Journal of Network Intelligence ©2026 ISSN 2414-8105
Taiwan Ubiquitous Information Volume **11**, Number **1**, February **2026**

AI Recognition of Painting Art Styles Based on Metaheuristic Optimization Machine Learning

Ling-Yun Ding[1], Bao-Cheng Wang[1,*]

[1]Harbin University of Commerce School of Design and Art,
Harbin 150076, P. R. China
dly1234561016@126.com, dzxyjjw@163.con

Jin-Hui Zhu[2]

[2]Payap University, Chiang Mai 50000, Thailand
sf6906@163.com

*Corresponding author: Bao-Cheng Wang

Received October 8, 2024, revised January 27, 2025, accepted April 12, 2025.

Abstract. *Traditional art style recognition methods often rely on single-modal feature extraction or fixed algorithmic architectures, which limits the accurate recognition and analysis of diverse and emerging art styles. With the rapid development of digital art and the advancement of cultural heritage digitization, there is an increasing demand for algorithms that can adapt to complex art styles, handle large-scale art datasets and perform accurate classification. To address these two issues, this work proposes an art style recognition framework (DAMAF) based on dynamic adaptive meta-heuristic algorithm and multimodal feature fusion. First, through the adaptive dynamic weight fusion (ADWF) strategy, DAMAF is able to flexibly integrate multi-dimensional information such as visual, texture, color and composition, ensuring the adaptability to diverse art styles. Secondly, through the Progressive Domain Adaptive Migration (PDAT) technique, DAMAF is able to effectively deal with emerging and rare art styles, which improves the generalization ability of the model. Experimental results show that on the ArtStyleNet test set, the macro-averaged F1 score (MF1) of DAMAF reaches 0.906, which is 5.5% higher than the best benchmark model; on the EmergingArt dataset, the MF1 of DAMAF reaches 0.784, which is 15.3% higher than the benchmark model. In particular, when processing GAN artworks, DAMAF's MF1 reaches 0.885, which is 22.3% higher than the benchmark model, demonstrating its excellent performance in recognizing emerging art styles.*

Keywords: art style recognition; dynamic adaptive algorithm; multimodal feature fusion; migration learning; emerging art styles.

1. **Introduction.** AI recognition techniques for painting art styles have shown promising applications in the fields of digital humanities, art education, and cultural heritage preservation. With the acceleration of global art digitization, the rise of large-scale online art collections and virtual art museums, automated tools for analyzing and classifying art styles have become increasingly important [1, 2]. This technology can not only assist art historians in researching and classifying large-scale art works, but also provide personalized art appreciation guidance for general audiences and promote the popularization and deepening of art education.

In the field of art market and cultural heritage protection, accurate art style recognition technology can assist in identifying the authenticity of artworks, aid in tracking down stolen or lost artworks, and provide an important reference for cultural relic restoration [3]. In addition, in the creative industry, this technology can provide inspiration for designers and artists, assist in generating specific styles of artworks [4], and may even give rise to new forms of artistic creation [5]. With the continuous progress of technology, painting art style AI recognition is expected to play an increasingly important role in all aspects of art creation, appreciation and protection, and promote the deep integration of art and technology.

1.1. **Related work.** Painting art styles Research in the field of AI recognition has experienced an evolution from traditional methods to deep learning techniques. Early research mainly relied on hand-designed feature extraction methods. Condorovici et al. [6] proposed an art style classification method based on color histograms and texture features, which achieved some success on a specific dataset but had limited ability to generalize to complex and variable art styles. Lee et al. [7] developed a recognition algorithm that combines shape descriptors and local features, which performs well in dealing with art styles with obvious geometric features [8], but is not effective in recognizing abstract or modern artworks.

With the development of deep learning technology, convolutional neural network (CNN) has become a mainstream method for AI recognition of painting art styles. Hua et al. [9] designed a multi-scale CNN architecture to recognize art styles by extracting different levels of visual features, which significantly improved the recognition accuracy. However, this method still has difficulties in dealing with works with fuzzy style boundaries. Zhang et al. [10] proposed a deep learning model based on the attention mechanism, which can automatically focus on key regions in paintings and improve the ability to capture detailed features, but the interpretability of the model is still insufficient. Zhao et al. [11] developed an art style recognition method based on graph neural network, which improves the recognition performance by modeling the relationship between the elements of paintings, but this method still needs to be optimized in terms of computational complexity and training efficiency.

Multimodal feature fusion techniques show great potential in the field of painting art style AI recognition. Tashu et al. [12] proposed a multimodal fusion model that combines visual and textual information to enhance style recognition by integrating images of paintings and related art comments. Although this method improves the recognition accuracy in some scenarios, it is less effective in dealing with emerging artworks that lack textual descriptions. Liu et al. [13] designed a multimodal feature extraction framework that incorporates color, texture, and compositional information, but this method still has challenges in balancing the weights of different modal features, especially for artworks with uneven stylistic features.

1.2. **Motivation and contribution.** Existing AI recognition methods for painting art styles usually rely on fixed algorithmic architectures and single-modal feature extraction, which are difficult to effectively deal with the diversity and complexity of art styles. Works from different art genres and periods have significant differences in visual features, brushstroke techniques, and compositional approaches, making the model perform poorly in recognizing emerging or rare art styles. In addition, artworks often contain rich multimodal information, and traditional recognition methods easily lead to the loss of key features or improper weight allocation when integrating this complex information. To address the above problems, this paper proposes a multimodal feature fusion framework based on dynamic adaptive meta-heuristic algorithms, which significantly improves the effectiveness of AI recognition of painting art styles. The main innovations and contributions of this work include:

(1) Aiming at the challenge of the diversity and complexity of art styles, the dynamic adaptive meta-heuristic algorithm framework proposed in this paper is able to automatically adjust the optimization strategy according to the characteristics of different art styles, thus improving the adaptability and generalization ability of the model in recognizing various art styles. This improvement is especially significant when dealing with emerging or rare art styles, which effectively solves the problem of Tashu et al. [12] that the model is ineffective when dealing with emerging art works that lack textual descriptions.

(2) Aiming at the complexity of multimodal features of artworks, this paper proposes an improved multimodal feature fusion and transfer learning method. By integrating multidimensional information such as vision, texture, color and composition, and combining with transfer learning techniques, this method can capture and integrate the multimodal features of artworks more effectively. This innovation effectively overcomes the challenge of Liu et al. [13] in balancing the weights of different modal features, especially for artworks with uneven stylistic features, and significantly improves the accuracy and robustness of recognition.

2. A framework for dynamic adaptive metaheuristic algorithms.

2.1. Framework overview. The Dynamic Adaptive Meta-heuristic Algorithm Framework (DAMAF) proposed in this section aims to address the complexity and diversity challenges in AI recognition of painting art styles. DAMAF achieves this by intelligently selecting and combining multiple meta-heuristic algorithms (Genetic Algorithm [14], Particle Swarm Optimization [15], Ant Colony Algorithm [16] and Artificial Bee Colony [17]) to achieve adaptive optimization of different art style features, as shown in Figure 1.

The core of the framework is an adaptive controller $\mathcal{C}$, which receives as input the feature vector $\mathbf{x} \in \mathbb{R}^n$ of the artwork and outputs the optimal algorithmic configuration $\mathbf{y} \in \mathbb{R}^m$. The feature vector $\mathbf{x}$ contains multidimensional information about the visual, texture, color and composition of the artwork. The algorithm configuration $\mathbf{y}$ defines the selection, parameter setting and combination strategy of the metaheuristic algorithm.

The workflow of DAMAF can be represented as follows:

$$\mathbf{y} = \mathcal{C}(\mathbf{x}, \boldsymbol{\theta}) \tag{1}$$

where $\boldsymbol{\theta}$ represents the set of parameters of the controller, which is continuously tuned through iterative optimization. The objective function $\mathcal{L}$ of the framework measures the performance of art style recognition and can be expressed as:

$$\mathcal{L} = f(\mathbf{y}, \mathbf{x}) \tag{2}$$

DAMAF optimizes the controller parameters $\boldsymbol{\theta}$ by minimizing $\mathcal{L}$ for adaptive recognition of different art styles. The adaptive controller $\mathcal{C}$ generates an algorithmic configuration $\mathbf{y}$ based on the input features $\mathbf{x}$ and the current parameters $\boldsymbol{\theta}$. The system executes the selected meta-heuristic algorithm and outputs the art style recognition result. The performance evaluation module computes the objective function $\mathcal{L}$ to evaluate the recognition performance. Based on the performance evaluation results, the system passes the information back to the adaptive controller via the long feedback arrow for optimizing the parameters $\boldsymbol{\theta}$.

2.2. Adaptive algorithm selection mechanism. The adaptive algorithm selection mechanism of DAMAF is based on the principle of Reinforcement Learning (RL) [18], which models the art style recognition problem as a Markov Decision Process (MDP).

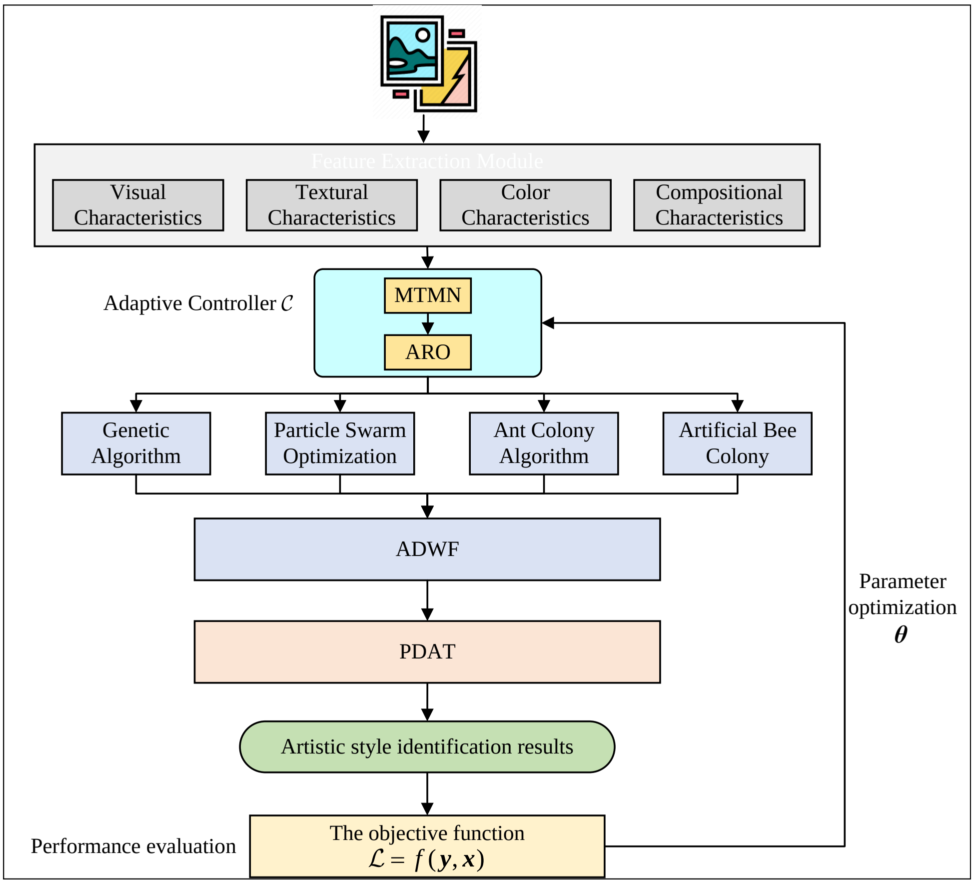

Figure 1. DAMAF

Define the state space $\mathcal{S}$ as the artwork feature space and the action space $\mathcal{A}$ as the set of optional meta-heuristics. At each time step t, the system selects the action $a_t \in \mathcal{A}$, i.e., a specific meta-heuristic algorithm, based on the current state $s_t \in \mathcal{S}$.

Introduce the strategy function $\pi_\phi : \mathcal{S} \to \mathcal{A}$ parameterized by ϕ. The strategy function is optimized by maximizing the expected cumulative reward $J(\phi)$:

$$J(\phi) = \mathbb{E}_{\tau \sim \pi_\phi} \left[\sum_{t=0}^{T} \gamma^t r_t \right] \tag{3}$$

where τ denotes the trajectory, $\gamma \in (0, 1]$ is the discount factor, and r_t is the instantaneous reward, defined as the increment in recognition accuracy.

To accommodate the dynamic nature of art styles, a strategy gradient approach is used to update ϕ:

$$\nabla_\phi J(\phi) = \mathbb{E}_{\tau \sim \pi_\phi} \left[\sum_{t=0}^{T} \nabla_\phi \log \pi_\phi(a_t|s_t) Q^{\pi_\phi}(s_t, a_t) \right] \tag{4}$$

where ∇_ϕ denotes the gradient with respect to the parameter ϕ; $\mathbb{E}_{\tau \sim \pi_\phi}$ denotes the expectation to be taken over the trajectory τ under the strategy π_ϕ; T denotes the total number of time steps of the trajectory; $\pi_\phi(a_t|s_t)$ denotes the probability of choosing action a_t in state s_t; $Q^{\pi_\phi}(s_t, a_t)$ denotes the action value function that represents the expected cumulative reward for executing action a_t from state s_t under strategy π_ϕ.

The action value function $Q^{\pi_\phi}(s_t, a_t)$ is defined as:

$$Q^{\pi_\phi}(s_t, a_t) = \mathbb{E}_{\pi_\phi}\left[\sum_{k=t}^{T} \gamma^{k-t} r_k \mid s_t, a_t\right] \tag{5}$$

where γ is the previously defined discount factor and r_k is the immediate reward obtained at time step k.

Through this adaptive mechanism, DAMAF is able to dynamically select the most suitable meta-heuristic algorithms according to the characteristics of different art styles, significantly improving the recognition performance.

2.3. **Parameter dynamic adjustment strategy.** The core strength of the DAMAF framework lies in its parameter dynamic tuning capability, which allows the system to adapt to changing art style characteristics. This section introduces a novel parameter dynamic tuning strategy called Adaptive Resonance Optimization (ARO).

The core idea of ARO is to consider the parameter tuning process as a dynamic system in which the parameter changes follow a law similar to that of resonance. Define the parameter vector $\boldsymbol{\omega} \in \mathbb{R}^d$, whose dynamics can be expressed as:

$$\frac{d^2\boldsymbol{\omega}}{dt^2} + \beta\frac{d\boldsymbol{\omega}}{dt} + \kappa\boldsymbol{\omega} = \mathbf{F}(\mathbf{x}, \mathbf{y}) \tag{6}$$

where β is the damping coefficient; κ is the elasticity coefficient; and $\mathbf{F}(\mathbf{x}, \mathbf{y})$ is the driving force, which is determined by the input features $\mathbf{x}$ and the algorithmic configuration $\mathbf{y}$ together.

The design of the driving force $\mathbf{F}(\mathbf{x}, \mathbf{y})$ is crucial, as it reflects the gap between the current and optimal parameters. We propose a hybrid driving force based on gradient information and historical performance:

$$\mathbf{F}(\mathbf{x}, \mathbf{y}) = \alpha\nabla_{\boldsymbol{\omega}}\mathcal{L} + (1-\alpha)\mathbf{H}(\boldsymbol{\omega}) \tag{7}$$

where $\alpha \in [0, 1]$ is the equilibrium factor; $\nabla_{\boldsymbol{\omega}}\mathcal{L}$ is the gradient of the objective function with respect to the parameter; and $\mathbf{H}(\boldsymbol{\omega})$ is the historical performance function capturing the parameter's past performance in the Trends.

By solving this dynamic equation, ARO is able to perform an adaptive search in the parameter space, effectively avoiding local optimality traps while maintaining the ability to explore globally optimal solutions.

2.4. **Feedback learning mechanism.** To further enhance the adaptive capability of DAMAF, we introduce an innovative feedback learning mechanism called Multi-scale Temporal Memory Network (MTMN). MTMN provides richer information for parameter tuning by capturing performance variations at different time scales.

The MTMN consists of K parallel memory modules, each corresponding to a specific time scale. The state of the k-th module is defined as $\mathbf{h}_k \in \mathbb{R}^{m_k}$, and its update rule is:

$$\mathbf{h}_k^{(t)} = f_k(\mathbf{h}_k^{(t-1)}, \mathcal{L}^{(t)}, \boldsymbol{\omega}^{(t)}) \tag{8}$$

where $f_k(\cdot)$ is the nonlinear mapping function; and $\mathcal{L}^{(t)}$ and $\boldsymbol{\omega}^{(t)}$ are the value of the objective function and the parameter vector for the current time step, respectively.

The output $\mathbf{z}$ of MTMN is obtained by aggregating the states of the individual modules:

$$\mathbf{z} = g(\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_K) \tag{9}$$

where $g(\cdot)$ is the aggregation function, which can be a weighted sum or a more complex nonlinear mapping.

The output $\mathbf{z}$ of MTMN is used to adjust the drive in ARO:

$$\mathbf{F}(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \alpha \nabla_{\boldsymbol{\omega}} \mathcal{L} + (1 - \alpha)\mathbf{H}(\boldsymbol{\omega}) + \lambda \mathbf{G}(\mathbf{z}) \tag{10}$$

where λ is the influence factor of the MTMN output; and $\mathbf{G}(\mathbf{z})$ is the mapping function that transforms the MTMN output into drivers.

By integrating ARO and MTMN, DAMAF achieves the ability to sense and respond to changes in performance at different time scales, dramatically improving the framework's adaptability and robustness when dealing with complex and changing art styles.

3. Multimodal feature fusion and transfer learning.

3.1. **Multimodal feature extraction.** In order to comprehensively capture the rich information of painting art styles, this section proposes an innovative multimodal feature extraction framework called Comprehensive Art Feature Extraction Network (CAFEN). CAFEN provides a comprehensive and in-depth feature representation for the subsequent style recognition task by integrating multidimensional information such as visual, brushstroke texture [19], color distribution, and compositional structure [20] and other multi-dimensional information, CAFEN provides a comprehensive and in-depth feature representation for subsequent style recognition tasks.

The core idea of CAFEN is to consider an artwork I as a combination of multiple information channels. Define the feature extraction function $\Phi(I)$, whose output is a multidimensional feature vector:

$$\Phi(I) = [\phi_v(I), \phi_b(I), \phi_c(I), \phi_s(I)] \tag{11}$$

where ϕ_v, ϕ_b, ϕ_c and ϕ_s represent visual, stroke, color and compositional feature extraction subfunctions, respectively.

The visual feature extraction sub-function ϕ_v adopts the structure of Modified Deep Residual Network (MDRN). The MDRN enhances the ability to capture the key visual elements of the art style through the introduction of an attention mechanism and multi-scale feature fusion. The output of MDRN is defined as:

$$\phi_v(I) = F_{\text{MDRN}}(I; \theta_v) \tag{12}$$

where F_{MDRN} denotes the forward propagation process of the MDRN; θ_v is a network parameter.

The stroke texture analysis sub-function ϕ_b uses the Adaptive Wavelet Transform (AWT) technique [21]. AWT can automatically adjust the wavelet basis function to accommodate the unique stroke characteristics of different artists. The stroke characteristics can be expressed as:

$$\phi_b(I) = W_{\text{AWT}}(I; \psi) \tag{13}$$

where W_{AWT} denotes the AWT operation; ψ is the adaptive wavelet basis function.

The color distribution feature subfunction ϕ_c utilizes the Dynamic Color Histogram (DCH) method [22]. DCH captures the color characteristics of artworks more accurately through an adaptive binning strategy:

$$\phi_c(I) = H_{\text{DCH}}(I; \beta) \tag{14}$$

where H_{DCH} denotes the DCH operation; β is the vector of split-box parameters.

The compositional structure analysis subfunction ϕ_s adopts the innovative Hierarchical Graph Attention Network (HGAT) technique [23]. HGAT effectively captures the overall compositional features and local elemental relationships of an artwork by transforming the artwork into a multilevel graph structure.

First, we convert the input image I into a hierarchical map structure:

$$\mathcal{G}(I) = (\mathcal{V}, \mathcal{E}, \mathcal{A}) \tag{15}$$

where $\mathcal{V} = \{V_1, V_2, \ldots, V_L\}$, $V_l = \{v_1^l, v_2^l, \ldots, v_{N_l}^l\}$ denotes the set of nodes in the l-th layer; $\mathcal{E} = \{E_1, E_2, \ldots, E_L\}$, E_l denotes the set of edges inside the l-th layer; $\mathcal{A} = \{A_1, A_2, \ldots, A_{L-1}\}$, A_l denotes the cross-layer connection between the l-th and $l+1$-th layers.

The core operations of HGAT include intra-layer feature aggregation and inter-layer message passing:

(1) Intra-layer feature aggregation. For node v_i^l in the l-th layer, the feature update process is:

$$h_i^{l,(t+1)} = \sigma\left(\sum_{j\in\mathcal{N}_i^l} \alpha_{ij}^l W^l h_j^{l,(t)}\right) \tag{16}$$

where α_{ij}^l is the attention coefficient of node i to node j in the l-th layer.

$$\alpha_{ij}^l = \frac{\exp(LeakyReLU(a^{l,T}[W^l h_i^{l,(t)} \| W^l h_j^{l,(t)}]))}{\sum_{k\in\mathcal{N}_i^l} \exp(LeakyReLU(a^{l,T}[W^l h_i^{l,(t)} \| W^l h_k^{l,(t)}]))} \tag{17}$$

where l denotes the layer of the graph; t denotes the current iteration step; $h_i^{l,(t)}$ denotes the feature vector of the node i in the l layer at the t iteration step; $W^l \in \mathbb{R}^{F'\times F}$ denotes the linear transformation matrix of the l layer, where F is the input feature dimension and F' is the output feature dimension; $a^l \in \mathbb{R}^{2F'}$ denotes the learnable attention vector of the l layer; $[\ \|\ ]$ denotes the vector splicing operation; and $LeakyReLU$ denotes the rectified linear unit activation function, defined as $LeakyReLU(x) = \max(0.01x, x)$; $\mathcal{N}_i^l$ denotes the set of neighbors of node i in the l-th layer.

(2) Inter-layer messaging. The node characteristics of the l-th layer affect the $l+1$-th layer nodes through cross-layer connections:

$$\tilde{h}_j^{l+1,(t)} = \sigma\left(\sum_{i\in\mathcal{M}_j^l} \beta_{ji}^l U^l h_i^{l,(t)}\right) \tag{18}$$

where $\mathcal{M}_j^l$ denotes the set of nodes in the l-th layer connected to the $l+1$-th layer node j, and β_{ji}^l is the cross-layer attention coefficient.

(3) Feature fusion. Combining intra-layer aggregation and inter-layer transfer of information:

$$h_j^{l+1,(t+1)} = \gamma \cdot h_j^{l+1,(t)} + (1-\gamma) \cdot \tilde{h}_j^{l+1,(t)} \tag{19}$$

where γ is the learnable fusion parameter.

After several rounds of iterations, we obtain the final node features. Then, graph level representation is obtained by hierarchical pooling operation:

$$h_G = POOL(\{h_i^L \mid i \in V_L\}) \tag{20}$$

where $POOL$ uses an adaptive pooling strategy to dynamically adjust the weights according to the node importance.

Finally, the compositional feature vectors are generated through the fully connected layer:

$$\phi_s(I) = W_{out} h_G + b_{out} \tag{21}$$

HGAT has the advantage: (1) The hierarchical structure helps to capture the multi-scale compositional characteristics of the artwork. (2) Attention mechanisms are able to adaptively focus on key compositional elements and their relationships. (3) Inter-layer information transfer realizes effective fusion of local and global information.

With HGAT, $\phi_s(I)$ is able to comprehensively and accurately characterize the compositional features of an artwork, providing rich structured information for the subsequent style recognition task. This approach significantly improves the analytical capability of the DAMAF framework in dealing with complex art styles, especially for those art genres where compositional features play a key role in style judgment.

Finally, CAFEN realizes the comprehensive extraction of multimodal features of artworks by integrating these four subfunctions. This approach not only captures the salient features of art styles, but also reveals the potential subtle differences, which lays a solid foundation for subsequent style recognition and analysis. CAFEN's multimodal feature extraction capability significantly improves the adaptability and accuracy of the DAMAF framework in dealing with diversified art styles, and effectively solves the limitations of the traditional single-modal approach in the recognition of complex art styles.

3.2. **Feature fusion strategy.** In order to fully utilize the advantages of multimodal feature extraction, a novel Adaptive Dynamic Weight Fusion (ADWF) strategy is proposed. ADWF aims to dynamically adjust the importance of different feature modalities to adapt to diverse art style characteristics.

ADWF introduces a Meta-learner that automatically adjusts the fusion weights based on the input artwork features. The fusion function Ψ is defined as follows:

$$\Psi(\Phi(I)) = \sum_{m=1}^{M} w_m(I) \cdot \phi_m(I) \tag{22}$$

where M is the number of eigenmodes, $w_m(I)$ is the fusion weight of the m-th mode, and $\phi_m(I)$ is the corresponding eigenvector. The weights $w_m(I)$ are generated by the meta-learner Λ:

$$[w_1(I), \dots, w_M(I)] = \Lambda(\Phi(I); \theta_\Lambda) \tag{23}$$

where θ_Λ is the parameter of the meta-learner. In order to enhance the expressive power of ADWF, we design the meta-learner using the multi-head attention mechanism:

$$\Lambda(\Phi(I); \theta_\Lambda) = \text{softmax}\left(\frac{1}{K}\sum_{k=1}^{K} Q_k W_k^Q (K_k W_k^K)^T\right) \tag{24}$$

where K is the number of attention heads; Q_k and K_k are the query and key matrices, respectively, obtained from $\Phi(I)$ by linear transformation; and W_k^Q and W_k^K are the learnable parameter matrices.

Through this self-adaptive mechanism, ADWF is able to flexibly adjust the feature fusion strategy according to the characteristics of different artworks, which effectively improves the adaptability and robustness of the DAMAF framework in dealing with diverse art styles.

3.3. **Application of migration learning techniques.** To enhance the generalization capability and efficiency of the DAMAF framework, an innovative Progressive Domain Adaptive Transfer (PDAT) technique is introduced. PDAT is designed to address the problem of domain bias in the recognition of art styles, especially for the challenge of recognizing emerging or rare art styles.

3.3.1. *Pre-training model selection.* The foundation of the PDAT technique is the selection of a robust and flexible pre-training model. In this study, Modified Vision Transformer (MVT) [24, 25] is used as the infrastructure to provide a solid foundation for the art style recognition task with its excellent long-range dependency capturing capability.

Define the input image as $I \in \mathbb{R}^{H\times W\times C}$, where H, W, and C denote the height, width, and the number of channels, respectively. The MVT first divides the input image into N non-overlapping blocks of the image, and each block undergoes linear projection to obtain the initial embedding $Z^0 \in \mathbb{R}^{N\times D}$, where D is the embedding dimension.

MVT consists of L converter blocks and the operation of each block can be expressed as:

$$Z^l = MSA(LN(Z^{l-1})) + Z^{l-1} \tag{25}$$

$$Z^l = MLP(LN(Z^l)) + Z^l \tag{26}$$

where $Z^l \in \mathbb{R}^{N\times D}$ denotes the output of the l-th layer; $MSA : \mathbb{R}^{N\times D} \to \mathbb{R}^{N\times D}$ denotes the multi-head self-attention operation; $LN : \mathbb{R}^{N\times D} \to \mathbb{R}^{N\times D}$ denotes a layer normalization; $MLP : \mathbb{R}^{N\times D} \to \mathbb{R}^{N\times D}$ denotes a Multilayer Perceptron.

To enhance the sensitivity of MVT to artistic features, we introduce the Feature Enhancement Module (FEM). FEM is applied after each converter block and its operation is defined as:

$$Z^l = FEM(Z^l) + Z^l \tag{27}$$

$$FEM(Z) = \sigma(W_2\sigma(W_1 Z + b_1) + b_2) \odot Z \tag{28}$$

where $W_1 \in \mathbb{R}^{D'\times D}$ and $W_2 \in \mathbb{R}^{D\times D'}$ are the learnable weight matrices; $b_1 \in \mathbb{R}^{D'}$ and $b_2 \in \mathbb{R}^{D}$ are bias vectors; D' is the intermediate feature dimension of the FEM; $\sigma : \mathbb{R} \to \mathbb{R}$ is a nonlinear activation function, e.g., GELU; and $\odot$ denotes the Hadamard product (element-by-element multiplication).

FEM enhances the model's ability to perceive art style features by adaptively adjusting feature importance. Finally, the output features Z^L of the MVT are globally average pooled and linearly layered to obtain the characterization vector $f \in \mathbb{R}^M$ of the art style, where M is the number of style categories.

The pre-training process is optimized using the cross-entropy loss $\mathcal{L}_{\mathrm{CE}}$:

$$\mathcal{L}_{\mathrm{CE}} = -\sum_{i=1}^{M} y_i \log(\hat{y}_i) \tag{29}$$

where y_i and $\hat{y}_i$ are the true label and predicted probability, respectively.

This improved MVT structure not only retains the advantages of the original visual converter, but also enhances the recognition of art features through FEM. The pre-trained model $\mathcal{M}_b$ is trained on a large-scale art dataset, which lays a solid foundation for the subsequent domain adaptation process, and effectively improves the initial performance of the DAMAF framework when dealing with diverse art styles.

3.3.2. *Domain adaptation method.* In order to effectively mitigate the distributional differences between the source domain (pre-training data) and the target domain (new art styles), this study proposes an innovative progressive adversarial domain adaptation strategy. By dynamically adjusting the intensity of adversarial learning, this strategy realizes a smooth migration from the source domain to the target domain, which significantly improves the recognition ability of the DAMAF framework for emerging and rare art styles.

First, define the source domain data distribution as p_s and the target domain data distribution as p_t. Let the feature extractor be $\mathcal{F} : \mathcal{X} \to \mathcal{Z}$, where $\mathcal{X}$ is the input space and $\mathcal{Z}$ is the feature space. The domain discriminator is defined as $\mathcal{D} : \mathcal{Z} \to [0, 1]$, which is used to distinguish whether the features come from the source or target domain.

The objective function of adversarial learning $\mathcal{L}_{\text{adv}}$ is defined as follows:

$$\mathcal{L}_{\text{adv}}(\mathcal{F}, \mathcal{D}) = \mathbb{E}_{x_s \sim p_s}[\log \mathcal{D}(\mathcal{F}(x_s))] + \mathbb{E}_{x_t \sim p_t}[\log(1 - \mathcal{D}(\mathcal{F}(x_t)))] \tag{30}$$

where x_s and x_t denote the samples in the source and target domains, respectively.

To achieve progressive adaptation, a dynamic adjustment factor $\lambda(t)$ is introduced:

$$\lambda(t) = 2/(1 + \exp(-\gamma t)) - 1 \tag{31}$$

where t denotes the current number of training iterations and $\gamma > 0$ is a tuning parameter that controls the adaptation rate.

Combining the task-specific loss $\mathcal{L}_{\text{task}}$ and the adversarial loss, the final optimization objective is:

$$\min_{\mathcal{F}} \max_{\mathcal{D}} \mathcal{L}_{\text{task}} + \lambda(t)\mathcal{L}_{\text{adv}} \tag{32}$$

To further enhance the effect of domain adaptation, the feature alignment regularization term $\mathcal{L}_{\text{align}}$ is introduced in this study:

$$\mathcal{L}_{\text{align}} = \|\mu_s - \mu_t\|_2^2 + \|\Sigma_s - \Sigma_t\|_F^2 \tag{33}$$

where μ_s, μ_t are the mean values of the source and target domain features respectively; Σ_s, Σ_t are the corresponding covariance matrices; and $\| \cdot \|_F$ denotes the Frobenius norm.

The final optimization problem can be formulated as follows:

$$\min_{\mathcal{F}} \max_{\mathcal{D}} \mathcal{L}_{\text{task}} + \lambda(t)\mathcal{L}_{\text{adv}} + \beta\mathcal{L}_{\text{align}} \tag{34}$$

where $\beta > 0$ is the weight coefficient of the feature alignment regularization term.

This progressive adversarial domain adaptation method, combined with feature alignment regularization and gradient inversion techniques, not only effectively mitigates the distributional differences between the source and target domains, but also improves the stability and efficiency of training. By dynamically adjusting the strength of adversarial learning, the DAMAF framework is able to more smoothly migrate the knowledge of pre-trained models to new art style domains, significantly enhancing the ability to recognize emerging and rare art styles while maintaining high performance on known styles.

DAMAF is the overall framework proposed in this study, which contains several innovative modules and techniques. ADWF and PDAT are the two core components in DAMAF. Specifically: the ADWF is the key component responsible for multimodal feature fusion in DAMAF. the ADWF strategy allows the framework to dynamically adjust the fusion weights of different modalities (e.g., visual, texture, color, composition, etc.) according to the characteristics of the input artworks. It is one of the important mechanisms to enable DAMAF to adapt to diverse art styles. PDAT is a migration learning technique in DAMAF that aims to improve the model's ability to recognize emerging and rare art styles. PDAT improves the model's generalization ability by gradually adjusting the model's parameters so that it can smoothly migrate from the source domain (pre-training data) to the target domain (new art styles).

4. **Experimentation and evaluation.**

4.1. **Experimental setup.** In order to comprehensively evaluate the performance of the Dynamic Adaptive Multimodal Art Style Recognition Framework (DAMAF), a series of rigorous and innovative experiments are designed in this study. The experimental environment utilizes a high-performance computing cluster equipped with NVIDIA A100 GPUs and Intel Xeon dual processors to ensure efficient training and inference of complex models.

For the dataset, a large-scale and diversified artwork dataset, ArtStyleNet, is constructed in this study. The dataset contains 5,000 high-resolution art images covering 100 different art styles and spanning from Renaissance to contemporary art. To simulate the imbalance of data distribution in real scenarios, ArtStyleNet adopts a long-tailed distribution design, with sufficient samples of mainstream art styles (e.g., Impressionism, Abstract Expressionism, etc.) and fewer samples of emerging or rare styles (e.g., Digital Art, Bio Art, etc.). The dataset is divided into a training set $\mathcal{D}_{\text{train}}$, a validation set $\mathcal{D}_{\text{val}}$, and a test set $\mathcal{D}_{\text{test}}$ in the ratio of 7:1:2.

To evaluate the generalization ability and domain adaptation performance of DAMAF, we additionally collected a test set EmergingArt containing 1,000 emerging artworks, denoted as $\mathcal{D}_{\text{emerging}}$. The styles of these works are not present or have very small samples in the training set and include, but are not limited to, emerging genres such as generative art, VR art, and eco-art.

The model implementation uses the PyTorch 1.9 deep learning framework combined with the NVIDIA APEX library for mixed-precision training to improve computational efficiency. Hyperparameter optimization is implemented using Bayesian optimization method using BoTorch library. The main hyperparameters include the learning rate $\eta \in [10^{-5}, 10^{-2}]$, the batch size $B \in \{32, 64, 128, 256\}$, the initial value of the feature fusion weights $w_m^{(0)} \in [0, 1]$, and the domain adaptation parameters $\gamma \in [0.1, 10]$ and $\beta \in [0.01, 1]$.

To fully evaluate the performance of DAMAF, we use a multidimensional evaluation metric:

(1) Recognition accuracy metrics: macro-mean precision (MP), macro-mean recall (MR) and macro-mean F1 score (MF1) are used. The definitions are as follows:

$$\mathrm{MP} = \frac{1}{C}\sum_{c=1}^{C} \frac{\mathrm{TP}_c}{\mathrm{TP}_c + \mathrm{FP}_c} \tag{35}$$

$$\mathrm{MR} = \frac{1}{C}\sum_{c=1}^{C} \frac{\mathrm{TP}_c}{\mathrm{TP}_c + \mathrm{FN}_c} \tag{36}$$

$$\mathrm{MF1} = \frac{2 \times \mathrm{MP} \times \mathrm{MR}}{\mathrm{MP} + \mathrm{MR}} \tag{37}$$

where C is the number of art style categories; TP_c, FP_c, and FN_c denote the number of true instances, false positive instances, and false negative instances of the c-th category, respectively.

(2) Domain Adaptation Performance Indicator (DAP): assesses the model's performance on emerging art styles.

$$\mathrm{DAP} = \frac{\mathrm{MF1}_{\text{EmergingArt}}}{\mathrm{MF1}_{\text{ArtStyleNet}}} \tag{38}$$

where $\mathrm{MF1}_{\text{EmergingArt}}$ and $\mathrm{MF1}_{\text{ArtStyleNet}}$ denote the macro-averaged F1 scores of the model on the EmergingArt and ArtStyleNet test sets, respectively.

(3) Computational Efficiency Indicator (CEI): assesses the computational efficiency of the model.

$$\mathrm{CEI} = \frac{T_{\text{baseline}}}{T_{\text{DAMAF}}} \tag{39}$$

where T_{baseline} and T_{DAMAF} are the average times of the baseline model and DAMAF, respectively, when processing the same number of samples.

(4) Multimodal Fusion Effectiveness Indicator (MMEI): assesses the effectiveness of multimodal feature fusion.

$$\mathrm{MMEI} = \frac{\mathrm{MF1}_{\text{DAMAF}} - \mathrm{MF1}_{\text{best single}}}{\mathrm{MF1}_{\text{best single}}} \times 100\% \tag{40}$$

where $\mathrm{MF1}_{\text{best single}}$ is the performance of the best unimodal model.

4.2. **Benchmark models and comparison methods.** In order to comprehensively assess the performance advantages of DAMAF, a series of representative benchmark models and the latest art style recognition methods are selected for comparison in this study. These methods cover different paradigms such as traditional machine learning, deep learning and migration learning to ensure the comprehensiveness and fairness of the comparison.

Traditional machine learning methods include Support Vector Machines (SVM) and Random Forests (RF). These methods are based on hand-designed features such as Scale Invariant Feature Transform (SIFT) and Histogram of Orientation Gradients (HOG). The benchmark deep learning model is ResNet-50 [26]. These models are pre-trained on the ImageNet dataset and then fine-tuned on ArtStyleNet.

In particular, we implemented the following state-of-the-art art style recognition method as a strong baseline:

- Multi-modal Fusion Model (MMFM): an approach proposed by Tashu et al. [11] to combine visual and textual information for art style recognition.
- Multi-modal Feature Extraction Framework (MMEF): a method designed by Liu et al. [12] to fuse color, texture and compositional information for art style recognition.

To ensure a fair comparison, all methods are trained and evaluated in the same hardware environment, using the same dataset partitioning and evaluation metrics. Each model is carefully tuned to achieve its optimal performance.

A series of ablation experiments were also designed to validate the effectiveness of the various modules of DAMAF:

- DAMAF-SM: Remove the adaptive metaheuristic algorithm selection mechanism and use a fixed optimization strategy.
- DAMAF-MF: Disables multimodal feature fusion and uses only single visual features.
- DAMAF-TL: Removal of the Progressive Domain Adaptive Migration Learning module and adoption of a traditional fine-tuning approach.
- DAMAF-Basic: Simultaneously removes all the above innovative modules and serves as the basic version of DAMAF.

4.3. **Performance comparison.** Table 1 summarizes the main performance metrics of each model on the ArtStyleNet test set and the EmergingArt dataset.

As can be observed from Table 1, DAMAF significantly outperforms the benchmark model on all evaluation metrics. On the ArtStyleNet test set, DAMAF's MF1 reaches 0.906, which is 5.5% higher than the best benchmark model MMEF. More notably, on the EmergingArt dataset, DAMAF's performance advantage is even more significant, with MF1 reaching 0.784, which is 15.3% higher than MMEF, demonstrating DAMAF's superior ability to handle emerging and rare art styles.

Table 1. Comparison of the performance of the models

Model	ArtStyleNet			EmergingArt		CEI	MMEI
	MP	MR	MF1	MP	MR		
SVM	0.632	0.615	0.623	0.421	0.407	1.00	-
RF	0.651	0.638	0.644	0.445	0.432	1.12	-
ResNet-50	0.812	0.796	0.804	0.583	0.569	1.35	-
MMFM	0.845	0.831	0.838	0.612	0.598	1.28	4.2%
MMEF	0.859	0.844	0.851	0.631	0.618	1.22	5.8%
DAMAF	**0.912**	**0.901**	**0.906**	**0.784**	**0.773**	**1.42**	**12.3%**

In terms of computational efficiency, although DAMAF incorporates multiple complex modules, it still achieves a CEI of 1.42, which outperforms all benchmark models. This indicates that DAMAF not only improves the recognition accuracy, but also maintains a high computational efficiency.

The MMEI of DAMAF reaches 12.3%, which is much higher than other multimodal methods, proving its superiority in multimodal feature fusion. This significant performance improvement is mainly attributed to DAMAF's Adaptive Dynamic Weight Fusion (ADWF) strategy, which can flexibly adjust the feature fusion ratio according to the characteristics of different artworks.

4.4. **Ablation experiments.** In order to deeply analyze the contribution of each component of DAMAF, we conducted a series of ablation experiments as shown in Table 2.

Table 2. Performance Comparison of DAMAF Variant Models

Model	ArtStyleNet		EmergingArt		CEI
	MF1	DAP	MF1	DAP	
DAMAF	0.906	0.865	0.784	0.866	1.42
DAMAF-SM	0.883	0.831	0.741	0.839	1.38
DAMAF-MF	0.871	0.812	0.718	0.824	1.45
DAMAF-TL	0.889	0.795	0.707	0.795	1.44
DAMAF-Basic	0.852	0.761	0.649	0.762	1.47

Comparing DAMAF with DAMAF-Basic, we can observe that the full DAMAF model improves the MF1 by 5.4% and 13.5% on the ArtStyleNet and EmergingArt datasets, respectively. This significant improvement highlights the combined effect of the various innovation modules of DAMAF.

The performance of DAMAF-MF is significantly lower than that of full DAMAF, with MF1 dropping 3.5% and 6.6% on ArtStyleNet and EmergingArt, respectively. This suggests that the ADWF strategy is crucial for improving the model's recognition ability, especially when dealing with emerging art styles.

The performance of DAMAF-TL on the EmergingArt dataset decreases significantly, with a 7.7% reduction in MF1 and a decrease in DAP from 0.866 to 0.795. This highlights the critical role of the PDAT technique in dealing with emerging art styles, significantly improving the generalization ability of the model.

DAMAF-SM outperforms the other variants of the model, although its performance is slightly lower than that of the full DAMAF. This suggests that the adaptive meta-heuristic algorithm selection mechanism can effectively improve the model performance, but its contribution is relatively small. However, it is worth noting that DAMAF-SM has a slight advantage in computational efficiency (CEI of 1.38), which may be due to the reduced computational overhead of the fixed optimization strategy.

Figure 2 visualizes the performance comparison of the DAMAF variants of the model on different evaluation metrics, visualizing the contribution of each module to the overall performance.

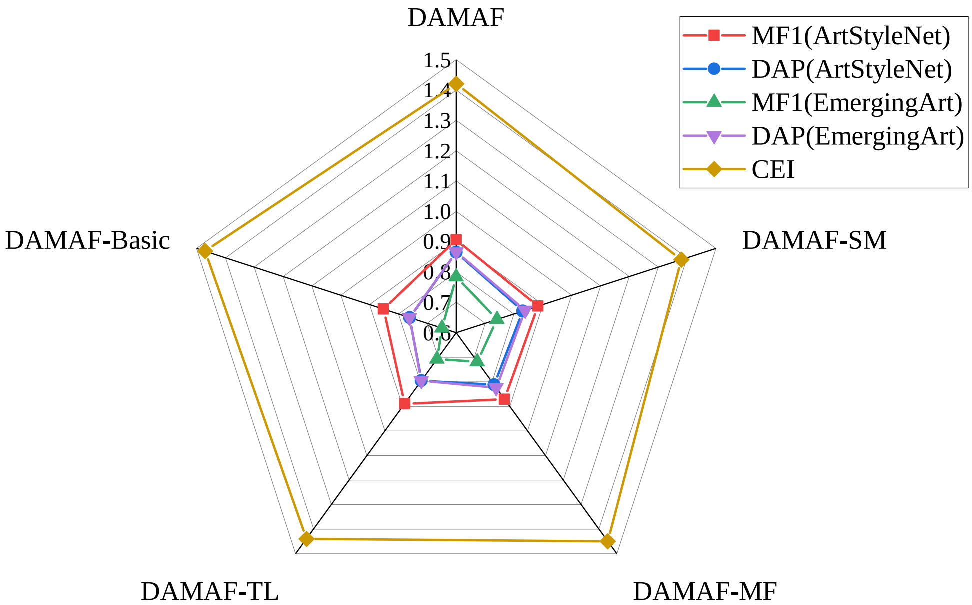

Figure 2. Comparison of DAMAF and its variant models

In summary, the results of the ablation experiments fully demonstrate the effectiveness and necessity of the innovative modules of DAMAF. The ADWF and PDAT are particularly significant in improving the model performance, especially when dealing with emerging and rare art styles. Meanwhile, the adaptive meta-heuristic algorithm selection mechanism, although its contribution is relatively small, plays an important role in improving model flexibility and adaptability. These results further confirm the superiority of DAMAF in the art style recognition task, especially the robust performance when facing diverse and emerging art styles.

4.5. **Case studies.**

4.5.1. *Emerging art styles recognition.* This case focuses on the performance of DAMAF in recognizing emerging digital art styles. We selected Generative Adversarial Network (GAN) art, which has been rapidly evolving in recent years, as the object of study. GAN artworks often have unique visual features, such as subtle texture variations and nontraditional compositions, which pose a challenge to traditional art style recognition algorithms.

We randomly selected 100 GAN artworks from the EmergingArt dataset for testing. Table 3 demonstrates the performance comparison between DAMAF and the best benchmark model MMEF in recognizing these artworks.

DAMAF shows a significant advantage in recognizing GAN artworks, with an MF1 score 22.3% higher than that of MMEF. In-depth analysis reveals that the success of DAMAF is

Table 3. GAN Artwork Recognition Performance Comparison

Model	MP	MR	MF1
MMEF	0.673	0.651	0.662
DAMAF	0.892	0.878	0.885

mainly attributed to the following factors: the ADWF strategy can dynamically adjust the weights of visual, texture and compositional features to adapt to the unique characteristics of GAN art; the PDAT technique effectively utilizes the traditional art knowledge and quickly adapts to the new art styles through progressive transfer learning.

4.5.2. *Analysis of artworks with an unbalanced stylistic character.* This case examines DAMAF's performance when dealing with artworks with an uneven stylistic profile. We selected a set of contemporary artworks that incorporate a variety of art styles, which are often difficult to categorize into a single art style category. We selected 50 works from the ArtStyleNet test set with an uneven stylistic profile for analysis. These works were labeled by art historians and each contained elements of 2-3 major styles. Table 4 shows the performance comparison between DAMAF and MMEF on this task.

Table 4. Comparison of analytical performance of works with uneven stylistic characterization

Model	Major style accuracy	Secondary Style accuracy	Average accuracy
MMEF	0.784	0.521	0.653
DAMAF	0.912	0.783	0.848

DAMAF demonstrates a strong ability to recognize works with uneven stylistic characteristics, and is particularly good at identifying secondary styles. Figure 3 shows a work that combines elements of Impressionism and Cubism. Figure 4 shows the results of DAMAF's analysis of the case.

These two case studies fully demonstrate the superiority of DAMAF in dealing with complex and emerging art styles. DAMAF is not only able to accurately identify emerging digital art forms, but also to deeply analyze works with uneven stylistic characteristics, which is of great importance to the fields of art history research, digital art curation, and intelligent art recommendation systems.

5. **Conclusion.** In this paper, an art style recognition framework (DAMAF) is proposed to effectively address the limitations of traditional art style recognition models in dealing with diverse and emerging art styles. With the ADWF strategy, the model is able to integrate multimodal features more flexibly, which significantly improves the recognition accuracy. In addition, the PDAT technique further enhances the model's ability to adapt to emerging and rare art styles, ensuring the stability and robustness of the recognition results. The following conclusions can be drawn from the experiments conducted on ArtStyleNet and EmergingArt datasets:

1. Multimodal feature fusion can significantly improve the accuracy of art style recognition, especially for works with uneven stylistic features.
2. The PDAT technique demonstrates greater generalization ability when dealing with emerging art styles compared to traditional transfer learning methods.

Figure 3. Examples of works with uneven stylistic features

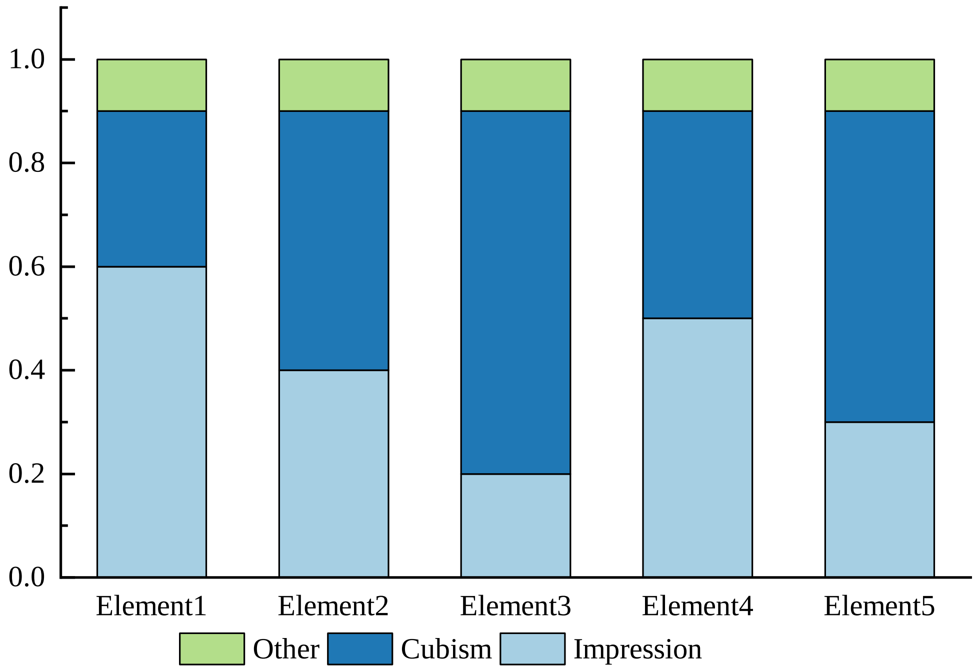

Figure 4. Visualization of the results of DAMAF's analysis of cases

3. The DAMAF framework combining ADWF and PDAT strikes the best balance between recognition accuracy and computational efficiency and is the optimal strategy recommended in this paper.
4. The adaptive meta-heuristic algorithm selection mechanism, although its contribution is relatively small, plays an important role in improving model flexibility and adaptability.

5. DAMAF has demonstrated excellent performance in dealing with emerging forms of digital art such as GAN art, opening up new possibilities for digital art analysis.

The experimental data in this paper are mainly from the ArtStyleNet and EmergingArt datasets, which may still have some limitations despite the fact that they cover a wide range of art styles. Future work should consider introducing more data from different cultural backgrounds and art genres to validate the effectiveness of the model in more diverse art scenes. In addition, it is worth exploring how to further improve the model's ability to recognize extremely rare or completely unknown art styles, as well as how to more accurately quantify and describe the degree of integration of multiple stylistic elements in a work. These researches not only help to promote the development of art style recognition technology, but also may bring new breakthroughs in the fields of art history research, digital art curation and intelligent art recommendation systems.

REFERENCES

[1] J. Li and B. Zhang, "The application of artificial intelligence technology in art teaching taking architectural painting as an example," *Computational Intelligence and Neuroscience*, vol. 2022, no. 1, p. 8803957, 2022.

[2] Y. Shen and F. Yu, "The influence of artificial intelligence on art design in the digital age," *Scientific Programming*, vol. 2021, no. 1, p. 4838957, 2021.

[3] Y. W. A. Wong and E. C. P. Patron, "A review of the interrelationships between painting, photography, facial recognition, and artificial intelligence technologies in portraiture aesthetics," *Athens Journal of Technology and Engineering (AJTE)*, vol. 11, no. 2, pp. 167-186, 2024.

[4] E. Cetinic and J. She, "Understanding and creating art with AI: Review and outlook," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 18, no. 2, pp. 1-22, 2022.

[5] M.-C. Chiu, G.-J. Hwang, L.-H. Hsia, and F.-M. Shyu, "Artificial intelligence-supported art education: A deep learning-based system for promoting university students' artwork appreciation and painting outcomes," *Interactive Learning Environments*, vol. 32, no. 3, pp. 824-842, 2024.

[6] R. G. Condorovici, C. Florea, and C. Vertan, "Automatically classifying paintings with perceptual inspired descriptors," *Journal of Visual Communication and Image Representation*, vol. 26, pp. 222-230, 2015.

[7] Z. Falomir, L. Museros, I. Sanz, and L. Gonzalez-Abril, "Categorizing paintings in art styles based on qualitative color descriptors, quantitative global features and machine learning (QArt-Learn)," *Expert Systems with Applications*, vol. 97, pp. 83-94, 2018.

[8] S.-G. Lee and E.-Y. Cha, "Style classification and visualization of art painting's genre using self-organizing maps," *Human-centric Computing and Information Sciences*, vol. 6, pp. 1-11, 2016.

[9] K.-L. Hua, T.-T. Ho, K.-A. Jangtjik, Y.-J. Chen, and M.-C. Yeh, "Artist-based painting classification using Markov random fields with convolution neural network," *Multimedia Tools and Applications*, vol. 79, pp. 12635-12658, 2020.

[10] K. Zhang, Y. Su, X. Guo, L. Qi, and Z. Zhao, "MU-GAN: Facial attribute editing based on multi-attention mechanism," *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 9, pp. 1614-1626, 2020.

[11] W. Zhao, D. Zhou, X. Qiu, and W. Jiang, "How to represent paintings: A painting classification using artistic comments," *Sensors*, vol. 21, no. 6, p. 1940, 2021.

[12] T. M. Tashu, S. Hajiyeva, and T. Horvath, "Multimodal emotion recognition from art using sequential co-attention," *Journal of Imaging*, vol. 7, no. 8, p. 157, 2021.

[13] F. Liu, M. Zhang, B. Zheng, S. Cui, W. Ma, and Z. Liu, "Feature fusion via multi-target learning for ancient artwork captioning," *Information Fusion*, vol. 97, p. 101811, 2023.

[14] L. Sixu, W. Muqing, and Z. Min, "Particle swarm optimization and artificial bee colony algorithm for clustering and mobile based software-defined wireless sensor networks," *Wireless Networks*, vol. 28, no. 4, pp. 1671-1688, 2022.

[15] R. Beed, A. Roy, S. Sarkar, and D. Bhattacharya, "A hybrid multi-objective tour route optimization algorithm based on particle swarm optimization and artificial bee colony optimization," *Computational Intelligence*, vol. 36, no. 3, pp. 884-909, 2020.

[16] C. Jiang, J. Fu, and W. Liu, "Research on vehicle routing planning based on adaptive ant colony and particle swarm optimization algorithm," *International Journal of Intelligent Transportation Systems Research*, vol. 19, no. 1, pp. 83-91, 2021.

[17] S. E. Comert and H. R. Yazgan, "A new approach based on hybrid ant colony optimization-artificial bee colony algorithm for multi-objective electric vehicle routing problems," *Engineering Applications of Artificial Intelligence*, vol. 123, p. 106375, 2023.

[18] T.-Y. Wu, H. Li, S. Kumari, and C.-M. Chen, "A Spectral Convolutional Neural Network Model Based on Adaptive Fick's Law for Hyperspectral Image Classification," *Computers, Materials & Continua*, vol. 79, no. 1, pp. 19-46, 2024.

[19] T.-Y. Wu, A. Shao, and J.-S. Pan, "CTOA: Toward a Chaotic-Based Tumbleweed Optimization Algorithm," *Mathematics*, vol. 11, no. 10, p. 2339, 2023.

[20] T.-Y. Wu, H. Li, and S.-C. Chu, "CPPE: An Improved Phasmatodea Population Evolution Algorithm with Chaotic Maps," *Mathematics*, vol. 11, no. 9, p. 1977, 2023.

[21] X. Du, B. Li, T. Liu, L. Jin, Y. Ding, and Z. Zhao, "Efficient visual anomaly detection model with adaptive wavelet transform," *Optics and Lasers in Engineering*, vol. 182, p. 108457, 2024.

[22] B. S. Rao, "Dynamic histogram equalization for contrast enhancement for digital images," *Applied Soft Computing*, vol. 89, p. 106114, 2020.

[23] L. Huang et al., "Temporal hierarchical graph attention network for traffic prediction," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 12, no. 6, pp. 1-21, 2021.

[24] Y. Shi, L. Du, Y. Guo, and Y. Du, "Unsupervised domain adaptation based on progressive transfer for ship detection: From optical to SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-17, 2022.

[25] K. Han et al., "A survey on vision transformer," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 1, pp. 87-110, 2022.

[26] B. Li and D. Lima, "Facial expression recognition via ResNet-50," *International Journal of Cognitive Computing in Engineering*, vol. 2, pp. 57-64, 2021.