Hindawi

Research Article

A Study of Artificial Intelligence-Based Poster Layout Design in Visual Communication

Huiyu Huo [1,2] **and Feng Wang** [1]

[1] *School of Design, Jiangnan University, Jiangsu, WuXi 214122, China*
[2] *College of Media and Art, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu 210023, China*

Correspondence should be addressed to Huiyu Huo; 7180306001@stu.jiangnan.edu.cn

In visual communication design, the basic design elements include four types of text, graphics, colour, and layout. While the first three can be called visual elements, layout design is the functional arrangement of visual elements. Layout design is as much about making the viewer receive visual information as it is about making them feel attractive. The combination of artificial intelligence and layout design has now become a popular direction in the field of visual communication design. However, the automatic layout design process achieved through an a priori design framework still requires human involvement and is a semi-intelligent application. To solve the above problems, this study proposes a poster layout design method based on artificial intelligence. The layout composition method consists of a learner and a generator. Firstly, the learner uses the spatial transformation network to learn the classification of layout composition elements and form the initial layout design templates for different composition cases. Secondly, the generator optimises the initial template based on the LeNet architecture using the golden ratio and trilateration parameters to produce multiple optimised templates. The templates are then stored in a library of corresponding templates according to their composition and framing style. The experimental results show that the proposed poster layout composition method achieves a higher accuracy rate than existing methods.

1. Introduction

Visual communication design, often also referred to as graphic design, is a creative act that encompasses graphics, text, colour, and layout. The purpose of visual communication design is to convey emotion and information [1–5]. In the Internet era, visual communication design has long since gone beyond the design laws of traditional graphic design. Cross-border integration of visual communication design has become an important development trend. In visual communication design, the basic design elements include four types of text, graphics, colour, and layout [6–8]. The first three can be called visual elements, while layout design is a relatively independent design art.

At the present stage, artificial intelligence (AI) has developed into an intersection of computer science, psychology, philosophy, art, and other technical sciences [9, 10]. With the development of deep learning and neural networks, AI has started to enter society, such as voice assistants and image recognition technology. How will artificial intelligence impact the field of design? In particular, design can be understood as a purposeful act of creation [11–13]. Artificial intelligence, on the other hand, can be understood as a tool with recognition functions for helping humans solve problems. Understood from this perspective, AI can be a kind of designer's aid. In recent years, more and more scholars are also exploring the combination of AI and design.

Currently, the combination of artificial intelligence and visual communication design has produced a number of intelligent applications [14–17], such as intelligent logo design, intelligent colour matching, and intelligent image search. In the case of magazine media, for example, intelligent automatic layout design has been implemented through an a priori design framework. However, this intelligent layout

design process still requires human participation and is a semi-intelligent application. Therefore, the core problem of this study is how to carry out automatic layout design through AI.

The first task in using AI for layout design is to make the computer understand the constituent elements of the layout. This research will use deep learning techniques to locate and identify the constituent elements of the layout in the target image. Machine learning is a discipline dedicated to the study of how computers can simulate the human brain [18–22]. Machine learning can use the empirical knowledge learned to continuously improve the performance of the system itself. With the continuous improvement of the computer hardware base (e.g., the rapid development of GPUs) and the large amount of data generated by various social networks, deep learning techniques have become a very important research direction in machine learning. Deep learning is a method for learning representations of data, with the advantage of replacing manual feature extraction with efficient hierarchical feature extraction. In recent years, many large Internet companies such as Microsoft, Google, Facebook, and Baidu have also set up relevant research teams. Deep learning has been widely used in various AI tasks such as speech recognition, computer vision, and natural language processing.

In traditional machine learning, the target image processing task is divided into two main steps. The first step relies on human experts to design the features manually. The second step is to provide the extracted features to a selected classifier for model training. This approach to image processing has been well used in practice in some respects. However, the process of manual feature extraction is time-consuming and labour-intensive and inevitably introduces some subjective differences due to differences in perception. It is difficult to fully standardise the execution criteria between different people or at different moments in time for the same person, which can easily result in the loss of feature information. This has limited the development of traditional machine learning to some extent. In recent years, image processing based on deep learning can solve the above problems very well. The biggest advantage is that no manual selection of extracted features is required.

Deep learning techniques can also significantly outperform traditional machine learning methods in terms of accuracy using convolutional neural network (CNN) models to autonomously learn appropriate image features. Andrearczyk et al. [23] proposed a convolutional neural network for processing handwritten digit recognition tasks. To solve large-scale computer vision recognition tasks, Zhang et al. [24] proposed an LeNet convolutional neural network model. The advantage of this network model is the use of ReLU as the activation function of the LeNet, which solves the problem of possible gradient dispersion in the sigmoid activation function in deeper network architectures. Wang et al. [25] proposed the dropout method, which effectively alleviates the problem of model overfitting during the training process of deep networks. The innovation of this method is that the neurons in the fully connected layer are inactivated with a certain probability (typically 0.5) during the training phase. The dropout method removes some of the neurons from the forward and backward parameter propagation, greatly reducing the interdependence between neurons and thus ensuring that mutually independent important features are extracted. Samir et al. [26] proposed a local response normalisation (LRN) to create a competitive mechanism for the activity of local neurons. LRN enhances the generalisation ability of the model by expanding the values of local neurons with larger responses and suppressing other neurons with smaller feedback. In addition to this, researchers have also tried to improve image processing accuracy through a number of conventional improvements. For example, the ReLU activation function has been modified into the very good fitting Maxout activation function, which not only inherits all the advantages of the ReLU activation function but also avoids the problem of "necrosis" of neurons due to negative gradients being set to zero.

Although all of these improvements improve the learning ability of the model to some extent, the AI-based layout design process involves more elements with greater spatial variation. For this problem, the above methods are obviously not effective solutions. Therefore, a layout composition method based on a learner and a generator is proposed. The main function of the learner is to use spatial transformer networks (STNs) [27] to classify layout composition elements and finally to realise the recognition and localisation of layout composition elements for learning the layout of selected poster cases and forming the initial layout template. Once the learning of the classification of the layout elements is completed, the element positions can be located. The main function of the generator is to optimise the initial template using the LeNet convolutional neural network. As the class, number, and position of the elements are already determined in the initial template, these parameters are only used to adjust the position of the elements. The final output of the generator is a parameter-optimised template, which is stored in a template library.

The main innovations and contributions of this study include the following:

(1) In this study, the popular LeNet convolutional neural network is used as the base network architecture and the STN is inserted in its input layer. The STN can be trained together with the LeNet and the original poster case can be automatically detected for spatial transformation during the training process, thus enabling the network model to extract more effective layout composition elements.

(2) There is a transition state in the spatial evolution process of the layout composition. To address this problem, an angular similarity softmax (A-softmax) loss function is used to replace the original softmax loss function of the LeNet architecture, which aims to make the intra-class distance between layout composition elements of the same category smaller and smaller, while the interclass distance between layout composition elements of different categories becomes larger and larger.

The rest of the study is organised as follows: in Section 2, the poster layout composition approach analysis is studied in

detail, while Section 3 provides the detailed layout composition method based on A-softmax convolutional neural network. Section 4 provides the experiments and analysis of results. Finally, the study is concluded in Section 5.

2. Poster Layout Composition Approach Analysis

2.1. Common Ways of Composing Poster Layouts. There is currently no unified view on the composition of posters in the academic community, so there are various classification methods. This study combines the existing classification methods and summarises the commonly used poster composition methods into the following three types, namely central composition, tilted composition, and three-part composition.

Center composition places the subject on the central axis of the layout. Posters with central composition, with the main body occupying the visual center, quickly attract the attention of the audience, concise layout, and a sense of stability. Three-part composition is to place the text and the main image according to the three-part line of the layout. In the three-part composition, the subject occupies two-thirds and the rest is the text. On the whole, the three-part composition has the effect of highlighting the main body and balancing the layout. In design examples, some poster works may be difficult to be classified into a certain composition method. Because of the disorder of free layout, it is difficult to summarise and define the characteristics of composition. Therefore, to simplify the analysis, this study only studies the layout design of a single composition mode.

2.2. Selection of Layout Adjustment Parameters. At present, the most commonly used layout design method is layout segmentation based on grid system, which is characterised by the use of mathematical calculation to divide the layout, thus guiding the layout of visual elements. In the development of modern design, grid-based layout segmentation is widely used in the field of poster design. Mathematical parameters commonly used in layout design include golden ratio, Fibonacci sequence [28], Vandergraff layout structure principle, and photography three-part method. The ratio of length to width in the golden ratio is about 0.618. In layout design, the most common use is to use golden rectangle to create grid. The three-part method is a composition technique used in photography. The photographer divides the length and width of the shot into three equal parts to form four intersection points and focuses the lens on the intersection points and near the three-point line.

In this study, the golden ratio and the three parts were chosen to optimise the layout template parameters. This is mainly because these two layout composition parameters have a more established and common use in guiding graphic layouts. In particular, when using the golden ratio division in poster layout design, the layout is first divided into golden ratio areas. Then, based on a reasonable reading order, the subject is placed at the visual focal point and the position and size of the elements are adjusted according to the golden ratio. When using the three parts, the position of the subject is first clarified, and the position of the text is used to determine the position of the subject to achieve different visual effects, such as the left and right placement of images and text, central placement, and diagonal placement.

3. Layout Composition Method Based on A-Softmax Convolutional Neural Network

Poster images present a wide variety of structures. Learning the classification of layout composition elements in poster images is the key to layout composition. However, due to the subjective nature of human extracted features, traditional machine learning methods cannot adequately characterise poster images, and complex layout composition elements are difficult to be fully described. In recent years, convolutional neural networks have been widely used in various image classification tasks due to their adaptive feature hierarchical learning capability. Compared with traditional machine learning methods, the process of feature extraction by convolutional neural networks is much simpler. The features extracted by convolutional neural networks are more capable of characterisation, and the classification accuracy has been substantially improved. Therefore, in this study, an improved LeNet architecture is proposed to achieve automatic classification learning of layout composition elements in poster images. In addition, several optimisation templates are output based on the trained classification network model.

The steps of the layout composition method based on A-softmax convolutional neural network are as follows: 1) the learner learns the layout composition elements categorically using STN to form the initial layout design templates for different composition cases and 2) the generator optimises the initial templates using the golden ratio and trilateration parameters on the basis of LeNet architecture to output multiple design solutions.

Due to the large spatial variation of layout composition elements in poster images, STN is embedded in the input layer of the LeNet as shown in Figure 1. Using STN, the whole network automatically learns more efficient layout composition elements during the training process. A-Softmax loss function is used to supervise the optimisation of the network model, forcing the network model to learn more discriminative layout composition elements and finally obtaining satisfactory classification learning results.

3.1. STN-Based Learner. As mentioned above, the proposed layout composition method is composed of a learner and a generator. The core function of the learner is to identify the classified layout composition elements. The STN is used for the recognition and classification learning of the layout composition elements. The aim is to locate and classify feature elements from an image and to obtain information on the position and size of the elements.

This study uses the popular LeNet convolutional neural network as the base network architecture and inserts an STN into its input layer, which can be trained together with the
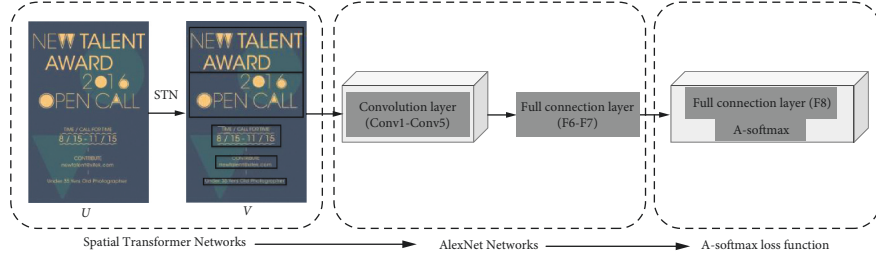
FIGURE 1: Layout composition method based on A-softmax convolutional neural network.
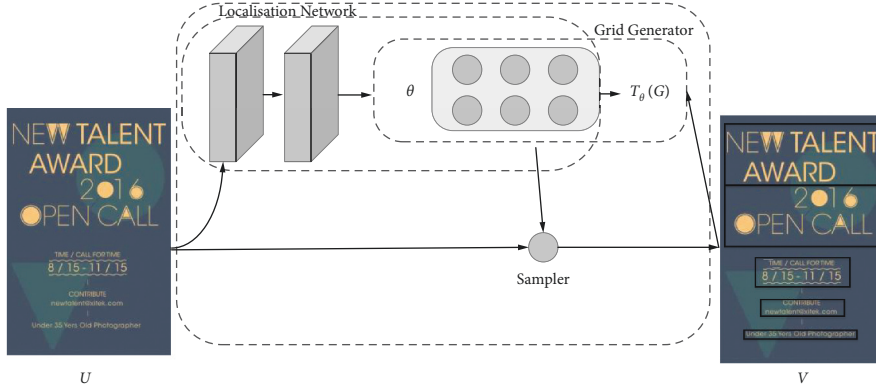

FIGURE 2: Spatial transform network model.

LeNet and can automatically detect the original poster case for spatial transformation during the training process, thus enabling the network model to extract more efficient layout composition elements. As shown in Figure 2, the STN can be divided into three parts.

As shown in Figure 2, the localisation network in the first part is a custom convolutional neural network architecture used to generate 2D affine transformation parameters. The input of learner is an initial set of poster images $U \in \mathbb{R}^{H \times W}$ with a width of H and a height of W. The output is a matrix transformation parameter θ of size $2 * 3$ obtained from the regression layer.

In the second section, the grid generator is used to solve for the mapping of each pixel coordinate with image V and initial image U. The mapping relationship between (x_i^t, y_i^t) and (x_i^s, y_i^s) is shown as follows:

$$
(x_i^s, y_i^s) = T_\theta(G_i) = A_\theta \begin{pmatrix} x_i^t \\ y_i^t \\ 1 \end{pmatrix} = \begin{pmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \end{pmatrix} \begin{pmatrix} x_i^t \\ y_i^t \\ 1 \end{pmatrix},
$$

$$(1)$$

where T_θ represents the mapping matrix function, which is composed of the transformation parameters θ obtained in the first part. G_i denotes the mapping space grid coordinates. A_θ denotes the transformation matrix.

Finally, the sampler samples the pixel coordinates in the target image using the coordinate results obtained in the second part. Since some of the positions in the mapped initial image may be fractional, the sampling needs to be determined jointly by the other pixel values around that

coordinate, hence the bilinear interpolation method used in this study.

$$
V_i = \sum_n^H \sum_m^W U_{nm} \max\left(0, 1 - |x_i^s - m|\right) \max\left(0, 1 - |y_i^s - n|\right),
$$

$$(2)$$

where n and m denote the positions of the surrounding coordinates of the coordinates (x_i^s, y_i^s) in the initial image U, U_{nm} denotes the pixel value at a point in the initial image U, and V_i denotes the sampling value of the target image V at the ith pixel point. The above is the forward propagation process of the STN. Because the network can be trained to learn with a neural network model, the back propagation equation is shown as follows:

$$
\frac{\partial V_i}{\partial U_{nm}} = \sum_n^H \sum_m^W \max\left(0, 1 - |x_i^s - m|\right) \max\left(0, 1 - |y_i^s - n|\right). \quad (3)
$$

$$
\frac{\partial V_i}{\partial x_i^s} = \sum_n^H \sum_m^W U_{nm} \max\left(0, 1 - |x_i^s - n|\right) \begin{cases} 0 & |m - x_i^s| \geq 1 \\ 1 & m \geq x_i^s \\ -1 & m < x_i^s \end{cases} . \quad (4)
$$

Derivation of transformation parameters θ can be obtained as follows:

$$
\frac{\partial V_i}{\partial \theta} = \begin{pmatrix} \dfrac{\partial V_i}{\partial x_i^s} \cdot \dfrac{\partial x_i^s}{\partial \theta} \\ \\ \dfrac{\partial V_i}{\partial y_i^s} \cdot \dfrac{\partial y_i^s}{\partial \theta} \end{pmatrix}. \quad (5)
$$

The use of convolution and maximum pooling in traditional CNNs achieves some degree of translation invariance, but the artificially set transformation rules make the network overly dependent on a priori knowledge. As a result, CNNs are neither truly translation-invariant nor feature-invariant to non-artificial geometric transformations such as rotations and distortions. However, the STN with its derivative nature does not require redundant annotation and can adaptively learn how to transform the space for different data. STNs enable spatially invariant classification network models, which have been applied to tasks such as numerical recognition and face recognition.

3.2. LeNet-Based Generator. The final output of the learner is the original layout, while the generator optimises the original layout based on the corresponding optimisation parameters and produces several optimised layout templates. The generator optimises the initial template based on the LeNet architecture using the golden ratio and trilateration parameters to produce multiple optimised templates. The golden ratio is a strictly mathematical and artistic visual feature that is commonly used in layout design, sculpture, painting, and other fields. In layout design, the use of the golden ratio can often help to achieve harmony in the layout. The use of the golden ratio is divided into two main parts: firstly, it is used for the layout of visual elements to optimise their placement; secondly, it is used to adjust the size of text elements to create a logical reading hierarchy. The three-part method is a technique commonly used in painting and photography. The three-part method starts by dividing the image into three equal parts, then placing the focus of the image on the three-quarter line or its intersection, and setting the proportion of the object on the screen according to the different weights.

LeNet convolutional neural networks have powerful image characterisation capabilities and are widely used in many image classification tasks [29]. In this study, the LeNet is used as the base classification network architecture for feature learning of poster images. The network consists of one input layer (input), five convolutional layers (C1~C5), three maximum pooling layers (S1, S2, and S5), and three fully connected layers (F6, F7, and F8), as shown in Figure 3.

Convolution layers are convolved with different types of convolution kernels to extract different features. Three maximum pooling layers are located after C1, C2, and C5 to increase the nonlinearity between features and the spatial invariance of the features. The fully connected layers (F6 and F7) can be progressed to extract more refined features. The final layer (F8) is used to generate the category labels for the posters. In addition, ReLU activation functions are added after each convolutional and fully connected layer for solving the gradient dispersion problem when the network is deeper.

To prevent the network from overfitting during training, dropout layers were added after the two fully connected layers (F6 and F7) to randomly ignore a portion of the neurons. Also, during the training phase of the network, the original poster image with an input size of $256 * * 256$ was
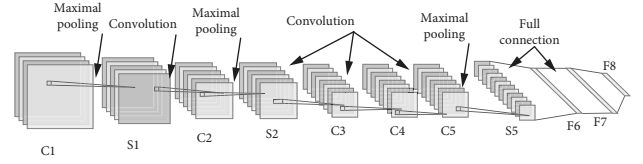

FIGURE 3: LeNet architecture.

randomly cropped $(224 * * 224)$ in the input layer to enhance the complexity of the data. The network also places the LRN behind the activation functions in layers C1 and C2. The aim is to improve the speed of convergence of the classification network and to enhance the generalisation ability of the network model.

3.3. A-Softmax Loss Function. In machine learning, commonly used loss functions include contrastive loss and triple loss. These two loss functions can reduce the intra-class spacing and increase the interclass distance between classes to a certain extent. However, these two loss functions require a certain degree of selectivity in the training samples, and CNNs are usually used to train large datasets, which is undoubtedly an increased workload and time-consuming. In contrast, the softmax loss function is widely used in many convolutional neural network architectures due to its simplicity and strong classification accuracy.

The original softmax loss function cannot force the intra-class distance between the same classes to become smaller. To effectively address this problem, this study uses angular similarity to extend the softmax loss function to a more general A-softmax loss function. The A-softmax loss function enables smaller intra-class distances between the same classes and larger interclass distances between different classes. In general, when the category label of the ith input feature x_i is y_i, the definition of the loss function is shown as follows:

$$L = \frac{1}{N} \sum_i L_i = \frac{1}{N} \sum_i -\log\left(\frac{e^{f_{y_i}}}{\sum_j e^{f_j}}\right), \tag{6}$$

where N is the number of training samples and f_j is the score vector of the jth category. Since f is the product of the activation function output weight W and the input x_i, f_{y_j} can be written as follows: $f_{y_j} = W_{y_i} x_i$. When ignoring offsets, f_j can also be expressed as follows:

$$f_j = \|W_j\| \|x_i\| \cos(\theta_j), 0 \leq \theta_j \leq \pi, \tag{7}$$

where θ_j is the angle between the weight vector W and the input feature x_i. Thus, the softmax loss function can be expressed as follows:

$$L_i = -\log\left(\frac{e^{\|W_{y_i}\| \|x_i\| \psi(\theta_{y_i})}}{\sum_j e^{\|W_j\| \|x_i\| \cos(\theta_j)}}\right). \tag{8}$$

For a binary classification problem, the original softmax loss function will generally satisfy the condition $\|W_1\| \|x\| \cos(\theta_1) > \|W_2\| \|x\| \cos(\theta_2)$, resulting in the correct

result for the input data x as category 1. However, the A-softmax loss function is motivated by the desire to constrain the above equation more tightly by adding a positive integer variable m. Such a constraint places a higher demand on the process by which the model learns the parameters W_1 and W_2, resulting in a wider categorical decision boundary between category 1 and category 2. Extending to the more general multiple category classification problem, the A-softmax loss function can be defined as follows:

$$L_i = -\log\left(\frac{e^{\|W_{y_i}\|\|x_i\|\psi(\theta_{y_i})}}{e^{\|W_{y_i}\|\|x_i\|\psi(\theta_{y_i})} + \sum_{j\neq y_i} e^{\|W_j\|\|x_i\|\cos(\theta_j)}}\right), \quad (9)$$

where $\psi(\theta)$ can be expressed as follows:

$$\psi(\theta) = \begin{cases} \cos(m\theta), & 0 < \theta \leq \dfrac{\pi}{m} \\ \\ D(\theta), & \dfrac{\pi}{m} < \theta \leq \pi \end{cases}. \quad (10)$$

When $m = 1$, the expression of the A-softmax loss function (9) is equivalent to the expression of the original softmax loss function (8). When m is larger, the wider the decision boundary for classification, the more difficult it is for the model to learn. In equation (10), $D(\theta)$ is a monotonically decreasing function used to ensure that $\psi(\theta)$ is a continuous function. To simplify the forward and backward propagation of the A-softmax loss function during training, $\psi(\theta)$ is defined as follows:

$$\psi(\theta) = (-1)^k \cos(m\theta) - 2k, \theta \in \left[\frac{k\pi}{m}, \frac{(k+1)\pi}{m}\right], \quad (11)$$

where k is a positive integer.

To reveal the inherent distance distribution of the poster image data so as to reflect the proximity between features, feature extraction was carried out in this study using the fully connected layer (F7) with a more complete feature abstraction. A simple Euclidean distance formula was used to calculate the distance to the test data, resulting in a distance matrix of $2184 * 2184$. This distance matrix is sometimes referred to as the phase difference matrix and is used to reflect the relative confusion between features. The formula for the Euclidean distance is shown as follows:

$$D(x)_{ij} = \left(\sum_k \left(x_{ik} - x_{jk}\right)^2\right)^{1/2} k \in [1, 4096], \quad (12)$$

where $D(x)_{ij}$ denotes the Euclidean distance between the ith image and the jth image feature in dataset x. k denotes a particular feature dimension for which the feature vector needs to be computed.

4. Experiments and Analysis of Results

4.1. Experimental Design and Evaluation Indicators.
The aim of the experiment was to verify whether the proposed layout composition method could generate a useable design solution. The experimental design is based on the following idea: first, a product poster case that meets the relevant requirements is selected from the current poster template library of the AI design tool. Then, a learner is used to simulate the target detection model so as to locate the coordinates of the different elements. Secondly, the generator is used to generate output multiple optimised design solutions. Finally, the order of the designed posters is disordered and the simulated design results are left unmarked, allowing the tester to score the unmarked design solutions.

As mentioned above, the A-softmax loss function is used in this study to supervise a network model generated by combining STN and LeNet. The parameter weights are optimised by stochastic gradient descent (SGD) [30] and back propagation principles (BPPs) [31]. The initial learning rate was 0.00003, which was automatically reduced to 0.1 times the initial value after the network model was iterated 5000, 8000, and 12000 times, respectively. The momentum and weight decay were set to 0.9 and 0.0005, respectively. The PC used for the experiments had a 3.8 GHz AMD 5800X CPU, and the system was Windows 10.

Considering that the A-softmax loss function makes it difficult for the network model to converge during the training process, this study adds a decay factor to the learning strategy λ.

$$f_{y_i} = \frac{\lambda\|W_{y_i}\|\|x_i\|\cos(\theta_{y_i}) + \|W_{y_i}\|\|x_i\|\psi(\theta_{y_i})}{1 + \lambda}. \quad (13)$$

The decay factor λ is a large positive integer value at the beginning of the gradient descent. As the number of generations of the network model increases, the value of the decay factor λ decreases and stops at the minimum value. In the experiments in this study, the initial value of λ is 100000 and the minimum value is 15.

No experts or scholars have yet established evaluation criteria for the layout design of posters. To quantitatively analyse the experimental results, a seven-level Likert scale was chosen to obtain the results of five (overall comprehensive evaluation, readability of textual information, consistency of information perception weights, rationality of element placement relationships, and rationality of visual paths) evaluations. The main reason for this is that the Likert scale provides quick access to the test taker's level of agreement with a viewpoint or feeling and is the most commonly used subjective evaluation tool.

4.2. Example of Layout Composition.
The posters that meet the filtering criteria are first selected from the poster template library. The filtering criteria include (1) the requirement that the elements do not block or overlap each other; (2) the composition of the poster is a central composition or a three-part composition; (3) the influence of other elements in the poster, such as colour, on the visual perception of the poster layout is as small as possible; and (4) the poster is a solid colour background or a textured background that is less intrusive to the identification of the elements. Although the layout composition method proposed in this study does not take into account the influence of colour elements,

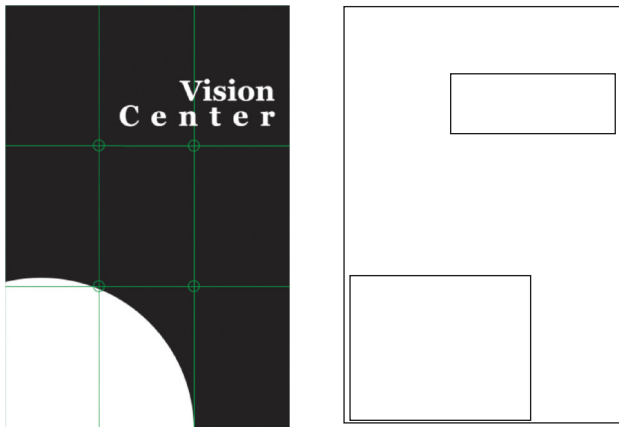Figure 4: Example of a central composition design and its initial template.

Figure 5: Example of a three-part composition design and its initial template.

colour can interfere more with the subjective evaluation of the audience when making an evaluation of the output design scheme. Therefore, the influence of this irrelevant variable on the experimental results should be minimised. Examples of the resulting layout composition are shown in Figures 4 and 5.

After completing the design of the experiment, the experiment used a seven-level Likert scale to obtain the evaluation results. Considering that the purpose of this experiment is to verify whether the output results of the model are useable or not. Therefore, the testers were required to have a certain foundation in design aesthetics, and factors such as the gender and occupation of the testers might have a large impact on the evaluation results, so the variables needed to be controlled reasonably. Ultimately, 55 students from art and design-related majors were selected for this experiment, of which the ratio of male to female was $1:1$.

4.3. Analysis of Results. In the experiment, the posters will be presented to the testers randomly to reduce subjectivity. The average score of the 5 different indicator scores will be calculated through the experiment. Higher scores on the
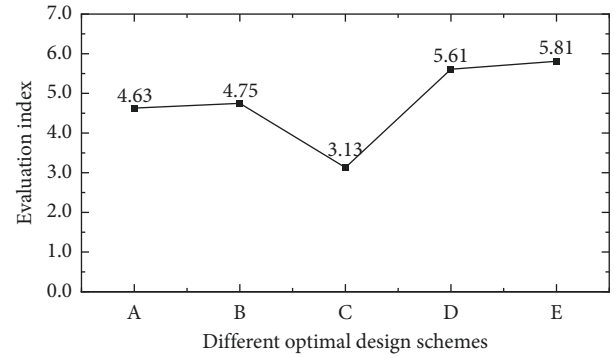

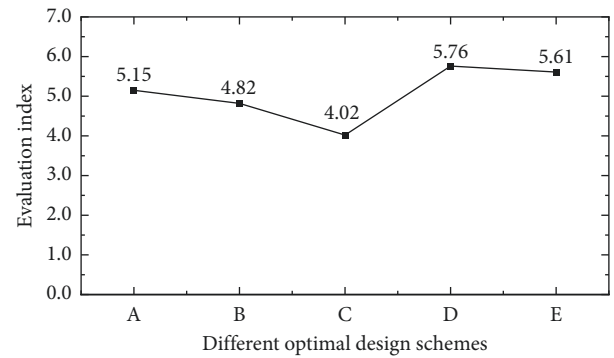Figure 6: Score statistics for overall comprehensive evaluation.

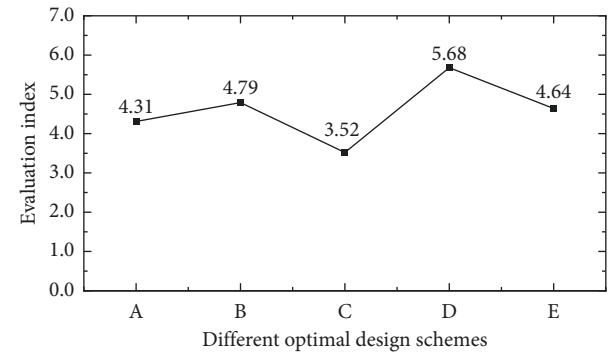Figure 7: Score statistics for readability of textual information score statistics.

Figure 8: Score statistics for consistency of information perception weights.

indicators indicate a more positive evaluation, and lower scores indicate a more negative evaluation. The statistics of the scores for the five indicators are shown in Figure 6 to Figure 10.

As can be seen from Table 1, the A-softmax loss function successfully forces the network model to distinguish between more types of layout constituent elements than the original softmax loss function. As the value of m increases, the average correct rate increases, thus demonstrating the effectiveness of the supervised optimisation network model training with the A-softmax loss function in this study.
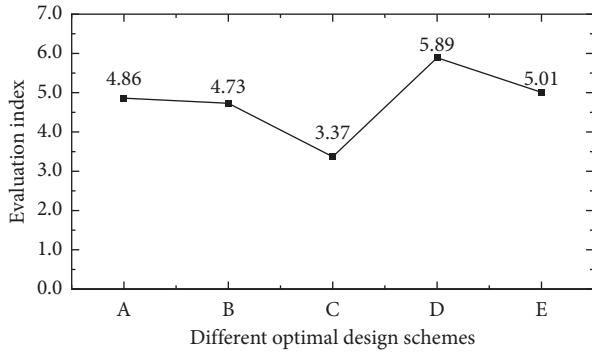
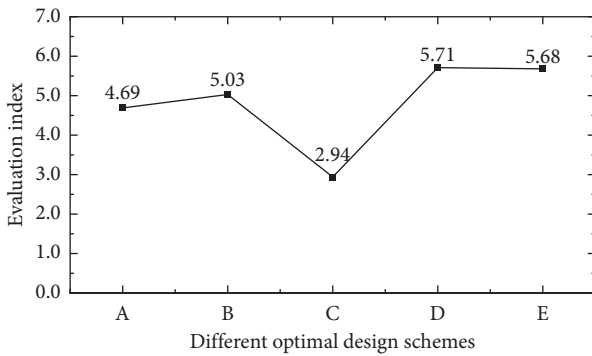Figure 9: Score statistics for rationality of element placement relationships.

Figure 10: Score statistics for rationality of visual paths.

Table 1: Comparison of the precision the LeNet combined with the A-softmax loss.

Design solutions	$m = 1$	$m = 2$	$m = 3$	$m = 4$
A	0.984	0.994	0.996	0.997
B	0.938	0.949	0.955	0.949
C	0.836	0.856	0.88	0.888
D	0.83	0.838	0.834	0.834
E	0.841	0.834	0.851	0.858
Average	0.885	0.894	0.903	0.905

Table 2: Comparison of the precision of the LeNet combined with the STN and the A-softmax loss.

Design solutions	$m = 1$	$m = 2$	$m = 3$	$m = 4$
A	0.991	0.991	0.994	0.99
B	0.955	0.963	0.967	0.984
C	0.88	0.874	0.904	0.862
D	0.828	0.858	0.862	0.878
E	0.919	0.927	0.937	0.934
Average	0.914	0.922	0.932	0.929

As can be seen from Table 2, the average correct rates all improved considerably, proving that the STN was correctly trained to spatially transform the original images during the training process. Combining the STN with the LeNet can further improve the accuracy of the model. The proposed method at $m = 3$ achieved the best design results. At $m = 3$,

Table 3: Compared with existing methods.

Design solutions	STN + LeNet + A-softmax	LeNet	VGGNet	Inception-v4
A	0.994	0.984	0.998	0.976
B	0.967	0.938	0.908	0.85
C	0.904	0.836	0.812	0.706
D	0.862	0.83	0.824	0.83
E	0.937	0.904	0.895	0.851
Average	0.932	0.898	0.887	0.842

the STN + LeNet + A-softmax model proposed in this study was compared with the LeNet, VGGNet [32], and Inception-v4 [33] network models, as shown in Table 3.

5. Conclusion

In this study, a poster layout composition method based on STN + LeNet + A-softmax is proposed. The layout composition method consists of a learner and a generator. The learner learns to classify the layout composition elements using STN and finally realises the recognition and localisation of the layout composition elements to form the initial layout template. The generator uses LeNet convolutional neural network to optimise the initial template, and the final output is the parameter-optimised template. The results show that the combination of the STN and the LeNet allows the network model to extract more effective layout components. At the same time, the A-softmax loss function guides the network model to learn the more discriminative layout elements. Compared with existing methods, STN + LeNet + A-softmax achieves better accuracy [34].

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding this study.

References

[1] R. Chalfen, "A sociovidistic approach to children's film-making: the philadelphia project," *Studies in Visual Communication*, vol. 7, no. 1, pp. 3–32, 1981.

[2] L. M. Malmsheimer, ""Imitation white man": images of transformation at the carlisle Indian school," *Studies in Visual Communication*, vol. 11, no. 4, pp. 54–75, 1985.

[3] M. Schreiber, "Audiences, aesthetics and affordances analysing practices of visual communication on social media[J]," *Digital Culture & Society*, vol. 3, no. 2, pp. 89–94, 2018.

[4] Z. Zomay, B. Keskin, and C. Ahin, "Grsel letiim tasarm blümü rencilerinin sektrel logolardaki renk tercihleri - color preferences of visual communication design students in sectoral logos[J]," *OPUS Uluslararası Toplum Araştırmaları Dergisi*, vol. 17, no. 37, pp. 4181–4198, 2021.

[5] D. K. Kilgo and R. R. Mouro, "Protest coverage matters: how media framing and visual communication affects support for black civil rights protests[J]," *Mass Communication & Society*, vol. 8, no. 1, pp. 1169–1178, 2021.

[6] G. S. Dhanesh and N. Rahman, "Visual communication and public relations: visual frame building strategies in war and conflict stories[J]," *Public Relations Review*, vol. 47, no. 1, pp. 2511–2521, 2021.

[7] D. Shinar, "Art and communications in the west bank: visual dimensions of Palestinian nation building," *Studies in Visual Communication*, vol. 10, no. 2, pp. 2–15, 1984.

[8] J. Lin, "Research on the performance of impressionist painting color visual communication based on wireless communication and machine vision," *Security and Communication Networks*, vol. 15, no. 1, pp. 1–6, 2021.

[9] G. C. Liu and C. H. Ko, "Photoshop and illustrator use in instructional design: a case study of college visual communication course design," *The International Journal of Design Education*, vol. 15, no. 1, pp. 119–130, 2021.

[10] H. Li, "Visual communication design of digital media in digital advertising," *Journal of Contemporary Educational Research*, vol. 5, no. 7, pp. 36–39, 2021.

[11] R. Liu, B. Yang, E. Zio, and X Chen, "Artificial intelligence for fault diagnosis of rotating machinery: a review," *Mechanical Systems and Signal Processing*, vol. 108, no. 8, pp. 33–47, 2018.

[12] J. H. Thrall, X. Li, Q. Li et al., "Artificial intelligence and machine learning in radiology: opportunities, challenges, pitfalls, and criteria for success," *Journal of the American College of Radiology*, vol. 15, no. 3, pp. 504–508, 2018.

[13] E. J. Piedad, Y. T. Chen, H. C. Chang, and C. C. Kuo, "Frequency occurrence plot-based convolutional neural network for motor fault diagnosis," *Electronics*, vol. 9, no. 10, p. 1711, 2020.

[14] C. S. Wang, I. H. Kao, and J. W. Perng, "Fault diagnosis and fault frequency determination of permanent magnet synchronous motor based on deep learning," *Sensors*, vol. 21, no. 11, p. 3608, 2021.

[15] K. Satpathi, Y. M. Yeap, A. Ukil, and N Geddada, "Short-time fourier transform based transient analysis of VSC interfaced point-to-point DC system," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 5, pp. 4080–4091, 2018.

[16] Y. Cheng, G. Zhang, C. Yu, P Bo, W Xuewei, and X Feng, "Application of short-time fourier transform in feeder fault detection of flexible multi-state switch[J]," *Journal of Physics: Conference Series*, vol. 1754, no. 1, pp. 149–160, 2021.

[17] E. Dupoux, "Cognitive science in the era of artificial intelligence: a roadmap for reverse-engineering the infant language-learner," *Cognition*, vol. 173, pp. 43–59, 2018.

[18] S. Garg, K. Kumar, N. Prabhakar, A. Ratan, and A Trivedi, "Optical character recognition using artificial intelligence," *International Journal of Computer Application*, vol. 179, no. 31, pp. 14–20, 2018.

[19] M. Hossein and H. Sajad, "Artificial intelligence design charts for predicting friction capacity of driven pile in clay[J]," *Neural Computing & Applications*, vol. 30, pp. 1–17, 2018.

[20] Y. F. Hu and Q. Li, "An adjusTab. envelope based EMD method for rolling bearing fault diagnosis[J]," *IOP Conference Series: Materials Science and Engineering*, vol. 1043, no. 3, pp. 327–339, 2021.

[21] G. R. Naik, S. E. Selvan, and H. T. Nguyen, "Single-channel EMG classification with ensemble-empirical-mode-decomposition-based ICA for diagnosing neuromuscular disorders," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 7, pp. 734–743, 2016.

[22] Z. Qin, H. Chen, and J. Chang, "Signal-to-Noise ratio enhancement based on empirical mode decomposition in phase-sensitive optical time domain reflectometry systems," *Sensors*, vol. 17, no. 8, p. 1870, 2017.

[23] V. Andrearczyk and P. F. Whelan, "Using filter banks in Convolutional Neural Networks for texture classification," *Pattern Recognition Letters*, vol. 84, pp. 63–69, 2016.

[24] C. Zhang, X. Yue, R. Wang, N. Li, and Y. Ding, "Study on traffic sign recognition by optimized lenet-5 algorithm[J]," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 34, no. 1, pp. 1–21, 2020.

[25] H. Wang, W. Yang, Z. Zhao, T. Luo, J. Wang, and Y. Tang, "Rademacher dropout: an adaptive dropout for deep neural network via optimizing generalization gap," *Neurocomputing*, vol. 357, no. 9, pp. 177–187, 2019.

[26] S. Samir, E. Emary, K. El-Sayed, and H. Onsi, "Optimization of a pre-trained AlexNet model for detecting and localizing image forgeries," *Information*, vol. 11, no. 5, p. 275, 2020.

[27] H. M. Kasem, K. W. Hung, and J. Jiang, "Spatial transformer generative adversarial network for robust image super-resolution," *IEEE Access*, vol. 7, pp. 182993–183009, 2019.

[28] K. Naoki and U. Hiroshi, "A construction of simple and smaller-state real-time generator for exponential sequences [J]," *Artificial Life and Robotics*, vol. 25, no. 1, pp. 64–72, 2020.

[29] S. A. Wagle and R. Harikrishnan, "Comparison of plant leaf classification using modified AlexNet and support vector machine," *Traitement du Signal*, vol. 38, no. 1, pp. 79–87, 2021.

[30] T. Ye, M. Jian, L. Chen, and Z. Wang, "Rolling bearing fault diagnosis under variable conditions using LMD-SVD and extreme learning machine[J]," *Mechanism and Machine Theory*, vol. 90, pp. 175–186, 2015.

[31] H. X. Zhou, H. Li, T. Liu, T. Liu, and Q. Chen, "A weak fault feature extraction of rolling element bearing based on attenuated cosine dictionaries and sparse feature sign search," *ISA Transactions*, vol. 97, pp. 143–154, 2020.

[32] M. Jangra, S. K. Dhull, and K. K. Singh, "ECG arrhythmia classification using modified visual geometry group network (mVGGNet)," *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 3, pp. 3151–3165, 2020.

[33] D. B. Toan, J. Lee, and J. Shin, "Incorporated region detection and classification using deep convolutional networks for bone age assessment[J]," *Artificial Intelligence in Medicine*, vol. 97, pp. 1–8, 2019.

[34] D. J. Frey, A. Mishra, M. T. Hoque, M. Abdelguerfi, and T. Soniat, "A machine learning approach to determine oyster vessel behavior," *Machine Learning and Knowledge Extraction*, vol. 1, no. 1, pp. 64–74, 2018.