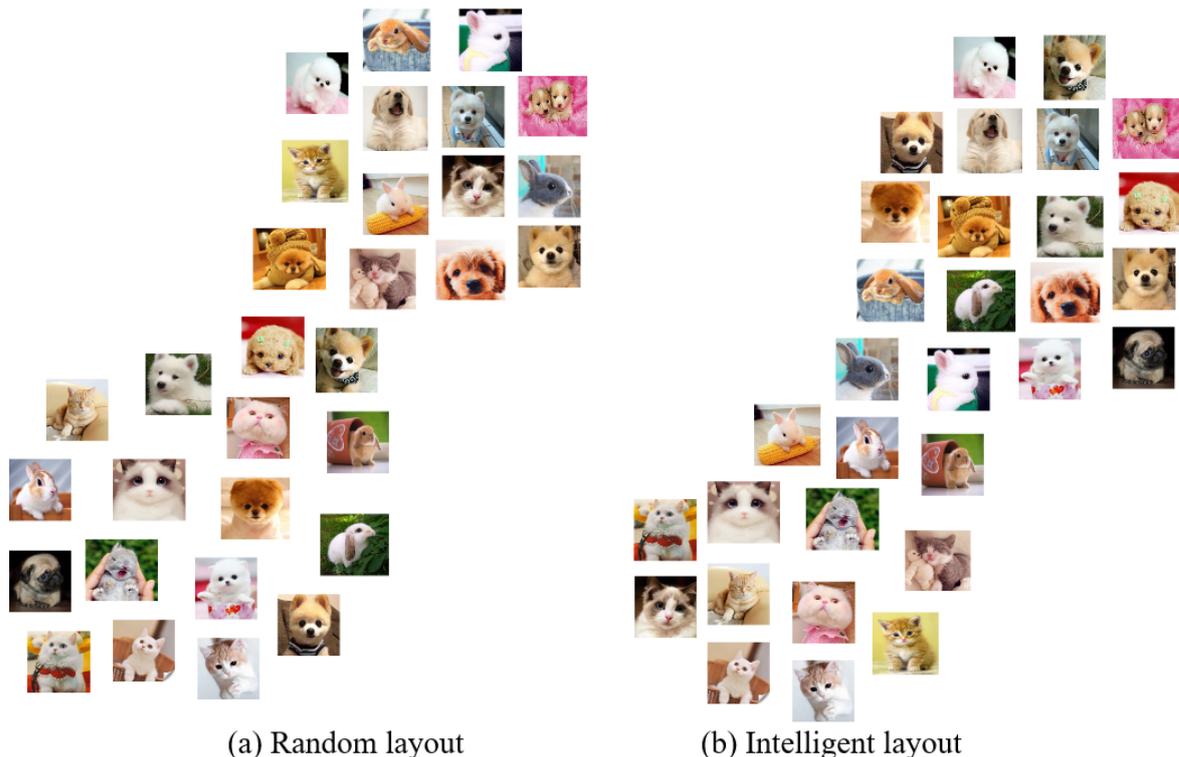


Figure 9. Comparison of similarity of layouts

graphic elements in the layout results obtained by the random layout method. All graphic elements are randomly distributed in the contour of the canvas. However, with the intelligent layout method, similar images are basically arranged together, thus enhancing the regularity of the overall visual effect. The experimental results show that the intelligent layout method can, to a certain extent, ensure that images with similar content are clustered together.

5.3. **Experiment II.** The final results of the random layout and the proposed method were compared for the same planar graphic element input and contour, respectively, as shown in Figure 10. As can be seen, after the user has selected the concerned image through the interactive interface, the concerned graphical element is placed in the centre of the contour in the layout effect obtained using the intelligent layout method. In addition, other images with similar content to the concerned graphic element are arranged nearby. However, the user's concern is not reflected in the layout effect obtained by the random arrangement method. All graphic elements are randomly arranged in the canvas contour. The comparison shows that the intelligent layout method is able to place the concerned

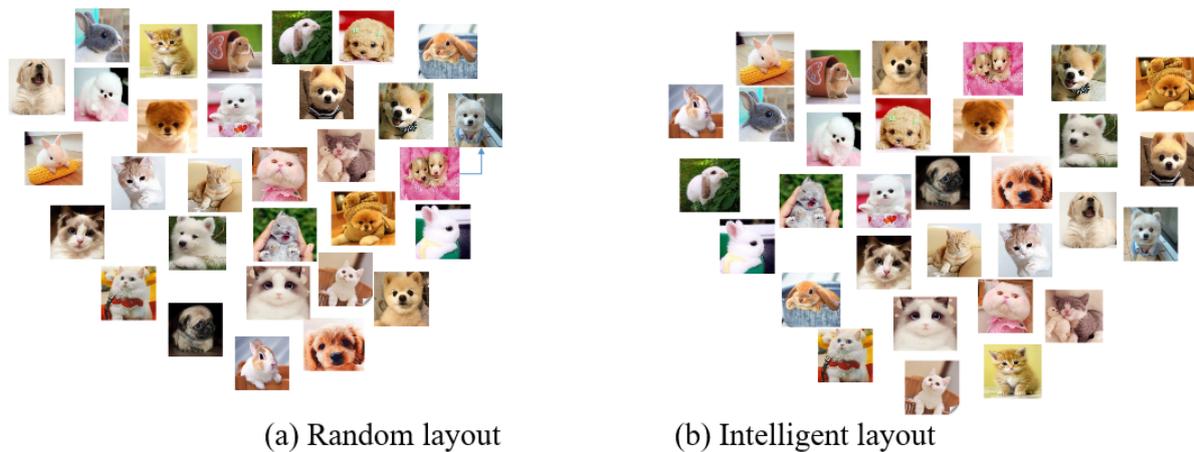


Figure 10. Comparison of the layout of the concerned graphic elements

graphical elements in exactly the most prominent position in the canvas contour by means of user interaction.

5.4. **Experiment III.** The final results of the random layout and the proposed method were compared for the same flat graphical element input and contour, respectively. The average time taken by the user to find a particular graphic element was counted and compared, and the results are shown in Figure 11. In the results obtained using the

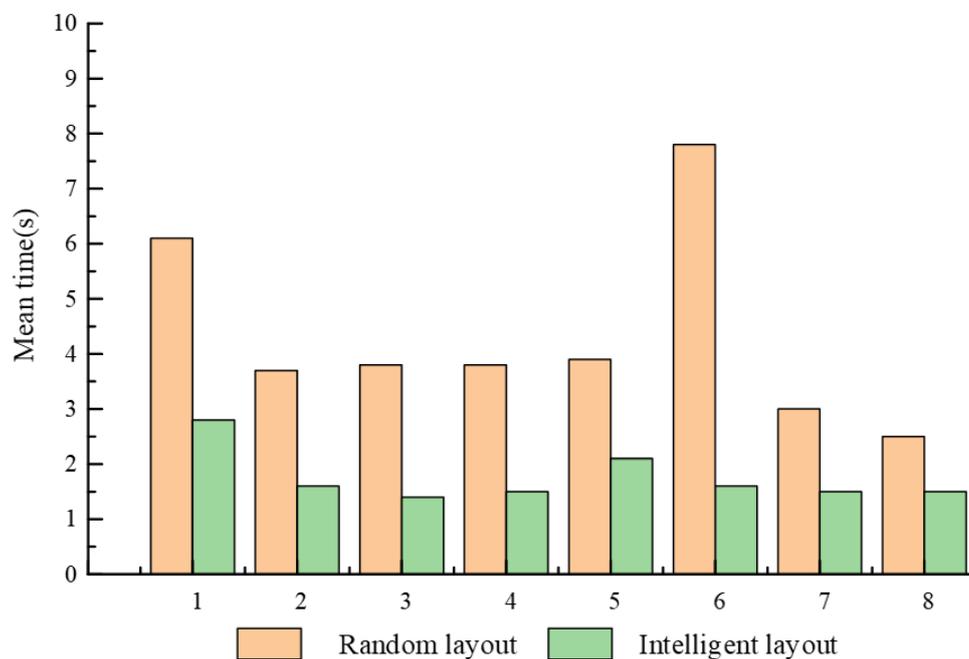


Figure 11. Graphical element query time comparison

random layout algorithm, all graphic elements are randomly distributed in the canvas contour in an irregular manner. The average time required by the user to find a graphic element is more than 4 seconds. However, in the intelligent layout method, the average time required for the user to find a graphic element is significantly reduced. The histogram of the average time statistics shows that the intelligent layout method allows the user to navigate through the various images more quickly.

5.5. **Experiment IV.** The final results of the random layout and the proposed method were compared for the same flat graphic element input and contour, respectively. User satisfaction with the two types of visual communication was counted and compared separately, as shown in Figure 12. The statistical bar chart of user satisfaction scores shows

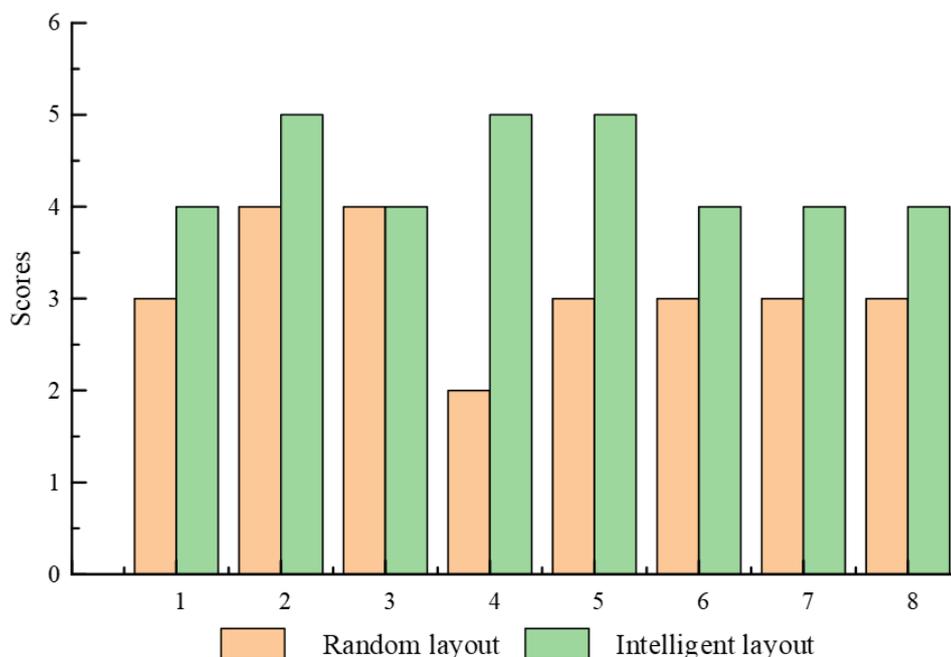


Figure 12. Comparison of user satisfaction scores

that for the results generated using the random layout algorithm, the average user satisfaction scores is around 3. The highest score is 6 and the lowest score is 0. The higher the score the higher the user satisfaction. For the intelligent layout method, the average user satisfaction scores are around 4.5. The results of the experiment show that the intelligent layout method can improve user satisfaction with the visual representation.

6. **Conclusion.** In order to facilitate the user's browsing of images and to fully express the user's interest, this paper designs a content-dependent intelligent layout method based on variable appearance, which is capable of arranging and displaying a collection of input planar graphic elements in a rational manner according to the content. Firstly, a modified Gabor wavelet feature extraction algorithm is used to extract multiple features of the graphic elements, and a feature distance matrix is constructed by calculating the distance between the graphic elements. Then, the IsoMap algorithm is used to map the feature distance matrix to a two-dimensional space. Finally, the DNN was used to identify and translate the reduced-dimensional matrix according to the user concerned images, so as to obtain the matching results between the planar graphic elements and their spatial locations. The effectiveness of the proposed intelligent layout method is verified through comparative experiments. The experimental results show that similar graphic elements are clustered together in the results generated by the proposed intelligent layout method. At the same time, user-focused images are placed in the centre of the contour. Subsequent studies will try to use different feature extraction methods, such as Sift features and Gist features, to pass more accurate image feature information.

REFERENCES

- [1] Y. Ma, Y. Peng, and T.-Y. Wu, "Transfer learning model for false positive reduction in lymph node detection via sparse coding and deep learning," *Journal of Intelligent & Fuzzy Systems*, vol. 43, no. 2, pp. 2121-2133, 2022.
- [2] F. Zhang, T.-Y. Wu, Y. Wang, R. Xiong, G. Ding, P. Mei, and L. Liu, "Application of Quantum Genetic Optimization of LVQ Neural Network in Smart City Traffic Network Prediction," *IEEE Access*, vol. 8, pp. 104555-104564, 2020.
- [3] F. Zhang, T.-Y. Wu, J.-S. Pan, G. Ding, and Z. Li, "Human motion recognition based on SVM in VR art media interaction environment," *Human-centric Computing and Information Sciences*, vol. 9, no. 1, 40, 2019.
- [4] F. Zhang, T.-Y. Wu, and G. Zheng, "Video salient region detection model based on wavelet transform and feature comparison," *EURASIP Journal on Image and Video Processing*, vol. 2019, no. 1, 58, 2019.
- [5] L. M. Corrigan, "Visual Rhetoric and Oppositional Consciousness: Poster Art in Cuba and the United States," *Intertexts*, vol. 18, no. 1, pp. 71-91, 2014.
- [6] J. Pavelka, "Remediation of Sign Texts as the Theme of Cultural Studies," *Procedia - Social and Behavioral Sciences*, vol. 217, pp. 1233-1240, 2016.
- [7] K. D. Allan, "City of Degenerate Angels: Wallace Berman, Jazz, and *Semina* in Postwar Los Angeles," *Art Journal*, vol. 70, no. 1, pp. 70-91, 2011.
- [8] J. Pavelka, "Remediation of Sign Texts as the Theme of Cultural Studies," *Procedia - Social and Behavioral Sciences*, vol. 217, pp. 1233-1240, 2016.
- [9] E. K. Wang, C.-M. Chen, M. M. Hassan, and A. Almogren, "A deep learning based medical image segmentation technique in Internet-of-Medical-Things domain," *Future Generation Computer Systems*, vol. 108, pp. 135-144, 2020.
- [10] K. Wang, C.-M. Chen, M. S. Obaidat, S. Kumari, S. Kumar, and J. Long, "Deep Semantics Sorting of Voice-Interaction-Enabled Industrial Control System," *IEEE Internet of Things Journal*, vol. 10, no. 4, pp. 2793-2801, 2023.
- [11] K.-K. Tseng, C. Wang, T. Xiao, C.-M. Chen, M. M. Hassan, and V. H. C. de Albuquerque, "Sliding large kernel of deep learning algorithm for mobile electrocardiogram diagnosis," *Computers & Electrical Engineering*, vol. 96, 107521, 2021.
- [12] S. Frolov, A. Sharma, J. Hees, T. Karayil, F. Raue, and A. Dengel, "AttrLostGAN: Attribute Controlled Image Synthesis from Reconfigurable Layout and Style," *Lecture Notes in Computer Science, Springer International Publishing*, vol. 2021, pp. 361-375, 2021.
- [13] J. Du, and Y. Long, "Landscape Image Layout Optimization Extraction Simulation of 3D Pastoral Complex under Big Data Analysis," *Complexity*, vol. 2020, pp. 1-11, 2020.
- [14] H. Tang, and N. Sebe, "Layout-to-Image Translation with Double Pooling Generative Adversarial Networks," *IEEE Transactions on Image Processing*, vol. 30, pp. 7903-7913, 2021.
- [15] M. F. Sadique, and S. M. R. Haque, "Content-Based Image Retrieval Using Color Layout Descriptor, Gray-Level Co-Occurrence Matrix and K-Nearest Neighbors," *International Journal of Information Technology and Computer Science*, vol. 12, no. 3, pp. 19-25, 2020.
- [16] S. S. Kumar, P. Rajendran, P. Prabakaran, and K. P. Soman, "Text/Image Region Separation for Document Layout Detection of Old Document Images Using Non-linear Diffusion and Level Set," *Procedia Computer Science*, vol. 93, pp. 469-477, 2016.
- [17] M. Dehghan, K. Faez, M. Ahmadi, and M. Shridhar, "Handwritten Farsi (Arabic) word recognition: a holistic approach using discrete HMM," *Pattern Recognition*, vol. 34, no. 5, pp. 1057-1065, 2001.
- [18] X. Ma, and X. Meng, "Image Position and Layout Effects on User Engagement of Multi-image Tweets," *Proceedings of the Association for Information Science and Technology*, vol. 58, no. 1, pp. 490-494, 2021.
- [19] G. Chen, X. Song, H. Zeng, and S. Jiang, "Scene Recognition with Prototype-Agnostic Scene Layout," *IEEE Transactions on Image Processing*, vol. 29, pp. 5877-5888, 2020.
- [20] W. Zhang, Q. Zhang, W. Zhang, J. Gu, and Y. Li, "From Edge to Keypoint: An End-to-End Framework for Indoor Layout Estimation," *IEEE Transactions on Multimedia*, vol. 23, pp. 4483-4490, 2021.
- [21] C. L. Srinidhi, O. Ciga, and A. L. Martel, "Deep neural network models for computational histopathology: A survey," *Medical Image Analysis*, vol. 67, pp. 101813, 2021/01, 2021.

- [22] T. Bouwmans, S. Javed, M. Sultana, and S. K. Jung, "Deep neural network concepts for background subtraction: A systematic review and comparative evaluation," *Neural Networks*, vol. 117, pp. 8-66, 2019.
- [23] E. Wang, J. J. Davis, R. Zhao, H.-C. Ng, X. Niu, W. Luk, P. Y. K. Cheung, and G. A. Constantinides, "Deep Neural Network Approximation for Custom Hardware," *ACM Computing Surveys*, vol. 52, no. 2, pp. 1-39, 2019.
- [24] D. Bau, J.-Y. Zhu, H. Strobel, A. Lapedriza, B. Zhou, and A. Torralba, "Understanding the role of individual units in a deep neural network," *Proceedings of the National Academy of Sciences*, vol. 117, no. 48, pp. 30071-30078, 2020.
- [25] S. Bianco, R. Cadene, L. Celona, and P. Napoletano, "Benchmark Analysis of Representative Deep Neural Network Architectures," *IEEE Access*, vol. 6, pp. 64270-64277, 2018.
- [26] O. Gupta, and R. Raskar, "Distributed learning of deep neural network over multiple agents," *Journal of Network and Computer Applications*, vol. 116, pp. 1-8, 2018/08, 2018.
- [27] Y. Wan, and D. Shi, "Joint Exact Histogram Specification and Image Enhancement Through the Wavelet Transform," *IEEE Transactions on Image Processing*, vol. 16, no. 9, pp. 2245-2250, 2007.
- [28] X. Li, "Numerical solution of fractional differential equations using cubic B-spline wavelet collocation method," *Communications in Nonlinear Science and Numerical Simulation*, vol. 17, no. 10, pp. 3934-3946, 2012.
- [29] Y. Zhang, W. Li, L. Zhang, X. Ning, L. Sun, and Y. Lu, "Adaptive Learning Gabor Filter for Finger-Vein Recognition," *IEEE Access*, vol. 7, pp. 159821-159830, 2019.
- [30] Y. Yuan, L.-N. Wang, G. Zhong, W. Gao, W. Jiao, J. Dong, B. Shen, D. Xia, and W. Xiang, "Adaptive Gabor convolutional networks," *Pattern Recognition*, vol. 124, 108495, 2022.
- [31] S. Serte, and H. Demirel, "Gabor wavelet-based deep learning for skin lesion classification," *Computers in Biology and Medicine*, vol. 113, pp. 103423, 2019/10, 2019.
- [32] S. Sweeney, "App-titude: Zooming In on One Multitalented App," *The ASHA Leader*, vol. 19, no. 1, pp. 38-39, 2014.
- [33] D. Young, "Creating Innovative Student Projects with App Smashing," *Educational Horizons*, vol. 93, no. 1, pp. 12-15, 2014.
- [34] Y. Gan, Y. Zhang, Z. Sun, and H. Zhang, "Qualitative photo collage by quartet analysis and active learning," *Computers & Graphics*, vol. 88, pp. 35-44, 2020.
- [35] Z. Yang, Q. Dai, and J. Zhang, "Visual perception driven collage synthesis," *Computational Visual Media*, vol. 8, no. 1, pp. 79-91, 2021.
- [36] J. Naranjo-Torres, M. Mora, R. Hernández-García, R. J. Barrientos, C. Fredes, and A. Valenzuela, "A Review of Convolutional Neural Network Applied to Fruit Image Processing," *Applied Sciences*, vol. 10, no. 10, 3443, 2020.
- [37] A. Aiyer, K. Pyun, Y.-z. Huang, D. B. O'Brien, and R. M. Gray, "Lloyd clustering of Gauss mixture models for image compression and classification," *Signal Processing: Image Communication*, vol. 20, no. 5, pp. 459-485, 2005.
- [38] S. Rebay, "Efficient Unstructured Mesh Generation by Means of Delaunay Triangulation and Bowyer-Watson Algorithm," *Journal of Computational Physics*, vol. 106, no. 1, pp. 125-138, 1993.
- [39] E. Odhiambo Omuya, G. Onyango Okeyo, and M. Waema Kimwele, "Feature Selection for Classification using Principal Component Analysis and Information Gain," *Expert Systems with Applications*, vol. 174, 114765, 2021.