

Improved DenseNet-DCGAN for Enhanced Digital Restoration of Embroidery Cultural Heritage

Guiying Dong^{1*}, Qian Mao²

¹College of Art and Design, Communication University of China Nanjing, Nanjing 210000, China

²Library, Nanjing University, Nanjing 210000, China

E-mail: adong118@126.com; maoqian328@163.com

*Corresponding author

Keywords: DCGAN, DenseNet, embroidery, image classification, image restoration

Received: July 14, 2025

At present, embroidery image restoration technology still has deficiencies in terms of color uniformity and detail restoration. To address these issues, the study improves the densely connected convolutional network and the deep convolutional generative adversarial network through spatial pyramid pooling, and proposes a novel method for embroidery image classification and restoration. The experimental results showed that the research method largely restored the details and colors of the original image and effectively addressed the uneven color issue. The average prediction accuracy, recall rate, and specificity of the image classification model on Suzhou embroidery, Hunan embroidery, Guangdong embroidery, and Shu embroidery reached 96.3%, 98.5%, and 99.4%, respectively. The structural similarity index of the image restoration model has reached 0.99. The restored image was almost indistinguishable to the naked eye in terms of details, texture, and color. The research method has significant advantages in classifying embroidery images and high-quality restoration tasks, and can provide reliable technical support for the digital protection and intelligent restoration of traditional embroidery cultural relics.

Povzetek: Za klasifikacijo in digitalno obnovo vzemenin so razviti izboljšani DenseNet in DCGAN z dodanim SPP, razširjenimi konvolucijami ter CBAM. Izboljšani model skoraj povsem naravno obnovi teksturo in barve.

1 Introduction

Embroidery works have attracted countless people's attention with their exquisite craftsmanship, rich patterns, and profound cultural connotations. However, over time, many embroidery artifacts have suffered from natural or human damage, such as fading, breakage, and insect infestation, which seriously threaten the preservation and inheritance of embroidery artifacts [1]. The traditional restoration of Embroidered Cultural Relics (ECR) mainly relies on manual skills. Although this method can finely handle every damage, it is limited by lower work efficiency and dependence on the superb skills of the restorer [2]. In addition, the subjectivity in the manual repair process may also lead to deviations in the consistency and accuracy of the repair effect. In this context, the emergence of Artificial Intelligence (AI) technology, especially Deep Learning (DL) technology, has provided new solutions for the restoration of cultural relics. By training DL models, staff can automatically detect and classify the types of damage to cultural relics, providing a scientific basis for restoration work. At present, many researchers have explored it. For example, Maitin et al. proposed a direct reconstruction technique without image segmentation using DL technology to reconstruct missing architectural elements in Greek temple ruins images from virtual image paintings. This method has successfully reconstructed the missing

architectural elements in the images of Greek temple ruins, improving the efficiency of restoration and enhancing the consistency and accuracy of the restoration effect [3]. Alessandro et al. used a trained multidimensional DL neural network to associate color images with X-ray fluorescence imaging raw data to complete the restoration of AI digital cultural heritage, achieving digital restoration of graphic artworks [4].

With the further advancement of DL technology, Generative Adversarial Networks (GANs) have made breakthrough progress in image recognition, providing a good solution for cultural relic image restoration [5]. Praveen et al. proposed a new GAN-based art restoration method to digitally repair damaged artworks and assist in physical restoration. This method performed well in digital restoration and could effectively restore the original appearance of artworks, providing important guidance for physical restoration [6]. Zheng et al. proposed an Example Attention Generative Adversarial Network (EA-GAN) that fuses with reference examples, which addressed the issue of significant reconstruction errors in traditional character restoration methods. Compared with existing internal drawing networks, EA-GAN could obtain the correct text structure through the guidance of additional examples in the "example attention block". The Peak Signal-to-Noise Ratio (PSNR) and

Structural Similarity Index (SSIM) values have increased by 9.82% and 1.82% [7].

In summary, numerous scholars have achieved significant results in cultural relic image restoration. However, there are still issues with GAN in terms of image feature extraction, such as poor network training stability and poor generated image quality. At present, there is relatively limited discussion on embroidery classification and restoration in cultural relic image classification and restoration. Given this, this study innovatively constructs an ECR-Image Classification Model (ICM) based on Densely Connected Convolutional Network (DenseNet) and an ECR-Image Restoration Model (IRM) based on Deep Convolutional GAN (DCGAN). Based on these models, improvements are made by introducing Local Binary Patterns (LBPs), Canny operator edge extraction, and Convolutional Block Attention Module (CBAM). The fusion of these technologies aims to enhance the model's capacity to capture details in ECR images, improve the precise reconstruction of textures and edges during the restoration process, and achieve higher quality ECR image restoration results. The main novelizations and contributions of this paper include: (1) For the first time, DenseNet is combined with Spatial Pyramid Pooling (SPP) and applied to classify embroidery images, improving the recognition performance under cross-style and complex patterns; (2) The structure of the DCGAN generator and discriminator is innovatively adjusted. By integrating a dilated convolutional layer, the receptive field of the model is expanded, which helps to capture image features more comprehensively and achieve high-quality restoration of embroidery texture and color. (3) A large-scale dataset containing eight types of traditional embroidery images is constructed, providing fundamental support for subsequent research. The research results have practical value for the digital inheritance and AI-assisted restoration of traditional embroidery culture.

2 Methods and materials

2.1 Construction of ECR-ICM based on SPP-IDenseNet

ECR image classification is the prerequisite and foundation for ECR image restoration. By classifying ECR images, different embroidery types, styles, and eras can be quickly identified and distinguished, providing a scientific foundation for protecting the cultural relics. This study first explores the classification of ECR images. DenseNet was proposed by Gao et al. in 2017. It is a novel

DL model architecture that can establish dense connections between network layers through DenseBlocks, thereby improving the information flow and gradient flow of the network, alleviating the problem of gradient vanishing, and promoting feature reuse [8-9]. The structure of DenseBlock in DenseNet is displayed in Fig.1.

In Fig.1, the connection mechanism of DenseBlock is more aggressive compared to the Residual Network (ResNet). Each layer is connected to all previous layers, providing each layer with a rich input that integrates the features of all previous layers [10]. This design ensures the uniformity of feature map size within DenseBlock and greatly promotes feature reuse through dense connections between layers, enabling the network to learn and transmit information more effectively [11]. However, DenseNet still has certain shortcomings in the image classification process, such as the problem of input image size limitation and the problem of network training not converging [12-13]. Therefore, this study improves it through techniques such as SPP, LBP, and Canny operator, and proposes a novel ECR-ICM model, namely the SPP-IDenseNet model. The training process for the embroidery image classification of this model is shown in Fig.2.

In Fig.2, this study first randomly selects a batch of data from the training set based on a preset batch size, and normalizes it to standardize the standard deviation of the Red-Green-Blue (RGB) color channels for each embroidery image. Subsequently, the normalized image is input into the network for forward propagation to extract features and predict categories. Secondly, by comparing the predicted categories of the network with the actual categories, the value of the loss function is calculated. Next, by adjusting the weights through the backward propagation process of the network, the model's performance is optimized. After completing a batch of training, the system will check if the entire dataset has been traversed. If the traversal is not completed, the model will continue to process the next training batch and repeat the above steps. Once the training traversal of the entire dataset is completed, the model will save the weight parameters of the current round and evaluate whether the predetermined training round has been reached. If the training rounds have not been completed, the model will restart the training process and continue iterative optimization. After reaching the predetermined training round, the model training terminates, and the weight parameters at this time will be used for subsequent image classification tasks. The calculation of the RGB three channel pixel values $Output_R$, $Output_G$, and $Output_B$ of the normalized image is shown in formula (1).

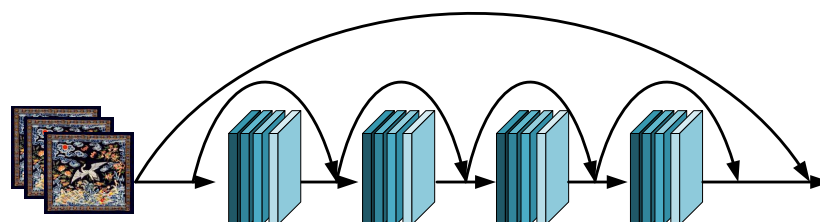


Figure 1: Schematic structure of DenseBlock (Source from: <https://colorhub.me/photos/e7RVB>).

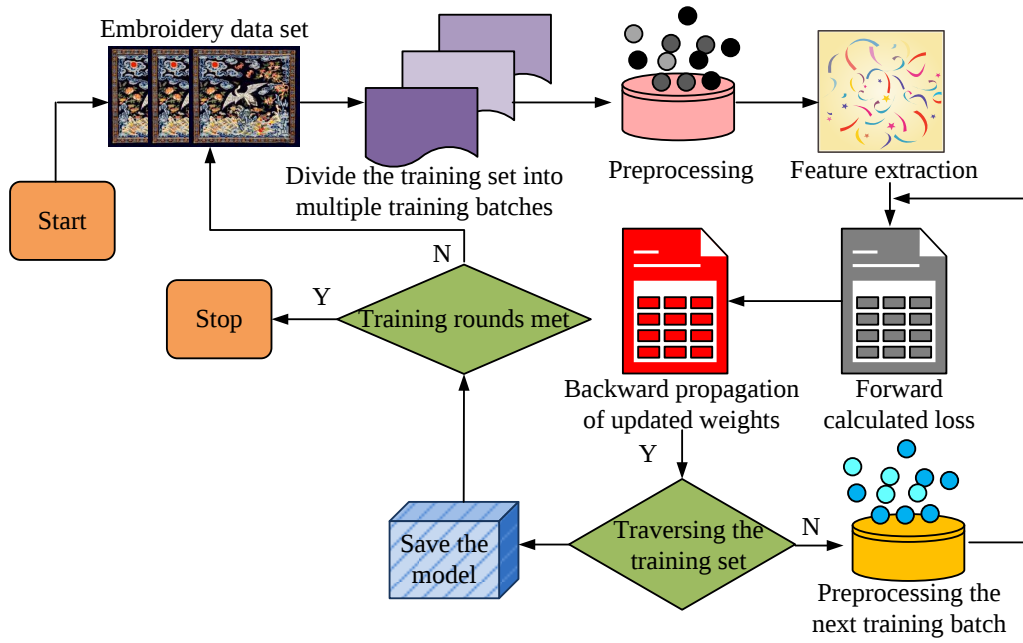


Figure 2: Training process of SPP-IDenseNet model for embroidery image classification (Source from: <https://colorhub.me/photos/e7RVB>).

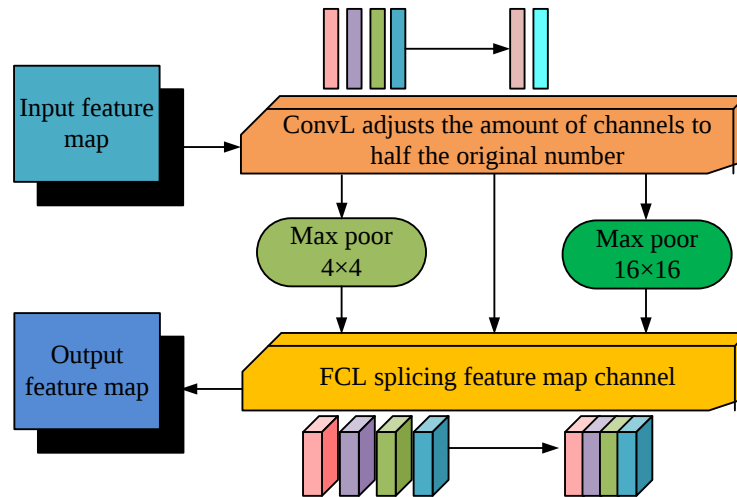


Figure 3: Schematic diagram of SPP structure.

$$\begin{cases} Output_R = \frac{Input_R - mean_R}{std_R} \\ Output_G = \frac{Input_G - mean_G}{std_G} \\ Output_B = \frac{Input_B - mean_B}{std_B} \end{cases} \quad (1)$$

In formula (1), $Input_R$, $Input_G$, and $Input_B$ are the RGB three channel pixel values of the image before normalization processing. $mean_R$, $mean_G$, and $mean_B$ are the mean values of the RGB channels. std_R , std_G , and std_B represent the standard deviation of the RGB three channels. The output feature M_n is shown in formula (2).

$$M_n = F(M_1 \oplus M_2 \oplus \dots \oplus M_{n-1}) \quad (2)$$

In formula (2), n is the hierarchy of the model network. F and $\oplus$ are convolution operations and feature interconnection operations. The loss l during the training process is shown in formula (3).

$$l = L(Y_1, Y'_1) \quad (3)$$

In formula (3), L is the loss function. Y_1 and Y'_1 are the real category and the predicted category. The updated network weight θ' is shown in formula (4).

$$\theta' = \theta - l_r \times \nabla g(l) \quad (4)$$

In formula (4), θ is the network weight before the update. l_r and $\nabla g(\cdot)$ are learning efficiency and derivative calculation. In response to the issue of input image size limitation in DenseNet model image classification tasks, this study uses SPP to enable the model to adapt to input images of different sizes. The structure of SPP is shown in Fig.3.

In Fig.3, this study integrates the SPP structure between the convolutional layer and the Fully Connected Layer (FCL) at the end of the DenseNet model. By dividing the feature map into grids of 1×1 , 4×4 , and 16×16 , and applying max pooling, this study achieves comprehensive capture of features of different resolutions. Subsequently, these multi-scale pooled feature maps will be merged into a fixed-length feature vector, providing rich information for the input of FCL. In addition, by pooling on windows of different sizes, this study generates feature maps with diverse resolutions and fine-tunes the channel dimensions through a 1×1 convolutional layer. The ReLU activation function used in DenseNet may cause neuron deactivation when the input is less than 0 [14]. Therefore, this study introduces the Leaky ReLU function and sets the negative slope coefficient to 0.01, effectively expanding the applicability of ReLU and promoting the stability and convergence of network training. The SPP module enhances the model's understanding of the structural hierarchy of embroidery patterns through multi-scale pooling operations and improves the receptive field coverage of complex patterns. LBP extracts fine-grained texture features from embroidery images, enabling the model to pay more attention to the local texture restoration of the defect area. Canny edge detection provides clear structural contour constraints, guiding the generator to maintain the coherence and integrity of the pattern edges. The three work in synergy, enhancing the quality and stability of image restoration from multiple dimensions, such as structure, texture, and edge.

2.2 Construction of ECR-IRM based on Improved DCGAN

The SPP-IDenseNet model designed above provides strong technical support for the digital restoration and intelligent management of ECR. However, further technological innovation and method improvement are needed in ECR image restoration to achieve more efficient and accurate restoration results. Therefore, this study explores the restoration of ECR images. GAN is a DL model containing two parts: the Generator and Discriminator. Although GANs are widely popular in computer image vision, in traditional GAN architectures, models do not rely on a determined distribution, but instead use internal feedback to adjust their parameters [15]. Although this approach enhances the flexibility of

the model, it may also cause training instability and sometimes even lead to model training crashes [16–17]. Therefore, this study further introduces a novel derivative GAN, namely DCGAN. This network can improve the quality of image generation and enhance the learning and representation capabilities of the model by combining the deep architecture of CNN with the GAN framework. The generator extends and reshapes 100-dimensional noise into a 3D feature map through FCL, and then gradually forms the final image size through upsampling and dimension adjustment of transposed convolutional layers. Batch normalization and ReLU are applied after each layer, and the output image is finally activated by Sigmoid to produce a specific tensor image [18–19]. The generator's loss function is shown in formula (5).

$$L_G = -E_{z \sim p(z)} [\log(D(G(z)))] \quad (5)$$

In formula (5), E is the expected operation symbol, usually taken as the average or expected value. z is a noise sample from the latent spatial prior distribution. $G(z)$ is the data generated by the generator through the noise sample z . $D(G(z))$ is the output of the discriminator to the data generated by the generator, which represents the probability of real data. At this point, the loss function of the discriminator is shown in formula (6).

$$L_D = -E_{x \sim p_{data}(x)} [\log(D(x_z))] - E_{z \sim p(z)} [1 - D(G(x_z))] \quad (6)$$

In formula (6), $D(x_z)$ means the discriminator's output for the real sample x_z , and $D(x_z)$ is the probability of the real data. Based on the above formulas, compared to traditional GANs, DCGAN uses convolutional and deconvolution layers to replace FCL in traditional GANs. This operation can capture the local structure and spatial message of embroidery images [20]. In addition, DCGAN also uses batch normalization techniques and expected values to accelerate the training process and stabilize the training of GAN. The aim is to further enhance the performance of DCGAN in embroidery image restoration tasks, improve the naturalness of restoration effects, and provide experts with more accurate texture and color information to assist them in more refined restoration work. Given this, the study also improves DCGAN and proposes a new type of ECR-IRM, namely IDCAN. The overall model structure is shown in Fig.4.

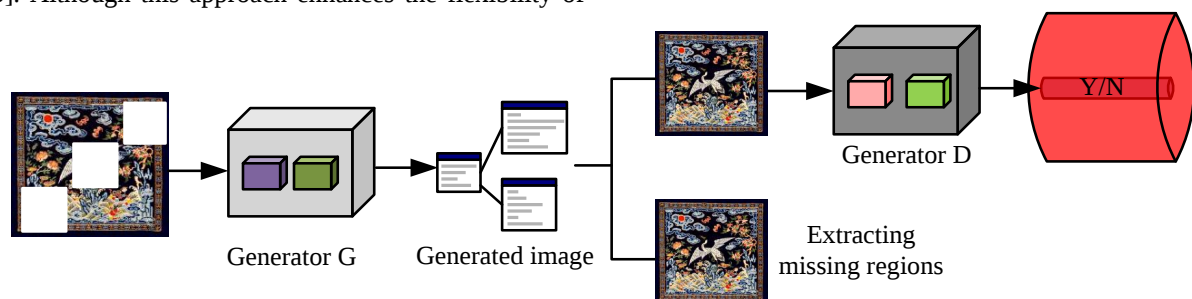


Figure 4: Overall structural framework of the IDCAN (Source from: <https://colorhub.me/photos/e7RVB>).

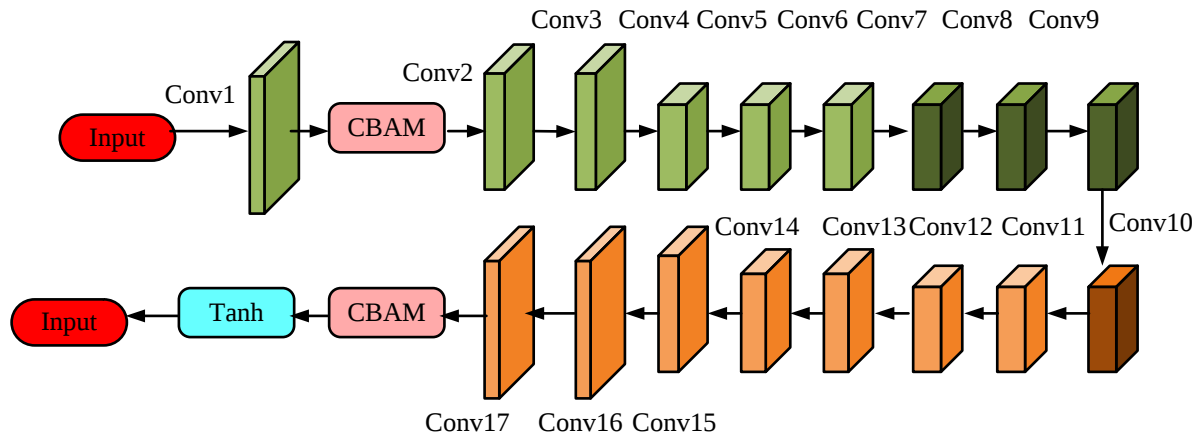


Figure 5: Specific structure of the generator in the IDCGAN model.

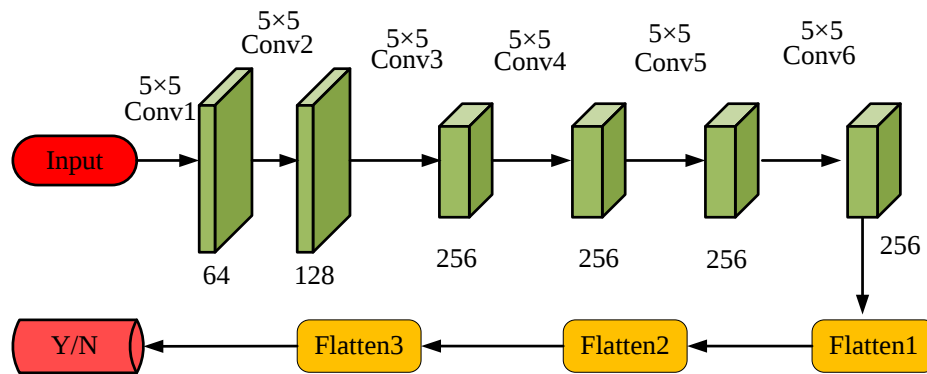


Figure 6: Specific structure of the discriminator in the IDCGAN.

In Fig.4, innovative adjustments are made to the generator architecture by integrating dilated convolutional layers to expand the model's receptive field, thereby helping the model capture image features more comprehensively. At the same time, CBAM is introduced to enhance the attention to key features at both the channel and spatial levels, thereby improving the accuracy of image restoration. The discriminator adopts a strategy of enhancing its depth and increasing the number of FCLs, thereby improving the network's ability to handle complex nonlinear problems, enabling the discriminator to more effectively recognize and distinguish between real and generated images. The loss function combines traditional MSE loss with adversarial loss. The calculation of mean square error loss L_{MSE} is shown in formula (7).

$$L_{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - g_i)^2 \quad (7)$$

In formula (7), g_i is the predicted value of the model on the training data x_i . The adversarial loss L_{adv} is shown in formula (8).

$$L_{adv} = \min_G \max_D E_{x \sim P_{data}} [\log_2 D(x)] + E_{z \sim P_{data}} [\log_2 (1 - D(G(z)))] \quad (8)$$

In formula (8), the algebraic meaning remains the same as before. The specific structure of the generator in the IDCGAN model is shown in Fig.5.

In Fig.5, the architecture of the generator in the IDCGAN model mainly consists of three key modules,

namely the convolution block, dilated convolution block, and CBAM. Hollow convolution blocks use convolutional layers with different void rates, namely 2, 4, 8, and 16, to achieve multi-scale capture of image features. When the hole rate is set to 1, the hole convolution degenerates into a standard convolution operation. This is reflected in the Conv6 to Conv10 layers of the generator, forming a series of ConvLs with different hole rates that ensure the flexibility and adaptability of the network. The introduction of CBAM adds the ability for dynamic weighting to the generator. It can weight features in both channel and spatial dimensions, highlighting the features that have the greatest impact on image quality. The framework of the discriminator in the IDCGAN model is shown in Fig.6.

In Fig.6, to improve the performance of the discriminator in addressing complex nonlinear problems, this study adds two FCLs based on the original discriminator architecture, making the discriminator contain a total of three FCLs. The interconnection of these layers enhances the discriminator's ability to learn features, thereby significantly improving model performance. Ultimately, the discriminator determines the authenticity of the input image through a binary classification task, distinguishing whether the image was generated by the generator or from a real dataset. The research is conducted based on a self-built embroidery image dataset. The images mainly come from digital museums, high-resolution cultural relic catalogues, and

cultural heritage archives, covering multiple historical periods and diverse embroidery styles. The initial dataset contains 1,800 images. After expansion, the dataset ultimately includes 8,957 images. For the unified model input, the image is cropped and scaled to 256×256 pixels, and normalization processing is carried out simultaneously. Ultimately, the dataset is divided into a training set and a test set in an 8:2 ratio. To simulate the common damage forms of ECR, the study also uses random occlusion to generate defect images. The occlusion forms include rectangles, free-shaped patterns, and speckled textures, and the area ratio is controlled at 10% to 40%. On this basis, image enhancement is carried out by applying methods such as rotation, flipping, scaling, and color perturbation to improve the robustness and generalization ability of the model. In addition, by analyzing the color and style distribution of the images, a balanced sampling strategy is adopted to control the category bias, ensuring the diversity and balance of the training data in terms of pattern style and damage type. All the code modules in the research are built based on the PyTorch framework. Some of the code is as follows:

3 Results

3.1 SPP-IDenseNet model performance testing

The study adopts five-fold cross-validation to evaluate the model's performance. The training set is evenly divided into five subsets of similar size. Four subsets are selected in sequence for model training, and the remaining subset is used as the validation set. This process is repeated five times to ensure that each subset participates in the verification. Through multiple rounds of training and validation, the mean and standard deviation of the accuracy, recall rate, and specificity of the calculation model are calculated, effectively avoiding the randomness brought by a single division and enhancing the statistical reliability and generalization ability of the evaluation results. Table 1 shows the experimental setup and environmental parameters.

According to the settings in Table 1, the effectiveness of the proposed model was first validated through ablation testing, as shown in Fig.8.

```
import torch
import torch.nn as nn
# Simple Generator example
class Generator(nn.Module):
    def __init__(self):
        super().__init__()
        self.net = nn.Sequential(
            nn.Linear(100, 256),
            nn.ReLU(),
            nn.Linear(256, 3*64*64),
            nn.Sigmoid()
        )
    def forward(self, x):
        return self.net(x).view(-1, 3, 64, 64)
# Simple Discriminator example
class Discriminator(nn.Module):
    def __init__(self):
        super().__init__()
        self.net = nn.Sequential(
            nn.Flatten(),
            nn.Linear(3*64*64, 256),
            nn.ReLU(),
            nn.Linear(256, 1),
            nn.Sigmoid()
        )
    def forward(self, x):
        return self.net(x)
# Training example (pseudo code)
# z = torch.randn(batch_size, 100)
# fake_images = generator(z)
# real_output = discriminator(real_images)
# fake_output = discriminator(fake_images)
```

Figure 7: Code.

Table 1: Environment and parameter configuration.

Serial number	Experimental environments and hyperparameter categories	Settings
1	Num epochs	200
2	Pre-training	No
3	Batch size	20
4	Num class	8
5	Optimizer	Adam
6	Learning rate	0.0001
7	Development Environment	Windows 10
8	CPU	Intel Core i9-10900K
9	GPU	NVIDIA RTX 3090
10	Memory	64GB
11	Graphics Memory	16GB GDDR6X
12	Programming Tools	PyTorch 1.6.0

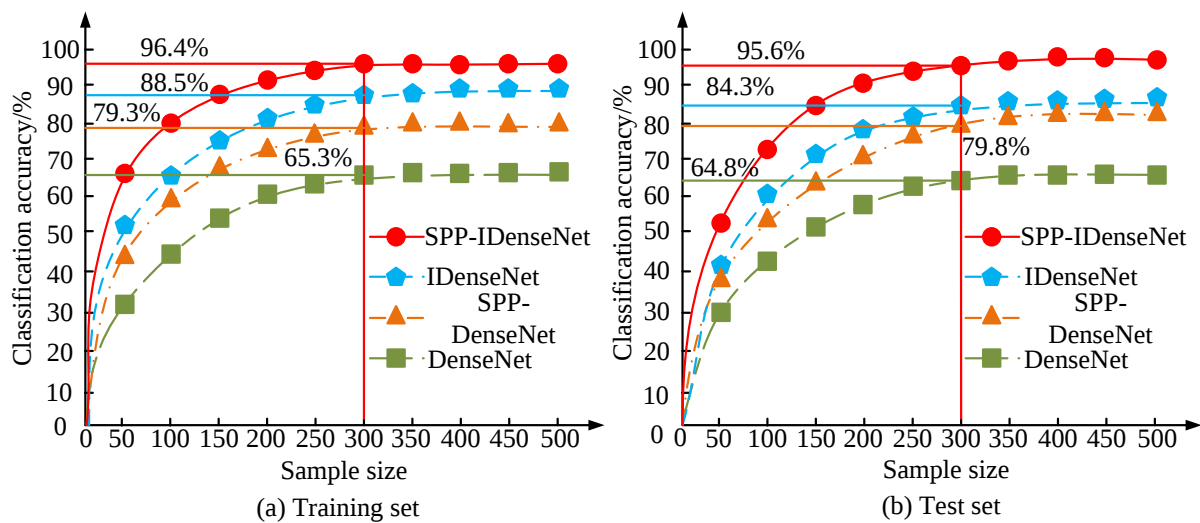


Figure 8: Ablation test results of SPP-IDenseNet.

Figs.8 (a) and (b) show the test results of the new model in two datasets. As the test samples continue to grow, the standalone DenseNet module shows lower classification accuracy in both datasets, with the highest being only 65.3%. After introducing the SPP module, LBP, Canny operator, and Gabor filter module successively, the classification effectiveness of the entire model has been significantly improved. The result indicates that when dealing with embroidery images with complex texture features, relying solely on global features for extraction has certain performance bottlenecks. The classification accuracy of SPP-IDenseNet is highest at 96.4% in the training set and 95.6% in the testing set. This study has improved various parts of the DenseNet model to varying degrees for classifying and recognizing ECR images, demonstrating the effectiveness of the improved method. In addition, popular ICMs of the same type, including Lightweight CNN (LCNN), Efficient CNN (ECNN), StyleGAN, and Global Image Spatial Texture (GIST), are introduced as comparative models. Performance tests are conducted using precision, recall, and specificity as indicators, as shown in Table 2.

In Table 2, due to their relatively simplified structures, LCNN and ECNN models have obvious deficiencies in feature expression ability and fine-grained classification.

Although the GIST model can capture certain texture information, it is limited by its feature extraction method based on compressed texture description. GIST's recognition ability for irregular shapes and multi-scale patterns is weak, resulting in limited classification performance. The SPP-IDenseNet model has demonstrated superior performance in all four types of embroidery image recognition tasks. This model enhances its feature perception ability for different scales and spatial structures by introducing SPP modules, and combines LBP and Gabor filters to model fine-grained textures, effectively improving the model's ability to recognize the microstructure of embroidery patterns. Meanwhile, the addition of the Canny edge detection operator enhances the ability to capture boundary and contour features, enabling the model to maintain high classification accuracy even in the face of complex background interference. The SPP-IDenseNet model has the highest accuracy rate of 96.3%, the highest recall rate of 98.5%, and the highest specificity of 99.4% on Suzhou embroidery, Hunan embroidery, Guangdong embroidery, and Shu embroidery. These indicators are numerically superior to other convolutional neural network models and have a more balanced distribution across categories. This result demonstrates the adaptability and effectiveness of

the SPP-IDenseNet model in handling the classification task of ECR images. The confusion matrix obtained on the embroidery image classification dataset before and after model improvement is shown in Fig.9.

Figs. 9 (a) and (b) show the confusion matrices before and after model improvement. The SPP-IDenseNet model has the highest classification and recognition accuracy for Shui ethnic ponytail embroidery, Xiqin, Hani, Su, Xiang, Shu, and other embroidery in the embroideries image classification dataset. Its classification accuracy in Yue embroidery types is relatively the lowest. Overall, the SPP-IDenseNet model achieves an average prediction accuracy of over 80% for the 8 styles of embroidery images in the embroidery image classification dataset. This indicates that the SPP-IDenseNet model demonstrates strong robustness in handling noise,

occlusion issues, and identifying embroidery images with similar features in the embroidery image classification dataset. This robustness makes the SPP-IDenseNet model a powerful tool for ECR image classification, which can effectively address the challenges in practical applications.

3.2 Performance simulation testing of ECR-IRM for IDCAN

This study uses the Tensorflow DL framework to implement the training and testing of the entire ECR-IRM. The weights β_1 and β_2 of the Adam optimizer are set to 0.5 and 0.9. The loss changes of IDCAN generator and discriminator at different network learning rates are shown in Fig.10.

Table 2: Multi-metric performance test results for different models.

Style	Model	Precision/%	Recall/%	Specificity/%
Suzhou embroidery	LCNN	63.5	65.7	80.2
	ECNN	67.2	69.8	81.6
	GIST	70.3	68.7	83.4
	StyleGAN	85.7	87.4	89.1
	Research method	95.8	98.5	94.2
Hunan embroidery	LCNN	55.2	56.3	89.6
	ECNN	58.7	60.4	90.2
	GIST	60.2	61.7	91.6
	StyleGAN	83.4	85.1	92.3
	Research method	96.3	90.2	99.4
Cantonese embroidery	LCNN	57.6	59.8	53.8
	ECNN	66.3	70.4	60.5
	GIST	71.6	69.7	70.8
	StyleGAN	80.2	82.5	75.4
	Research method	95.1	90.8	95.1
Sichuan embroidery	LCNN	58.8	60.5	55.6
	ECNN	62.8	68.8	58.3
	GIST	70.4	73.4	60.7
	StyleGAN	79.8	81.7	69.2
	Research method	92.4	96.7	90.3

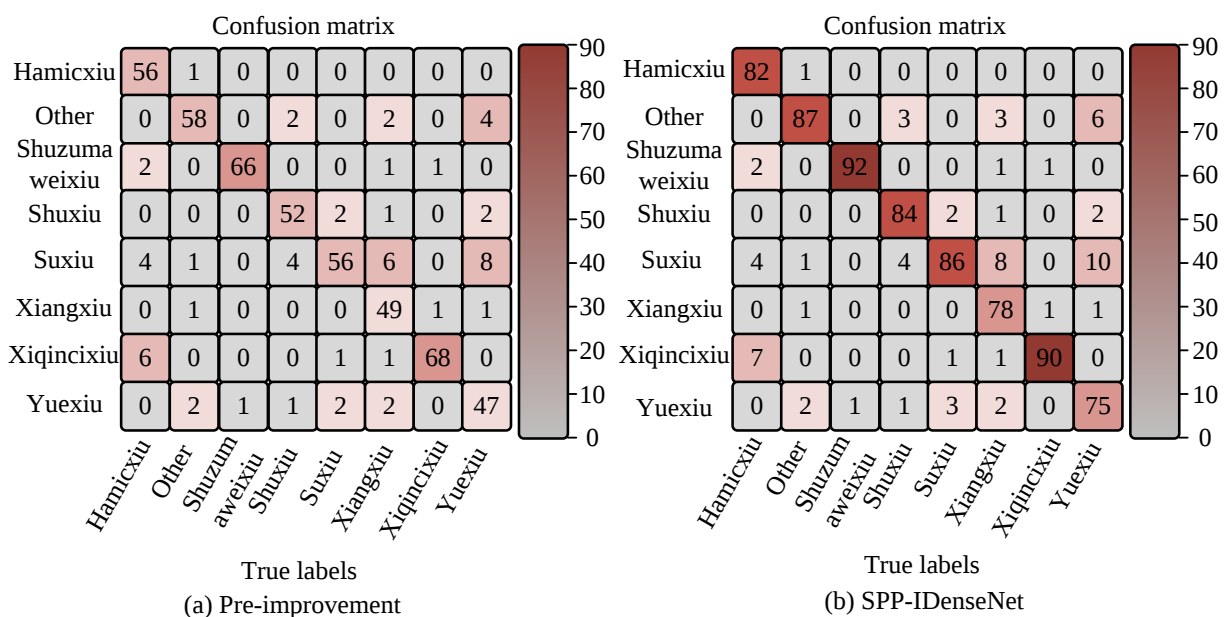


Figure 9: Confusion matrix plots before and after model improvement.

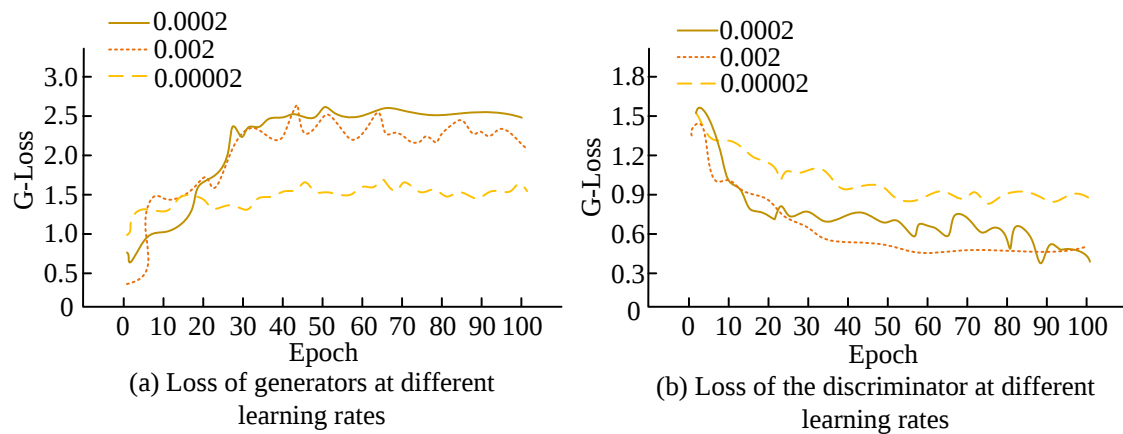


Figure 10: Loss variation of IDCGAN between generator and discriminator at different learning rates.

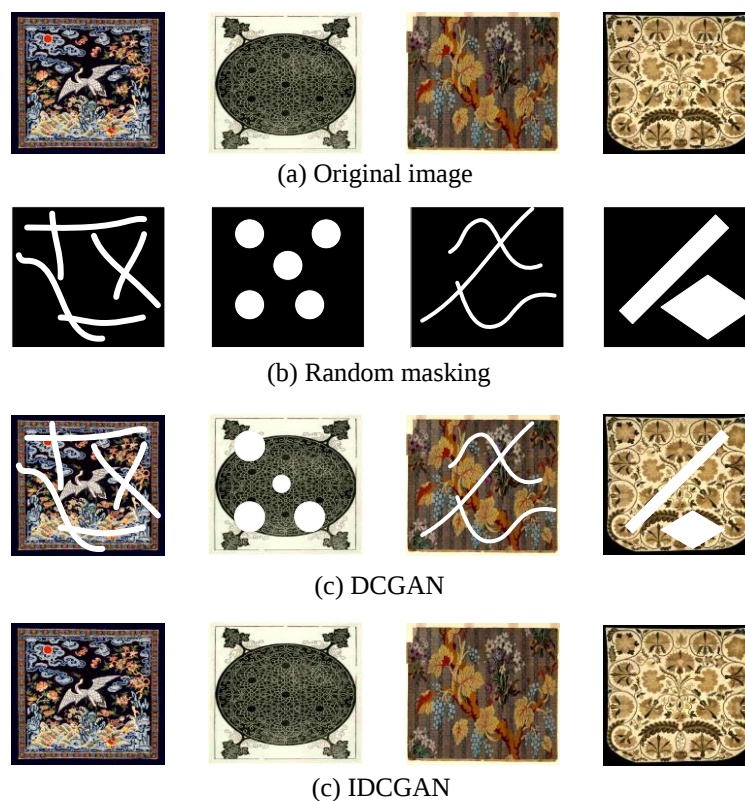


Figure 11: Repair effects of the model before and after the improvement (Source from: <https://colorhub.me/>).

In Fig.10 (a), the loss of the IDCGAN generator slowly increases with the growth of training cycles, and the curve with a learning rate of 0.00002 shows a low and stable loss value. The curves with learning rates of 0.002 and 0.0002 show higher loss values and larger fluctuations. In Fig.10 (b), the discriminator loss slowly decreases as the number of training cycles increases. The curve with a learning rate of 0.00002 decreases the fastest and tends to stabilize, indicating that a smaller learning rate helps the discriminator learn more effectively. In contrast, the curves with learning rates of 0.002 and 0.0002 exhibit significant fluctuations and higher loss values. Based on the comprehensive experimental data, this study ultimately sets the network learning rate of the IDCGAN model to 0.00002. To verify the impact of dilated convolutional layers, loss functions, and CBAM on

model performance, the repair effect of the improved model before and after random occlusion is compared, as shown in Fig.11.

Figs.11 (a) to (d) show the embroidery original image, images subjected to random occlusion, images restored by the DCGAN model, and images restored by the IDCGAN model. By comparing these images, the effectiveness of IDCGAN in handling different types of embroidery and varying degrees of occlusion can be demonstrated. IDCGAN can enhance the focus on key features, thereby enabling the restored image to largely restore the details and colors of the original image, effectively solving the problem of color non-uniformity. However, DCGAN's repair effect is not ideal when facing large-scale defects, and it cannot maintain good contextual consistency, resulting in poor repair performance. This discovery

validates the necessity of improving the DCGAN. To further test the effectiveness of the research model in embroidery image restoration, the Cycle-Consistency GAN (CCGAN), Conditional GAN (CGAN), and Stacked GAN (Stack-GAN) models are introduced for comparison. The test results of SSIM as the experimental indicator are shown in Fig.12.

Figs.12 (a) and (b) show the SSIM performance comparison of four models in two datasets. Both in the training and testing sets, the IDCGAN model performs the best, followed by Stack-GAN and CCGAN, while CGAN performs the worst. In the training set, the maximum SSIM values for CGAN, CCGAN, Stack-GAN, and the research model are 0.64, 0.72, 0.85, and 0.98, while in the testing set, they are 0.69, 0.78, 0.90, and 0.99. The above data indicates that the research model has significant advantages in maintaining image structure and quality. The reason behind this is that the dilated convolution technique effectively expands the receptive field, allowing it to capture richer contextual information in the image. In addition, CBAM further enhances the model's attention to key features by weighting important features in both channel and spatial dimensions. These improvements have

led to significant advantages of IDCGAN in embroidery image restoration. Finally, to confirm the resolution capability of the proposed model, this study also tests four models using image clarity as an indicator, as shown in Fig.13.

Figs.13 (a) to (d) show the clarity performance of CGAN, CCGAN, Stack-GAN, and IDCGAN models in the Yue embroidery image restoration task. Figure 13 (e) shows the clarity of the original image. The Yue embroidery restoration images generated by IDCGAN are visually very similar to the original images, and it is almost impossible to distinguish the quality differences with the naked eye. In contrast, there are significant differences between the restoration results of CGAN, CCGAN, and Stack-GAN and the original images. Especially for the restored images of the CGAN model, there is a significant decrease in clarity in comparison to the original images. In summary, the research model surpasses the comparative model in image resolution for Guangdong embroidery restoration, demonstrating its potential and advantages in embroidery image restoration processing.

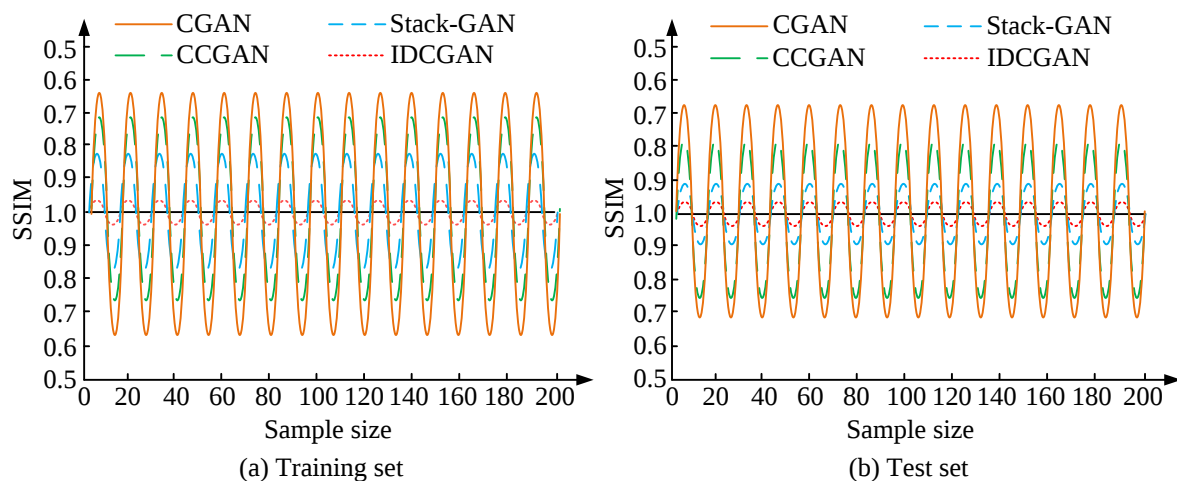


Figure 12: Schematic of SSIM test results for different models.

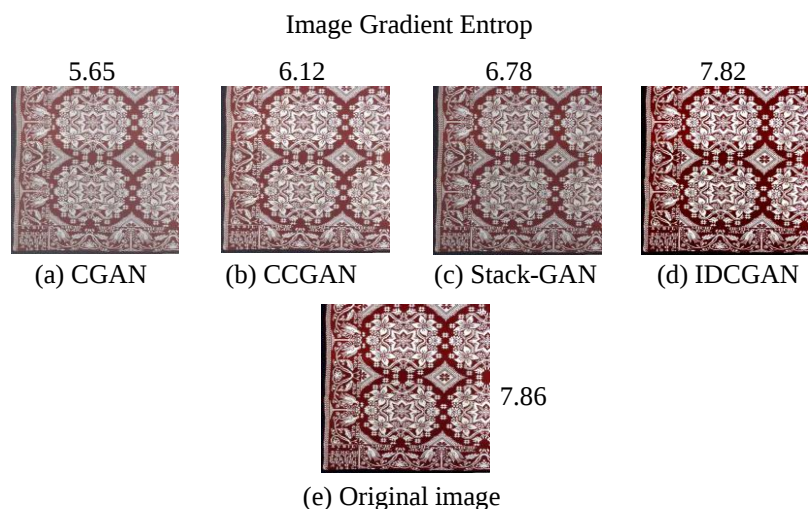


Figure 13: The clarity of restored images of Cantonese embroidery (Source from: <https://colorhub.me/photos/VXeo3>).

4 Conclusion

The study focused on the task of image restoration of ECR and innovatively constructed an ECR-ICM based on SPP-IDenseNet and an ECR-IRM based on the improved DCGAN. The experimental results showed that the SPP-IDenseNet model achieved an average prediction accuracy rate of over 80% for the embroidery images of eight styles. The IRM could enhance the focus on key features, thereby enabling the restored image to largely restore the details and colors of the original image, effectively solving the problem of uneven color. The SSIM value has reached 0.99. Furthermore, the research model could still maintain an excellent restoration effect even when dealing with large-area damaged embroidery images. The restored image of Cantonese embroidery generated was visually extremely similar to the original image, and it was almost impossible to distinguish the quality difference with the naked eye. The results show that the research model achieves innovation in technology and demonstrates significant advantages in practical applications. However, the research model also has certain limitations. On the one hand, the current models mainly target 2D embroidery images. At present, there is no adaptive research on complex 3D multi-level embroidery structures and heterogeneous multi material embroidery patterns, which limits their promotion and application in high-precision virtual restoration. On the other hand, due to the adoption of a deep generative network structure, the model has a certain dependence on computing resources during the training and inference stages. This may pose practical challenges in resource constrained cultural heritage conservation institutions or mobile deployments. Furthermore, for severely damaged or extremely blurry images, there is still a certain risk of distortion in the structural reconstruction of the research model. Future research can be carried out in the following directions: (1) Expansion of model generalization ability: By integrating 3D reconstruction and multimodal input, the restoration ability of 3D ECR can be enhanced; (2) Enhanced multi-material adaptability: Material perception module or style transfer mechanism can be introduced to achieve texture simulation and reconstruction of heterogeneous embroidery materials; (3) Lightweight deployment optimization: By applying techniques such as model pruning, quantization, and distillation, the network structure is compressed to adapt to edge devices or mobile terminal applications. Overall, the research method provides a feasible and effective technological path for ECR digital protection, which is expected to have practical applications in digital museum construction, virtual restoration of cultural heritage, and reconstruction of cultural creative models.

References

- [1] Xinyang Guan, Likang Luo, Honglin Li, He Wang, Chen Liu, Su Wang, and Xiaogang Jin. Automatic embroidery texture synthesis for garment design and online display. *The Visual Computer*, 37(9):2553-2565, 2021. <https://doi.org/10.1007/s00371-021-02216-0>
- [2] Xiaoli Fu, and Niwat Angkawisittpan. Detecting surface defects of heritage buildings based on deep learning. *Journal of Intelligent Systems*, 33(1):163-169, 2024. <https://doi.org/10.1515/jisys-2023-0048>
- [3] Ana M. Maitin, Alberto Nogales, Emilio Delgado-Martos, Giovanni Intra Sidola, Carlos Pesqueira-Calvo, Gabriel Furnieles, and Álvaro J. García-Tejedor. Evaluating activation functions in GAN models for virtual inpainting: A path to architectural heritage restoration. *Applied Sciences*, 14(16):6854-6854, 2024. <https://doi.org/10.3390/app14166854>
- [4] Alessandro Bombini, Fernando García-Avello Bofías, Chiara Ruberto, and Francesco Taccetti. A cloud-native application for digital restoration of cultural heritage using nuclear imaging: THESPIAN-XRF. *Rendiconti Lincei. Scienze Fisiche e Naturali*, 34(3):867-887, 2023. <https://doi.org/10.1007/s12210-023-01174-0>
- [5] Kanghyeok Ko, Taesun Yeom, and Minhyeok Lee. SuperstarGAN: Generative adversarial networks for image-to-image translation in large-scale domains. *Neural Networks*, 162(42):330-339, 2023. <https://doi.org/10.1016/j.neunet.2023.02.042>
- [6] Praveen Kumar, and Varun Gupta. Restoration of damaged artworks based on a generative adversarial network. *Multimedia Tools and Applications*, 82(26):40967-40985, 2023. <https://doi.org/10.1007/s11042-023-15222-2>
- [7] Wenjun Zheng, Benpeng Su, Ruiqi Feng, Xihua Peng, and Shanxiong Chen. EA-GAN: Restoration of text in ancient Chinese books based on an example attention generative adversarial network. *Heritage Science*, 11(1):55-62, 2023. <https://doi.org/10.1186/s40494-023-00882-y>
- [8] Mihai Bundeia, and Gabriel Mihail Danciu. Pneumonia image classification using DenseNet architecture. *Information*, 15(10):611-619, 2024. <https://doi.org/10.3390/INFO15100611>
- [9] Sherly Alphonse, S. Abinaya, and Nishant Kumar. Pain assessment from facial expression images utilizing Statistical Frei-Chen Mask (SFCM)-based features and DenseNet. *Journal of Cloud Computing*, 13(1):142-148, 2024. <https://doi.org/10.1186/S13677-024-00706-9>
- [10] Chunyang Zhu, Lei Wang, Weihua Zhao, and Heng Lian. Image classification based on tensor network DenseNet model. *Applied Intelligence*, 54(8):6624-6636, 2024. <https://doi.org/10.1007/S10489-024-05472-4>
- [11] S. Deepa, Beevi S. Zulaikha, Laxman L. Kumarwad, and Sabbineni Poojitha. Namib beetle firefly optimization enabled DenseNet architecture for hyperspectral image segmentation and classification. *International Journal of Image & Data Fusion*, 15(2):190-213, 2024. <https://doi.org/10.1080/19479832.2023.2284781>
- [12] Suresh Samudrala, and C. Krishna Mohan. Semantic segmentation of breast cancer images using DenseNet with proposed PSPNet. *Multimedia Tools*

- and Applications, 83(15):46037-46063, 2023. <https://doi.org/10.1007/S11042-023-17411-5>
- [13] M. Karthikeyan, and D. Raja. Deep transfer learning enabled DenseNet model for content-based image retrieval in agricultural plant disease images. Multimedia Tools and Applications, 82(23):36067-36090, 2023. <https://doi.org/10.1007/S11042-023-14992-Z>
- [14] Babu Rajendra Prasad, and Dr. B. Sai Chandana. Human face emotions recognition from thermal images using DenseNet. International Journal of Electrical and Computer Engineering Systems, 14(2):155-167, 2023. <https://doi.org/10.32985/IJECES.14.2.5>
- [15] Ning Wang, Yanzheng Chen, Yi Wei, Tingkai Chen, and Hamid Reza Karimi. UP-GAN: Channel-spatial attention-based progressive generative adversarial network for underwater image enhancement. Journal of Field Robotics, 41(8):2597-2614, 2024. <https://doi.org/10.1002/ROB.22378>
- [16] Noa Barzilay, Tal Berkovitz Shalev, and Raja Giryes. MISS GAN: A Multi-IlluStrator style generative adversarial network for image to illustration translation. Pattern Recognition Letters, 151(16):140-147, 2021. <https://doi.org/10.1016/J.PATREC.2021.08.006>
- [17] Manuel Domínguez-Rodrigo, Ander Fernández-Jaúregui, Gabriel Cifuentes-Alcobendas, and Enrique Baquedano. Use of generative adversarial networks (GAN) for taphonomic image augmentation and model protocol for the deep learning analysis of bone surface modifications. Applied Sciences, 11(11):5237-5247, 2021. <https://doi.org/10.3390/APP11115237>
- [18] Aram You, Jin Kuk Kim, Ik Hee Ryu, and Tae Keun Yoo. Application of generative adversarial networks (GAN) for ophthalmology image domains: A survey. Eye and Vision, 9(1):6-16, 2022. <https://doi.org/10.1186/S40662-022-00277-3>
- [19] Zhiguo Xiao, Jia Lu, Xiaokun Wang, Nianfeng Li, Yuying Wang, and Nan Zhao. WCE-DCGAN: A data augmentation method based on wireless capsule endoscopy images for gastrointestinal disease detection. IET Image Processing, 17(4):1170-1180, 2022. <https://doi.org/10.1049/IPR2.12704>
- [20] Betelhem Zewdu Wubineh, Andrzej Rusiecki, and Krzysztof Halawa. Classification of cervical cells from the Pap smear image using the RESDCGAN data augmentation and ResNet50V2 with self-attention architecture. Neural Computing and Applications, 18(24):1-15, 2024. <https://doi.org/10.1007/S00521-024-10404-X>