


A METHOD FOR GENERATING ADVERTISING DESIGN IMAGES BASED ON HIERARCHICAL FEATURES AND SIMULATED ANNEALING ALGORITHM

Jian Zhang* 

Department of Fine Arts, Henan Vocational College of Light Industry, Zhengzhou, China

**Corresponding author: Jian Zhang (migesha@126.com)*

Submitted: 13 Jun 2025 Accepted: 25 Aug 2025 Published: 08 Dec 2025

License: CC BY-NC 4.0 

Abstract With the development of intelligent design and computer-aided design technology, advertising image generation has gradually received attention and over 70% of digital advertisers regard automated creative generation as a key direction for improving efficiency and precision delivery. To address the shortcomings of existing advertising design methods in feature extraction and optimization efficiency, a novel advertising design image generation method combining hierarchical feature extraction and simulated annealing algorithm optimization is proposed. Research is based on a hierarchical feature model to extract multi-scale semantic information from advertising images, and optimize layout through simulated annealing algorithm to improve the visual consistency of design images. The experiment outcomes show that the raised model has the highest mean fitness, especially in the first set of hyperparameter settings, with mean fitness values of 3.00 and 2.95 on the training and testing sets, respectively. Meanwhile, the standard deviation and coefficient of variation are significantly lower than for other algorithms, with minimal fluctuations and the strongest robustness. In addition, among the three types of advertising images for product promotion, brand promotion, and directive sign advertisement, the generated advertising images have significant advantages in visual clarity, perceptual quality, and other aspects. As shown in the directive sign advertisement, the mean square error, peak signal-to-noise ratio, structural similarity, and learning perceptual image patch similarity of this model are 0.025, 66.97, 0.67, and 0.10, respectively, which are significantly better than the other two comparison methods. The research results indicate that the raised model is suitable for scenarios that require high-precision image generation, providing an effective solution for intelligent advertising generation.

Keywords: hierarchical features, simulated annealing algorithm, advertising design images, feature extraction, perceived quality.

1. Introduction

Advertising design plays a crucial role in modern marketing, and its effectiveness directly affects users' perception and acceptance of the brand [7]. With the rapid advancement of digital technology, advertising image generation is gradually shifting from traditional manual design to intelligence and automation [15]. This change has injected new vitality into the advertising industry, but also brought more complexity and challenges. Advertising images not only need to have visual appeal, but also need to effectively convey core information within a limited space, such as promotional content, brand identity, and product features. How to achieve efficient design of advertising images and meet diverse commercial needs has become a hot topic of concern for both academia and

industry. In existing research, the introduction of deep learning technology has significantly facilitated the advancement of intelligent advertising design. For example, models based on generative adversarial networks and transfer learning have shown significant effects in the fields of image generation and style transfer [10]. However, these methods mainly focus on texture refinement or single feature optimization of images, making it difficult to comprehensively handle the complex characteristics of multi-dimensional and multi-level advertising images. In the generation of advertising images, it is not only necessary to accurately extract design elements such as text, background, and subject, but also to achieve coordination between visual elements through reasonable layout optimization [11]. Existing methods commonly suffer from incomplete feature extraction, unstable generation quality, and low computational efficiency when dealing with complex advertising scenarios.

Zhang et al. [22] proposed the Emocolor system, which achieved matching between feelings and hues through hue arrangement recommendations based on emotional vocabulary and images. This method combined interactive genetic algorithm to optimize emotional color schemes, which could help professional designers generate color schemes that met users' emotional needs, and was successfully applied in fields such as advertising design. The research results indicated that emotion driven color design could significantly enhance the visual appeal and user satisfaction of advertising images.

In recent years, Denoising Diffusion Probabilistic Models (DDPM) have shown great potential in the field of image super-resolution, but sampling efficiency remains a key bottleneck restricting their practical applications. Song et al. [17] integrated rough set theory with DDPM and proposed a rough set DDPM super-resolution method. This method minimizes the roughness of the sample set through rough set theory, optimizes the segmentation of the sampling sequence, and uses particle swarm optimization algorithm to screen the optimal sub columns for iterative denoising. The results indicate that compared to traditional autoregressive models, this method can generate higher quality high-resolution images with fewer sampling steps, achieving a good balance between image quality and processing speed.

In addition, hierarchical feature extraction techniques are widely applied in image processing and feature recognition [18]. Lin et al. [8] raised a layered attribute selection approach utilizing label distribution learning to address the issue of high-dimensional feature space and class imbalance. This method successfully alleviated the problem of sample imbalance, improved the distinguishing capability of attribute subsets, and enhanced the performance of downstream categorization tasks through the distribution of labels in a hierarchical structure. Research showed that hierarchical feature selection could effectively capture the multi-level structural information of images, which was helpful for extracting complex features from advertising design images.

Simulated Annealing (SA) algorithm, as a classic stochastic optimization algorithm, is widely used in layout optimization and feature adjustment [14]. Iyappan et al. [6]

proposed a hybrid algorithm that combined SA and Spotted Hyena Optimization (SHO) to solve asset allocation and job dispatch problems in cloud environments. This method ensured load balancing of virtual machine resources by balancing exploration and development, avoiding overload or underload phenomena. The findings indicated that the algorithm was capable of significantly enhancing resource allocation and operational efficiency of tasks, while reducing energy consumption and response time, which was of great significance for asset allocation and job dispatch in cloud computing environments.

In summary, there are problems in the current advertising design image generation methods, such as low optimization efficiency and difficulty in fully extracting multi-level features from advertising images. In view of this, an innovative method combining structured hierarchical feature modeling with simulated annealing optimization mechanism has been proposed. By constructing a multi-layer semantic abstraction feature model, the system integrates semantic information such as main elements, copy, and background in advertising images, and achieves multi-scale semantic representation through hierarchical annotation and feature abstraction. At the same time, the simulated annealing algorithm based on KDE density estimation is introduced in the optimization stage, which is not only used for global search of design parameters, but also constructs the objective function and acceptance criteria through the density evaluation function, effectively improving the adaptability of the model to complex design constraints and the search performance for global optimal solutions. The research aims to improve the quality and efficiency of advertising image generation by utilizing hierarchical semantic features and powerful optimization techniques.

2. Methods and materials

2.1. Construction of advertising design image feature model based on hierarchical features

The design elements of advertising images serve as the cornerstone for constructing feature models and are crucial for delving deeper into the realm of advertising image design. By quantifying these design elements, we can enhance their organization and practical applicability. The model illustrating the design features of advertising images is depicted in Fig. 1 [4]. In this Figure it can be seen that, as the hierarchy increases, the abstraction level of features also increases, which is closer to people's subjective understanding. At the same time, the difficulty of objective quantification also increases. Currently, many deep learning driven image generation techniques rely on pixel data for training, enabling functions such as image creation, style conversion, and color adjustment. Advertising images contain rich and diverse design elements [1], which increases the difficulty for users to annotate the elements.

To simplify the annotation process, the study categorizes all design elements into

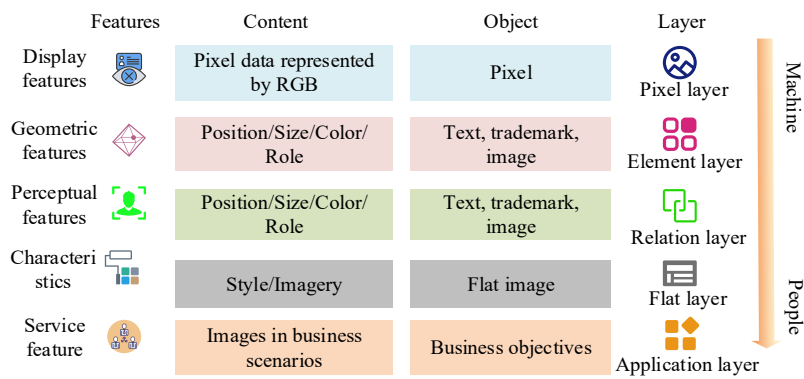


Fig. 1. Image design feature model.

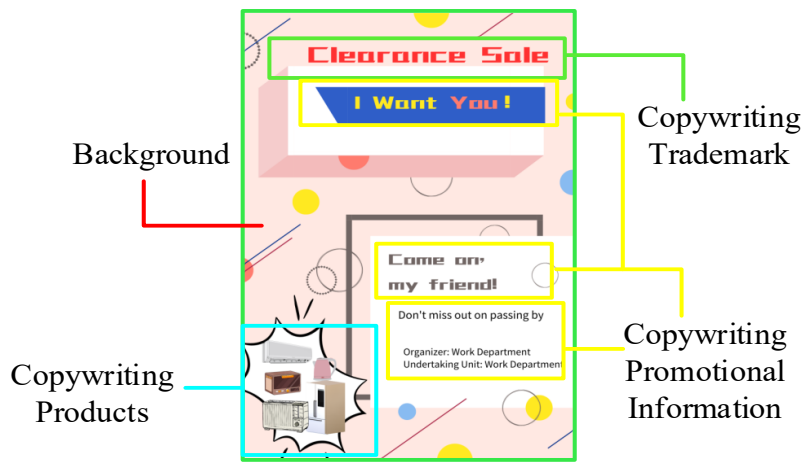


Fig. 2. Example of annotation results.

three main types: subject, copy, and background. An example of annotation is shown in Fig. 2, where the subject usually refers to the product, model, character, or product that occupies the core position. The copywriting section covers product descriptions, trademarks, and promotional information. In advertising images, the background refers to the entire background area, while the sub-background is located below the product or text, forming a sharp contrast with the main background, usually with colors or borders to highlight the product and text. The structured representation of various element

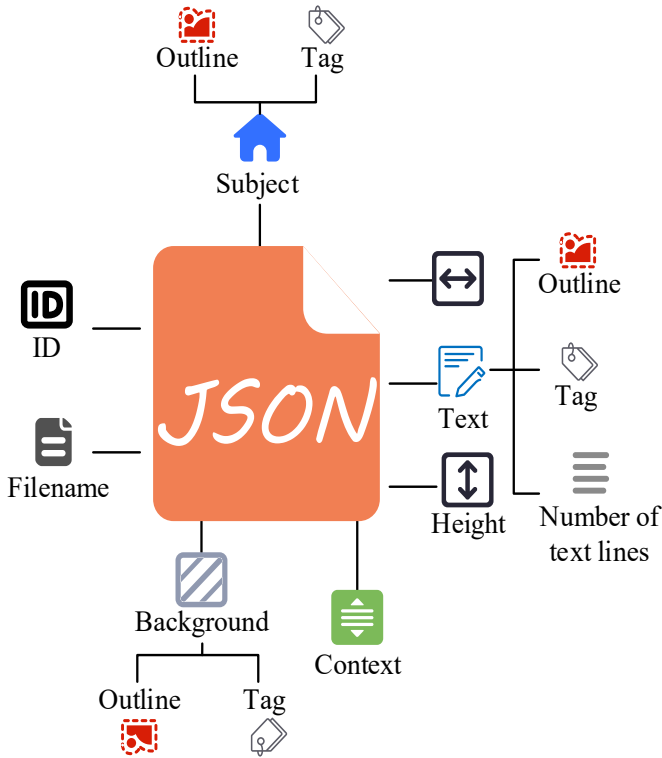


Fig. 3. Structured element feature data.

feature data is shown in Fig. 3 [19], where the annotation data of the advertising image is saved in JSON format in a CSV file for the purpose of constructing a feature model in the future. The three datasets focus on subject elements, copywriting elements, and background elements, each containing spatial location information of the corresponding elements.

2.2. Feature model optimization based on geometric features and SA algorithm

A probabilistic model, grounded in the geometric attributes of elements and utilizing quantified tomographic features, is established with the objective of being utilized within the realm of advertising image design. This model converts sampled data into a continuous probability distribution using KDE to predict feature distributions under specific

conditions [16]. The observed data are smoothed through KDE to achieve probability estimation of the true distribution. Assuming there is a set of independent sample points that follow the distribution $F(x)$, their corresponding probability density functions are shown in Equation (1) [21]:

$$F(x) = \int_{-\infty}^x f(t) dt, \quad (1)$$

where x represents the sample point, and $f(t)$ signifies the likelihood of the sample point occurring in x when $x = t$, that is, the density function. Given that the true probability density function $F(x)$ is unknown, a non-parametric distribution estimator can serve as a proxy for $F(x)$, as shown in Equation (2):

$$F_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{x_i \leq t}, \quad (2)$$

where n means the total quantity of sample points, from which the final probability density can be derived as shown in Equation (3):

$$f(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_e^i}{h}\right), \quad (3)$$

where $f(x)$ represents the final density function, x_e^i is the i th point of the sample data, and K is the kernel function. To ensure that the sum of integrals of the kernel function reaches the unit value, it is possible to replace the original kernel function $K(x)$ with another type of density function. h represents bandwidth, used to control the smoothness of the kernel function. The kernel density estimation (KDE) can be generalized from univariate variables to multidimensional variables, as expressed in Equation (4) [9]:

$$f_H(x) = \frac{1}{n|H|^{\frac{1}{2}}} \sum_{i=1}^n K\left(\frac{x - x_e^i}{H^{\frac{1}{2}}}\right), \quad (4)$$

where H represents a symmetric and positive $d \times d$ dimensional matrix, and according to the multivariate normal distribution K function, it is shown in Equation (5):

$$K\left(\frac{x - x_e^i}{H^{\frac{1}{2}}}\right) = \frac{1}{(2\pi)^{\frac{d}{2}}} \exp\left[-\frac{1}{2}(x - x_e^i)' H^{-1} (x - x_e^i)\right], \quad (5)$$

where d represents the dimension of the data, and $(2\pi)^{\frac{d}{2}}$ is the normalization factor of the Gaussian kernel. By analyzing the color characteristics of multiple brand advertising images and using KDE for modeling, the color design features of different brand advertisements are revealed. In visual design, there are interdependent and reinforcing

relationships between design elements. To establish the relationship between these elements and the target attributes, a predictive probability model can be built, utilizing conditional features to forecast the distribution of the target variables, as expressed in Equation (6):

$$F(x | C = a), \quad (6)$$

where a represents the conditional feature and C is the constraint condition. If a is regarded as the center position of the primary component in the picture, then function $F(x|C = a)$ can be applied to predict the probability distribution characteristics of the center position of the text element under given conditions. Each conditional feature corresponds to a target feature and clustering category. Multiple logistic regression is used to analyze triplet data and train a classifier to predict the probability of the target feature in different clustering categories.

A Gaussian kernel can be constructed based on clustering standard deviation to form a mixture Gaussian model that approximates a continuous probability density distribution, as shown in Equation (7):

$$P(x | C = a_e) = \sum_{i=10} \exp\left(\frac{-\|x - m_i\|^2}{2\sigma^2}\right) \cdot p(m_i | C = a_e), \quad (7)$$

where $P(x|C = a_e)$ represents the probability density of x under given condition $C = a_e$, σ represents the standard deviation, i represents a specific category in cluster analysis, and $p(m_i|C = a_e)$ is the multiple logistic classification function. In order to optimize KDE functions and conditional probability density, the study drew on the ideas of SA and adopted a geometric cooling strategy, with the updated formula shown in Equation (8):

$$T_{k+1} = \alpha \cdot T_k, \quad (8)$$

where T_k and T_{k+1} represent the temperatures after the k th and $(k + 1)$ th iterations, respectively. The cooling coefficient α is empirically set to 0.95, and the initial temperature is set to 1.0. The iteration is terminated when the temperature drops below 10^{-3} . In the initialization phase, the initial solution is randomly sampled based on the mean of the feature distribution smoothed by KDE, ensuring that the search starts from high probability regions. Using the results of the density estimation function as evaluation indicators, the design parameters that can maximize the evaluation indicators are identified through an iterative optimization process. Firstly, a candidate solution is randomly sampled from the neighborhood of the current solution, as shown in Equation (9) [13]:

$$x' = x_k + \varepsilon, \quad (9)$$

where x' represents the candidate solution, x_k represents the current solution, and ε is a random perturbation coefficient that follows a Gaussian distribution. The calculation

process for the incremental score of evaluation x' compared to x_k is shown in Equation (10):

$$\Delta F = [f(x') + \lambda P(x' | C)] - [f(x_k) + \lambda P(x_k | C)] , \quad (10)$$

where ΔF is the score increment between the potential and the current solutions, and λ is the balance coefficient. When $\Delta F > 0$, then $x_{k+1} = x'$. After updating the solution, the temperature is updated according to the cooling strategy, as shown in Equation (11):

$$T_{k+1} = \alpha T_k , \quad (11)$$

where T_k represents the temperature of the k th iteration, and α represents the temperature decay factor. The final expression of the model objective function is in Equation (12):

$$x^* = \arg \max_x [f(x) + \lambda P(x | C)] , \quad (12)$$

where x^* represents the optimal solution of the objective optimization problem. By simulating the random sampling and iterative optimization process of annealing, the global optimal solution of the objective function, i.e. the best feature, is searched in the solution space.

2.3. Advertising design image generation method based on hierarchical features and SA algorithm

The optimal features predicted based on the optimized feature model can not only enable image elements to be reasonably configured according to specific layout rules, but also enhance the overall visual effect of the image. By inputting these features, the color information of the image can be accurately restored and optimized. In this process, the study utilizes the global optimization characteristics of the SA algorithm to effectively avoid the problem of local optima and explore the optimal layout scheme and feature selection. The container layout results optimized according to the SA algorithm are shown in Fig. 4. The red border in this Figure indicates the placement area of the image container, while the yellow border is used to indicate the placement area of the document container. N_p represents the total number of document containers in a single image. In graphic advertising design, different layout features can be identified by clustering the positional relationships between elements, such as left and right, center, and top and bottom layouts. Meanwhile, considering the textual content in the image, the study explores the method of multi-line text layout based on container features predicted by the model, in order to finalize the arrangement of flat advertising visuals.

The user interface created through various compositions is shown in Fig. 5. Each layout cluster is visualized in this Figure to display its estimated container layout features. After selecting a specific cluster, the algorithm constructs a geometric feature model grounded on the clustering results. In order to evaluate the consistency between

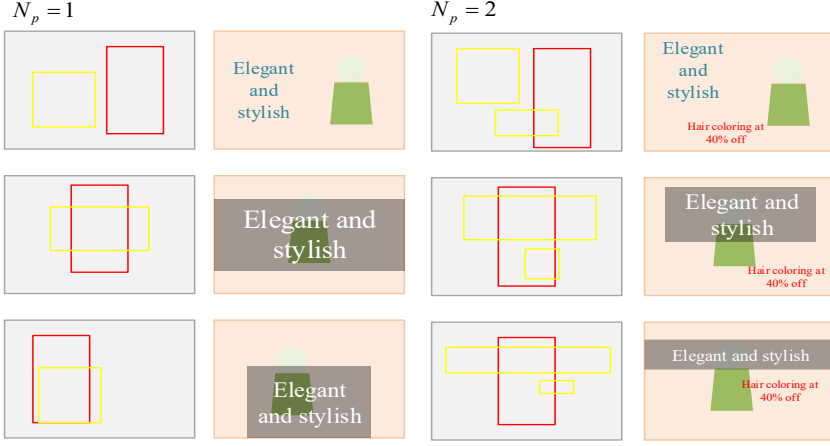


Fig. 4. Layout results based on element containers.

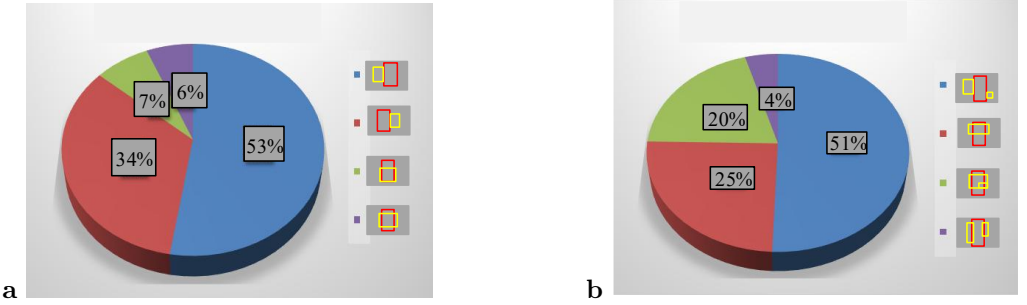


Fig. 5. Interactive interface for image layout clustering. (a) Single-text layout; (b) multi-text layout.

input characters and graphic layout features, the algorithm designs a probability model as shown in Equation (13):

$$p(T_i | L_i) \propto \prod_{k=1}^n \exp \left(\frac{-1}{2\sigma^2} \left| m^k - \frac{h_i^k}{d_i^k} \right|^2 \right), \quad (13)$$

where T_i represents the target variable, L_i represents the given condition, m^k represents the target value of the k th feature, h_i^k means the observed value of the i th sample on the k th feature, and k is the scale factor. In the case where an image contains multiple text containers, the algorithm randomly cuts characters into rows and selects the optimal value through feature sampling as the output of the layout geometry. During the visual

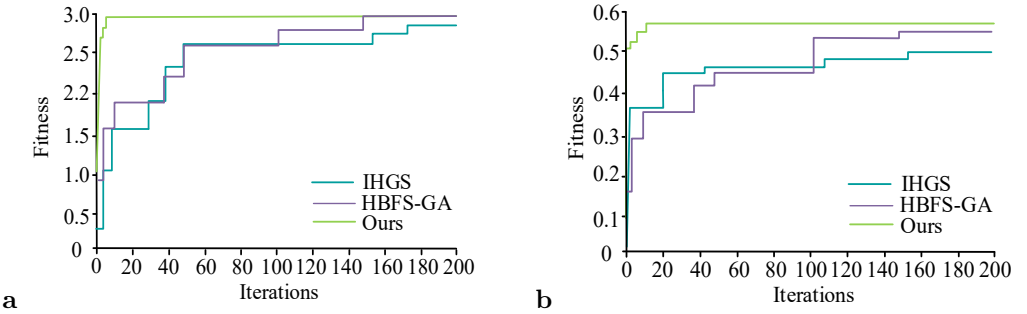


Fig. 6. The variation of fitness with the iteration times. (a) Training set; (b) test set.

interaction phase, designers have the ability to choose various sets of color attributes simply by clicking on buttons. These selected color perception attributes are then integrated into the t-Distributed Stochastic Neighbor Embedding (t-SNE) dimensionality reduction algorithm's space, resulting in the creation of corresponding two-dimensional visual representations [20].

3. Results

A detailed analysis was conducted on the variation of fitness of various algorithms with the number of iterations in the study. The Improved Hunger Games Search Algorithm (IHGS), the Hybrid Version of Binary Flamingo Search with a Genetic Algorithm (HBFS-GA), and the proposed model are run on the ADE20K dataset [23, 24, 25]. The performance optimization of different algorithms was verified by comparing the trend of fitness changes with iteration times. Simultaneously selecting advertising images of different categories, the application effects of three methods were compared and analyzed.

3.1. Performance testing of feature models based on hierarchical features and SA optimization

To confirm the performance of the optimization model proposed by the research, a performance comparison test was conducted between the feature model based on hierarchical features and SA optimization, IHGS [5], and HBFS-GA [2]. For the study, the ADE20K dataset was chosen and split into training and testing (Tr-Te) sets at a ratio of 7:3. The limit for iterations count was set to 200, and the fitness of different algorithms varied with the number of iterations, as shown in Fig. 6. In subfigures 6a and b the fitness changes of three algorithm models on the Tr-Te sets are shown. In the Tr-Te sets, the model proposed by the research exhibited faster convergence speed and higher fitness

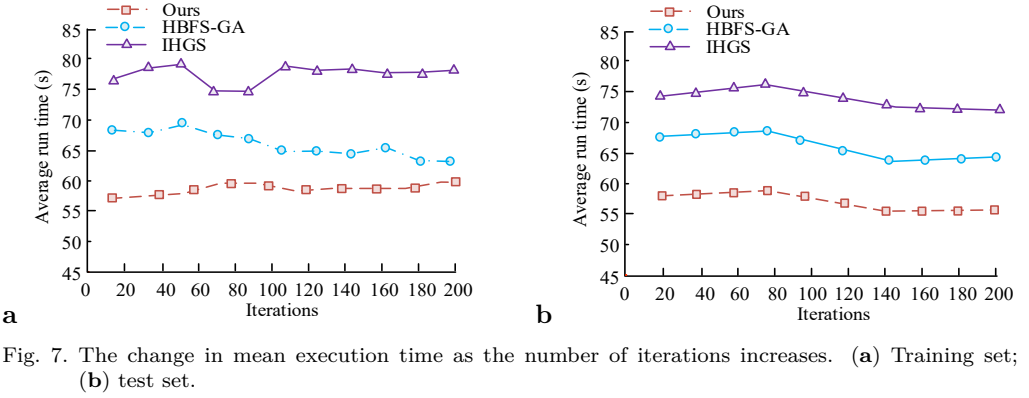
Tab. 1. Adaptation changes under different hyperparameter settings.

Algorithm	Hyperparameter Combination	Data Set	Mean Fitness	Standard Deviation	Coefficient of Variation (CV)
IHGS	Weight=0.5, particles=50	Training set	2.75	0.11	0.036
		Test set	2.73	0.12	0.044
HBFS-GA	Crossover rate=0.8, mutation rate=0.1	Training set	2.65	0.12	0.045
		Test set	2.61	0.13	0.054
Ours	$\alpha = 0.9, T_0 = 300, \lambda = 0.5$	Training set	3.06	0.02	0.006
	$\alpha = 0.7, T_0 = 500, \lambda = 0.7$	Test set	2.82	0.04	0.014

values than IHGS and HBFS-GA, demonstrating strong optimization ability and superior performance. In the test set, the fitness of this model reached 0.6, far exceeding the other two, demonstrating good generalization ability. In addition, the fitness curve of the model is smooth and stable, indicating stronger stability and global search ability, while the fluctuations of IHGS and HBFS-GA indicate instability in local optimal solution search. To evaluate the fitness changes of various algorithms under different hyperparameter settings, the influence of hyperparameter sensitivity on the target fitness value was tested, and the outcomes are summarized in Tab. 1. As it can be seen, the mean fitness of the IHGS Tr-Te sets was 2.75 and 2.73, respectively, with the result for the testing set slightly lower than that for the training set. The standard deviations of the Tr-Te sets were relatively small, with values of 0.11 and 0.12, and CV of 0.036 and 0.044, respectively, indicating small fluctuations in fitness and good stability. The mean fitness of HBFS-GA on the Tr-Te sets was 2.65 and 2.61. The standard deviations of the Tr-Te sets were 0.12 and 0.13, respectively, and the CV were 0.045 and 0.054, indicating slight fluctuations in fitness.

edremParagraph added.

Two sets of hyperparameters were set for the model proposed by the research. The mean fitness of the first set of hyperparameters on the Tr-Te sets was 3.06 and 2.95, significantly higher than that of IHGS and HBFS-GA. The standard deviations were 0.02 and 0.03, and the CV were 0.006 and 0.012, respectively, which were notably lower than the others, demonstrating that the model had the smallest variability and uncertainty



in fitness. In the second set of hyperparameters, the mean fitness of the Tr-Te sets were 2.85 and 2.82, which were also higher than IHGS and HBFS-GA, but lower than the first set of hyperparameters. In order to analyze the trade-off between optimization performance and time consumption of feature models based on hierarchical features and SA optimization, the running time of three algorithms when reaching the target fitness value was compared. The outcomes are shown in Fig. 7. The subfigures 7a and b, respectively, indicate the variation of the average running time (ART) of three algorithms on the Tr-Te sets with the number of iterations. In Fig. 7a, the ART of the proposed model remained the lowest, fluctuating between 55s and 60s. As the iteration numbers rose, the running time fluctuated less and gradually stabilized. The running time of HBFS-GA was within 65s – 70s, with small fluctuations. The running time of IHGS was the highest, with some fluctuations between 75s – 80s and poor stability. The running time of the model proposed by the research remained stable as the number of iterations increased, indicating that its SA computational complexity was relatively low. In Fig. 7b, compared to the performance system on the training set, the ART of the proposed model was the lowest, between 55s – 58s. The running time of HBFS-GA was within 65s – 68s, and the curve is relatively stable. The running time of IHGS was still the highest, between 75s – 80s, with a slight downward trend compared to the ART of the training set, but overall fluctuations were significant. Both in the Tr-Te sets, the ART of the proposed model was consistently significantly lower than that of HBFS-GA and IHGS. This indicated the time efficiency of introducing SA in optimizing advertising design feature tasks, especially showing good stability as the iteration numbers rose.

In terms of computational efficiency, the study tested three algorithms on four indicators: inference time, parameter count, memory usage, and floating-point operations. The outcomes are shown in Tab. 2. IHGS had relatively slow inference time, with inference times of 1.12s and 1.14s for the Tr-Te sets. The minimum number of parameters was

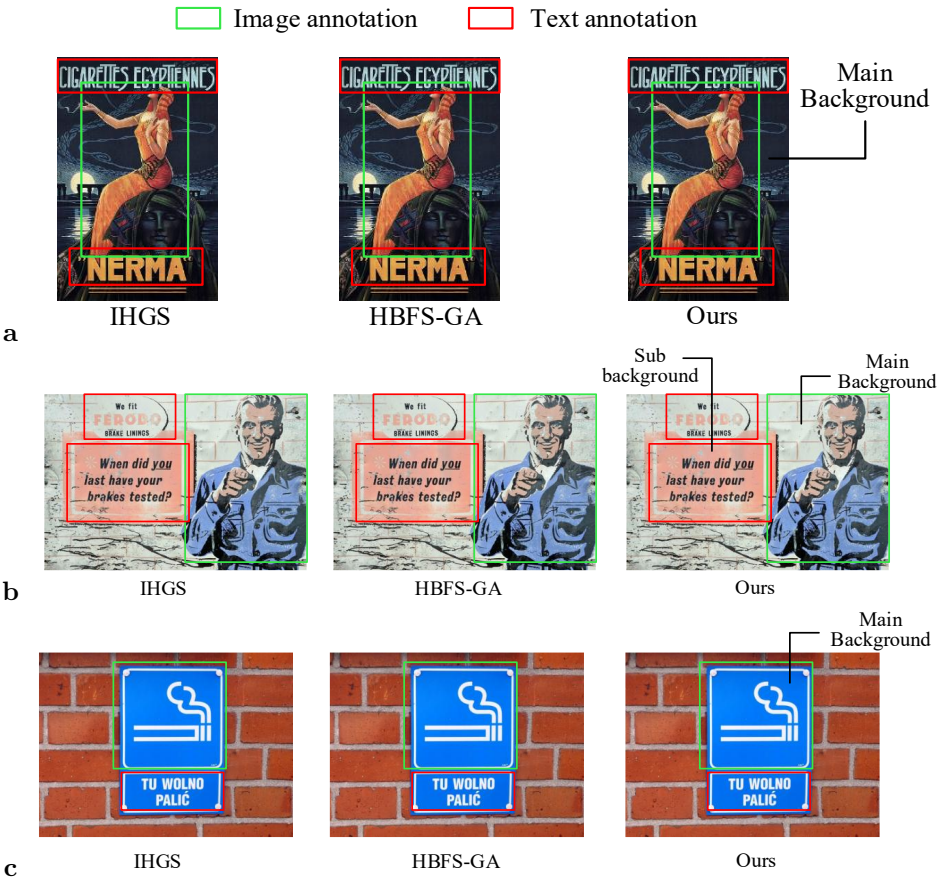
Tab. 2. Comparison of computational efficiency of different algorithms.

Algorithm	Data Set	Inference Time (s)	Parameter Count (Millions)	Memory Usage (MB)	Floating-Point Operations (GFLOPs)
IHGS	Training set	1.12	1.9	104	7.5
	Test set	1.14	2.1	112	7.4
HBFS-GA	Training set	1.07	2.6	124	8.8
	Test set	1.05	2.8	131	9.2
Ours	Training set	0.94	2.5	124	8.1
	Test set	0.92	2.4	125	8.4

1.9 M for the training set and 2.1 M for the testing set. Meanwhile, IHGS also had the lowest memory usage, with Tr-Te sets of 104 MB and 112 MB, respectively. Compared with IHGS, HBFS-GA improved inference time, with inference times of 1.07 s and 1.05 s for the Tr-Te sets, respectively. HBFS-GA had a large number of parameters on the Tr-Te sets, which were 2.6 M and 2.8 M, respectively. In terms of computational complexity, the floating-point operations of HBFS-GA significantly increased, with a training set of 8.8 GFLOPs and a testing set of 9.2 GFLOPs. Finally, the model proposed by the research performed the best in inference time, with inference times of 0.94 s and 0.92 s for the Tr-Te sets, making it the fastest among the three methods. The number of parameters was 2.5 M and 2.4 M, respectively, and the memory usage was similar to HBFS-GA. The Tr-Te sets were 124 MB and 125 MB. In terms of computational complexity, ours had lower floating-point operations than HBFS-GA but higher than IHGS, with a value for the training set of 8.1 GFLOPs and that for the testing set of 8.4 GFLOPs. The model raised by the research achieved a good balance between inference time, parameter size, and computational complexity, and had better computational efficiency.

3.2. Effect analysis based on hierarchical features and SA advertising image generation

To confirm the validity of hierarchical features and SA advertising image generation, this study selected three different types of advertising images from the ADE20K dataset: product promotion, brand promotion, and directive sign advertisement, and compared and analyzed the generation effects of different algorithms. In Fig. 8, the hierarchical feature extraction effects of different algorithms on elements in images were compared and analyzed. The red border represents the text annotation, and the green border



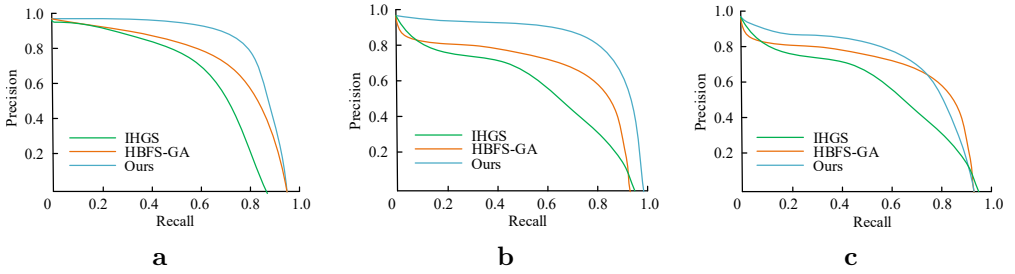


Fig. 9. Accuracy-recall results of data annotation. (a) Product promotion; (b) brand promotion; (c) directive sign advertisement.

to the generation effect of advertising images. From this, the model raised by the research exhibited significant advantages in advertising image generation tasks, especially in background feature extraction, precise annotation of copy and image subject, and feature integration ability, which were superior to IHGS and HBFS-GA.

Subsequently, the study employed Precision-Recall (PR) to evaluate the quality of data annotation, as shown in Fig. 9. The subfigures a, b and c show the PR curve results of different algorithms for product promotion, brand promotion, and directive sign advertisement, respectively. In Fig. 9a, the P-R curve of the proposed model is located at the top, indicating that the accuracy performed best throughout the entire recall range, especially maintaining high accuracy at high recall rates. In contrast, the curve of HBFS-GA shows a faster decrease in accuracy as the recall rate increased. The curve of IHGS is located at the bottom, and its accuracy rapidly decreased as the recall rate increased, resulting in the least ideal performance. Similarly, in Fig. 9b and c, the curve of the proposed model remains optimal, with significantly higher accuracy than other algorithms.

Finally, Mean Square Error (MSE), Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Learned Perceptual Image Patch Similarity (LPIPS) were used as evaluation metrics for image visual clarity. Among them, LPIPS is an important indicator for measuring the perceptual quality of images, which quantifies the perceptual differences between images through the feature space of deep networks. The smaller the value, the closer the generated image and the target image are perceived, which is more in line with human visual judgment. When LPIPS is 0, it indicates that the two images are perceived to be completely identical and almost indistinguishable by the human eye. The test results are gathered in Tab. 3 which shows the comparison of image generation effects of three algorithms on three different types. In product promotion, the MSE of the proposed model algorithm was the lowest, only 0.027, significantly better than IHGS and HBFS-G. The PSNR was the highest at 65.48, while IHGS was the lowest at only 60.45. In terms of SSIM, the proposed model reached 0.64, higher than the

Tab. 3. Comparative experiment outcomes.

Algorithm	Scene	MSE	PSNR	SSIM	LPIPS
IHGS	Product promotion	0.061	60.45	0.47	0.31
	Brand promotion	0.059	60.19	0.45	0.33
	Directive sign advertisement	0.057	61.42	0.46	0.32
HBFS-GA	Product promotion	0.054	63.97	0.51	0.25
	Brand promotion	0.053	62.57	0.54	0.26
	Directive sign advertisement	0.049	63.59	0.62	0.16
Ours	Product promotion	0.027	65.48	0.64	0.12
	Brand promotion	0.014	67.68	0.68	0.11
	Directive sign advertisement	0.025	66.97	0.67	0.10

0.51 of HBFS-GA and 0.47 of IHGS. In addition, the LPIPS performance of the model proposed by the research was the best, only 0.16, far lower than the 0.31 of IHGS. In brand promotion, the various indicators of the model proposed by the research showed significant advantages. Among them, MSE was the lowest, only 0.014, while IHGS was the highest, at 0.059. The PSNR was as high as 67.68, notably better than the others. In terms of SSIM, the model proposed by the research achieved 0.68, which was the best performance. LPIPS was the lowest at only 0.10, while IHGS was the highest at 0.33. In the directive sign advertisement, the model algorithm proposed by the research still performed the best, with an MSE of 0.025 and a PSNR of 66.97, both better than the other two algorithms. In terms of SSIM, the proposed model achieved 0.67, which was also significantly better than the other two algorithms, while LPIPS demonstrated excellent image generation quality with a minimum value of 0.10.

To evaluate the usability and deployment performance of the proposed advertising image generation model in real environments, the model is integrated into a prototype level advertising design assistance system to simulate the content production process of enterprises in actual marketing scenarios. Further research on the introduction of speeding up DDPM (SU-DDPM) [12] and multi-dimensional attention guided generative adversarial network (MDA-GAN) [3] two modern deep generative models were compared with feature models without SA and with SA, and the results are shown in Tab. 4, where, in the three types of actual advertising image application scenarios, the proposed model outperforms existing mainstream methods in terms of image quality and generation efficiency. Taking the LPIPS index as an example, the model achieved 0.12, 0.09, and 0.11 in three scenarios, all lower than SU-DDPM and MDA-GAN, indicating that it is closer to real images in terms of perceptual quality. In terms of SSIM, the model proposed by the research institute achieved 0.66, 0.68, and 0.67 in the three types of tasks, respectively, all higher than other methods, indicating a more complete preservation of

Tab. 4. Comparison of performance and efficiency of image generation models.

Application scenarios	Model	LPIPS	SSIM	Inference Time (s)
Social media	SU-DDPM	0.14	0.61	1.42
	MDA-GAN	0.17	0.59	0.88
	Ours w/o SA	0.18	0.57	0.79
	Ours	0.12	0.66	0.73
E-commerce Banner	SU-DDPM	0.13	0.62	1.45
	MDA-GAN	0.16	0.60	0.91
	Ours w/o SA	0.19	0.58	0.74
	Ours	0.09	0.68	0.76
Offline promotional poster	SU-DDPM	0.14	0.61	1.51
	MDA-GAN	0.17	0.58	0.92
	Ours w/o SA	0.19	0.58	0.75
	Ours	0.11	0.67	0.72

structural information. In terms of inference efficiency, although the inference time of the proposed model in the E-commerce Banner scenario is higher than that of the feature model without SA, it is generally lower than other methods in most cases. In addition, compared with the feature model without SA, the introduction of SA resulted in an average decrease of 0.07 in LPIPS and an average improvement of about 0.09 in SSIM, while the inference time showed almost no significant increase. This indicates that the SA optimization strategy maintains good computational efficiency while improving generation quality.

4. Discussion and conclusion

Aiming at the limitations of traditional advertising image generation methods, the research proposed an advertising design image generation method that combines hierarchical features and SA. By using a hierarchical feature model to represent advertising images in a hierarchical manner, gradually abstracting from low-level pixel features to high-level business features, comprehensively capturing multi-scale semantic information of advertising design, and utilizing SA to effectively avoid local optima through probabilistic random search in high-dimensional solution space, optimizing advertising design parameters. The experiment outcomes showed that the proposed model performed well in all indicators. On the test set, MSE, PSNR, SSIM, and LPIPS were 0.014, 67.68, 0.68, and 0.10, respectively, all of which were superior to the other two comparison methods, indicating that the structure of the generated image was closer to that of the target

image, reflecting the highest visual quality. In the context of optimizing efficiency, the raised model achieved a fitness value of 3.00 within 50 iterations on the training set, with the lowest standard deviation and CV of 0.02 and 0.006, respectively. In addition, the ART per iteration of the model was 55-60 seconds, which was significantly faster than the comparison algorithm, demonstrating its superior computational efficiency.

Overall, the model based on hierarchical features and SA optimization provided a robust and efficient framework for advertising image generation, demonstrating excellent performance in visual quality, computational efficiency, and adaptability. However, when facing industrial level application scenarios such as large-scale datasets or real-time generation tasks, current models still face certain challenges in terms of computational complexity and inference efficiency. Future research should further optimize algorithm structures or introduce technologies such as distributed computing and model compression to enhance system scalability and deployment flexibility. In addition, the model also needs to pay attention to potential ethical risks in practical applications. On the one hand, there may be aesthetic bias in the training data, which leads to the generation of images that excessively present certain cultural styles or gender stereotypes, affecting the diversity and inclusiveness of advertising content; On the other hand, generating images may visually exaggerate product effects or create misleading scenes, thereby harming consumer rights. Future research should strengthen the introduction of fairness mechanisms and control of content compliance while improving performance, ensuring comprehensive optimization of advertising image generation in terms of efficiency, scale, and ethics.

References

- [1] J. Du. Application of CAD aided intelligent technology in landscape design. *International Journal of Advanced Computer Science and Applications* 13(12):1030–1037, 2022. doi:10.14569/IJACSA.2022.01312118.
- [2] R. K. Eluri and N. Devarakonda. Feature selection with a binary flamingo search algorithm and a genetic algorithm. *Multimedia Tools and Applications* 82(17):26679–26730, 2023. doi:10.1007/s11042-023-15467-x.
- [3] B. Gu, X. Wang, W. Liu, and Y. Wang. MDA-GAN: Multi-dimensional attention guided concurrent-single-image-GAN. *Circuits, Systems, and Signal Processing* 44(2):1075–1102, 2025. doi:10.1007/s00034-024-02867-z.
- [4] R. Harada, K. O. Kim, and M. Takatera. Estimation of garment impression using regression models with design parameters and image features. *International Journal of Affective Engineering* 23(3):211–222, 2024. doi:10.5057/ijae.IJAE-D-23-00011.
- [5] E. H. Houssein, M. E. Hosney, W. M. Mohamed, A. A. Ali, and E. M. G. Younis. Fuzzy-based hunger games search algorithm for global optimization and feature selection using medical data. *Neural Computing and Applications* 35(7):5251–5275, 2023. doi:10.1007/s00521-022-07916-9.
- [6] P. Iyappan and P. Jamuna. Hybrid simulated annealing and Spotted Hyena Optimization

- Algorithm-based resource management and scheduling in cloud environment. *Wireless Personal Communications* 133(2):1123–1147, 2023. doi:[10.1007/s11277-023-10807-4](https://doi.org/10.1007/s11277-023-10807-4).
- [7] Y. Jiang. A novel DABU-Net model based on principle component analysis for intelligent collaborative robot design. *Journal of Applied Science and Engineering* 27(11):3533–3541, 2024. doi:[10.6180/jase.202411.27\(11\).0010](https://doi.org/10.6180/jase.202411.27(11).0010).
- [8] Y. Lin, H. Liu, H. Zhao, Q. Hu, X. Zhu, et al. Hierarchical feature selection based on label distribution learning. *IEEE Transactions on Knowledge and Data Engineering* 35(6):5964–5976, 2022. doi:[10.1109/TKDE.2022.3177246](https://doi.org/10.1109/TKDE.2022.3177246).
- [9] W. Liu, L. Yang, and B. Yu. Kernel density estimation based distributionally robust mean-CVaR portfolio optimization. *Journal of Global Optimization* 84(4):1053–1077, 2022. doi:[10.1007/s10898-022-01177-5](https://doi.org/10.1007/s10898-022-01177-5).
- [10] Y. Liu and M. Wu. Intelligent design of ethnic patterns in clothing using improved DCGAN for real-time style transfer. *International Journal of Advanced Computer Science and Applications* 14(11):1034–1044, 2023. doi:[10.14569/ijacsa.2023.01411105](https://doi.org/10.14569/ijacsa.2023.01411105).
- [11] M. Lu and Y. Xie. Intelligent detection system for electrical equipment based on deep learning and infrared image processing technology. *International Journal of Advanced Computer Science and Applications* 14(8):1147–1155, 2023. doi:[10.14569/IJACSA.2023.01408124](https://doi.org/10.14569/IJACSA.2023.01408124).
- [12] S. Lu, F. Guan, H. Zhang, and H. Lai. Speed-up DDPM for real-time underwater image enhancement. *IEEE Transactions on Circuits and Systems for Video Technology* 34(5):3576–3588, 2023. doi:[10.1109/TCSVT.2023.3314767](https://doi.org/10.1109/TCSVT.2023.3314767).
- [13] D. Manafi and M. J. Nategh. Optimization of setup planning by combined permutation-based and simulated annealing algorithms. *Arabian Journal for Science and Engineering* 48(3):3697–3708, 2023. doi:[10.1007/s13369-022-07209-2](https://doi.org/10.1007/s13369-022-07209-2).
- [14] M. Mashayekhi and H. Ghasemi. An enhanced simulated annealing algorithm for topology optimization of steel double-layer grid structures. *Advances in Computational Design* 9(2):115–136, 2024. doi:[10.12989/acd.2024.9.2.115](https://doi.org/10.12989/acd.2024.9.2.115).
- [15] J. Purohit and R. Dave. Leveraging deep learning techniques to obtain efficacious segmentation results. *Archives of Advanced Engineering Science* 1(1):11–26, 2023. doi:[10.47852/bonviewAAES32021220](https://doi.org/10.47852/bonviewAAES32021220).
- [16] M. Salehi, A. Bekker, and M. Arashi. A semi-parametric density estimation with application in clustering. *Journal of Classification* 40(1):52–78, 2023. doi:[10.1007/s00357-022-09425-9](https://doi.org/10.1007/s00357-022-09425-9).
- [17] T. Song, R. Wen, and L. Zhang. RoughSet-DDPM: An image super-resolution method based on rough set denoising diffusion probability model. *Tehnički vjesnik* 31(1):162–170, 2024. doi:[10.17559/TV-20230717000808](https://doi.org/10.17559/TV-20230717000808).
- [18] Y. Tian and Y. She. Uncertainty measure-based incremental feature selection for hierarchical classification. *International Journal of Fuzzy Systems* 26(6):2074–2096, 2024. doi:[10.1007/s40815-024-01708-0](https://doi.org/10.1007/s40815-024-01708-0).
- [19] T. Wang. Exploring intelligent image recognition technology of football robot using omnidirectional vision of internet of things. *The Journal of Supercomputing* 78(8):10501–10520, 2022. doi:[10.1007/s11227-022-04314-9](https://doi.org/10.1007/s11227-022-04314-9).
- [20] C. Xiao, S. Hong, and W. Huang. Optimizing graph layout by t-SNE perplexity estimation. *International Journal of Data Science and Analytics* 15(2):159–171, 2023. doi:[10.1007/s41060-022-00348-7](https://doi.org/10.1007/s41060-022-00348-7).
- [21] R. Yamasaki and T. Tanaka. Convergence analysis of mean shift. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 46(10):6688–6698, 2024. doi:[10.1109/TPAMI.2024.3385920](https://doi.org/10.1109/TPAMI.2024.3385920).

- [22] L. Zhang, M. Li, Y. Wang, B. Xing, X. Liu, et al. Emocolor: An assistant design method for emotional color matching based on semantics and images. *Color Research and Application* 48(3):312–327, 2023. doi:10.1002/col.22851.
- [23] B. Zhou, H. Zhao, X. Puig, S. Fidler, A. Barriuso, et al. Scene parsing through ADE20K dataset. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5122–5130, 2017. doi:10.1109/CVPR.2017.544.
- [24] B. Zhou, H. Zhao, X. Puig, S. Fidler, A. Barriuso, et al. ADE20K, 2025. <https://ade20k.csail.mit.edu/>. [Access: Dec 2024].
- [25] B. Zhou, H. Zhao, X. Puig, T. Xiao, S. Fidler, et al. Semantic understanding of scenes through the ADE20K dataset. *International Journal of Computer Vision* 127(3):302–321, 2019. doi:10.1007/s11263-018-1140-0.
- [26] spicetree687. Vintage poster print advertisement cigarettes. In: Free downloads – Photos, ID: 88671. <https://www.1001freedownloads.com/free-photo/vintage-poster-print-advertisement-cigarettes>.
- [27] Ghwtog. Advertising sign-plate advertise advertising sign. In: Free downloads – Photos, ID: 83116. <https://www.1001freedownloads.com/free-photo/advertising-sign-plate-advertise-advertising-sign>.
- [28] T. Mikołajczyk. Sign tablet an array of information smoking burn. In: Free downloads – Photos, ID: 82615. <https://www.1001freedownloads.com/free-photo/sign-tablet-an-array-of-information-smoking-burn>.



Jian Zhang born in 1981, from Xinxiang City, Henan Province, China. She graduated from the Fine Arts Department of Hubei Normal University in 2004 with a bachelor's degree in literature. She taught at Henan Light Industry Vocational College in 2004. Her main research areas include visual communication design and digital media technology.