Incomplete handwritten Dongba character image recognition by multiscale feature restoration

Xiaojun Bi[1,2] and Yanlong Luo[3*]

Abstract

Incomplete handwritten Dongba character often appears in heritage documents and its recognition is significant for heritage and philology. However, all previous methods always suppose that a complete Dongba character is used as input, and thus fail to achieve satisfactory performance when applied to incomplete Dongba character recognition. In this paper, an end-to-end network (DB2RNet) is proposed for incomplete handwritten Dongba character image recognition by multiscale feature restoration. Specifically, we first develop datasets that contain different levels of incomplete Dongba characters. A restoration module is proposed to restore the input incomplete Dongba character, and then a recognition module is employed to recognize Dongba character. By introducing an inter-module residual connection between the restoration module and recognition module, the DB2RNet can strengthen feature information transmission and boost the recognition performance. In addition, novel multiscale feature blocks are introduced, which can provide more effective texture and contextual information transmission for Dongba character image restoration, and thus yielding better restoration effects and better recognition results. Extensive experiments are conducted on Dongba character, Chinese character and Oracle character datasets and validate the effectiveness, superiority and robustness of our methods. Experiments results demonstrate that our proposed DB2RNet achieves competitive Dongba character restoration and recognition performance and outperforms the current state-of-the-art methods.

Keywords Dongba character, Image recognition, Image restoration, Inter-module residual connection, Multiscale feature

Introduction

Dongba character is a kind of hieroglyphic character created by the ancestors of the Naxi national minority in China. It is recognized as the only hieroglyphic character still in use in the world today, and it is a cultural treasure of great academic value [1]. Dongba character has been listed as "Memory of the world" by the United Nations Educational Scientific and Cultural Organization [2]. However, due to the poor preservation conditions of the Dongba ancient heritage documents, many documents were damaged and lots of Dongba characters were broken. Figure 1 shows ten incomplete Dongba characters from an ancient heritage document. All previous Dongba character recognition methods always suppose that a complete Dongba character is used and incomplete Dongba character recognition is more difficult and no relevant research has been proposed yet. Therefore, in this paper, we will focus on studying incomplete Dongba character recognition by multiscale feature restoration.

*Correspondence:
Yanlong Luo
lyanlong@hrbeu.edu.cn
[1] Key Laboratory of Ethnic Language Intelligent Analysis and Security Governance of MOE, Minzu University of China, Beijing 100081, China
[2] School of Information Engineering, Minzu University of China, Beijing 100081, China
[3] College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China

Fig. 1 Ten incomplete Dongba characters from an ancient heritage document. The strokes enclosed within the red rectangle constitute an example of incomplete Dongba character

Lots of complete handwritten Dongba character recognition (HDCR) algorithms have been introduced for enhancing the recognition accuracy. Traditional methods typically encompass three distinct stages: pretreatment, characteristic extraction and recognition. Binary processing, shape normalization and wavelet transform [3] are generally employed pretreatment methods [4]. A combination of topological features processing and projection method shows a good performance on HDCR [5]. The statistical features such as coarse grid [6] are more discriminable than structural features that mainly analyze parts of Dongba characters. The most usually adopted methods of classifiers include support vector machines (SVM) [7–9], random forests [10–12] and k-nearest neighbor [13, 14].

With the explosion of deep learning, a diversity of deep learning based methods have been designed to solve heritage recognition problems [15, 16] as well as HDCR problems [17, 18]. Multi-scale features fusion method of three separate branch networks has been proposed for HDCR [17]. Improved ResNet [19, 20] are proposed to handle HDCR problems. Luo et al. [21] demonstrate that features of the shallower layer include both more detailed and spatial information, and the information is very helpful for similar Dongba character recognition.

All the mentioned above studies concentrate on complete handwritten Dongba character recognition and complete HDCR have achieved much progress in recent years. However, these recent algorithms for Dongba character recognition are specialised for perfect/complete characters that are typically created by imitating characters in authoritative dictionaries [22, 23]. On the contrary, most real Dongba characters in ancient heritage documents usually exhibit incompletion. As far as we know, no recognition methods have been applied on incomplete handwritten Dongba character. It can be easily summarized from Fig. 1 that these incomplete characters lack crucial information. As a result, the existing

excellent image recognition methods do not perform well. Therefore, it's not hard for us to think about that it would be easier to recognize them if incomplete Dongba characters have been restored.

Although complete Dongba character recognition models based on neural networks have performed progress in the past years, no research paid attention to incomplete Dongba character recognition, let alone incomplete Dongba character restoration. Consequently, we propose to recognize handwritten Dongba character with incompletion through an attached restoration procedure by an end-to-end network. Nevertheless, because no public incomplete Dongba character image dataset is obtainable presently, incomplete handwritten Dongba character restoration and recognition seem much tricky. Therefore, this paper begins with constructing novel incomplete handwritten Dongba character image datasets via complete handwritten Dongba character dataset [21] and most widely used random mask dataset which contains different mask ratio images [24].

It is not hard to come up with the idea of first obtaining the completed Dongba character through a restoration model, and then recognizing Dongba character through a recognition model using the restored image. Although no incomplete Dongba character recognition and restoration methods have been proposed. Wan et al. [25], Liu et al. [26] and Chen et al. [27] have make progress in end-to-end network to solve sketch restoration-to-recognition problem and vehicle logo image restoration-to-recognition problem, which provide us with ideas to solve the incomplete Dongba character restoration-to-recognition problems.

SketchGAN [26] uses a cascade Encode-Decoder network to complete the input sketch in an iterative manner, and employ an auxiliary sketch recognition task to recognize the completed sketch. Joint-Caps-Net [27] proposes a joint framework that simultaneously performs both vehicle logo restoration and recognition within a shared

deep neural network architecture. ADFRNet [25] restores the imperfect sketch by the proposed attention-based feedback restoration loop, and then sent to a classifier. Although [25, 26] and [27] have made progress to solve image restoration and recognition problem. Two significant issues that have great impacts on end-to-end image restoration-to-recognition are still ignored by them.

On the one hand, they only fed the restoration images into the recognition model, leading to inadequate transmission of feature information and poor performance of the recognition model. It has been widely demonstrated that residual connection can strengthen feature propagation [28–30]. To solve this problem, we propose a novel end-to-end incomplete Dongba character image recognition method by multiscale feature restoration (DB2RNet), which can boost the performance of the recognition model. As shown in Fig. 2, DB2RNet not only fed the restoration image into the recognition model but also with the initial incomplete Dongba character. The residual connection between the restoration module and recognition module (inter-module residual connection) can strengthen feature information transmission and boost the recognition performance.

On the other hand, incomplete handwritten Dongba characters are short of texture and contextual information, and are commonly known to be ambiguous to be restored. What's more, it is demonstrated that the Dongba character have more detailed information [21]. Therefore, it is necessary to extract more detailed and global information to achieve a restoration of texture and contextual information. It is experimentally validated that the features of the deeper layer contain more global and semantic information, while the features of the shallower layer contain both more spatial and detailed information [21, 31]. Therefore, a novel multiscale feature block (MFB) is proposed in the restoration module, which can make use of the different size output feature maps of different encoders, and fuses these feature maps to get the Dongba character features which contain multiscale information. MFB provides rich fusion information for the following decoder layers, and can provide more effective texture and contextual information transmission for Dongba character image restoration, yielding better restoration effects and better recognition results.

To estimate the superiority of the proposed DB2RNet at recognizing and restoring incomplete Dongba characters, extensive experimental studies have been carried out. Ablation studies and comparisons with other state-of-the-art methods are carried out on the constructed datasets, and the sufficient qualitative and quantitative results demonstrate the superiority of the proposed method. In summary, several novel contributions are presented as follows:

1) To the best of our knowledge, we are the first to solve the problem of incomplete Dongba character recognition by restoration.
2) A novel end-to-end network architecture named DB2RNet is proposed, which handles both incomplete Dongba character image recognition and restoration. And the restoration task benefits the recognition task.
3) Inter-module residual connection is proposed to boost recognition performance. And multiscale feature restoration module is proposed to extract more texture and contextual information for restoration.
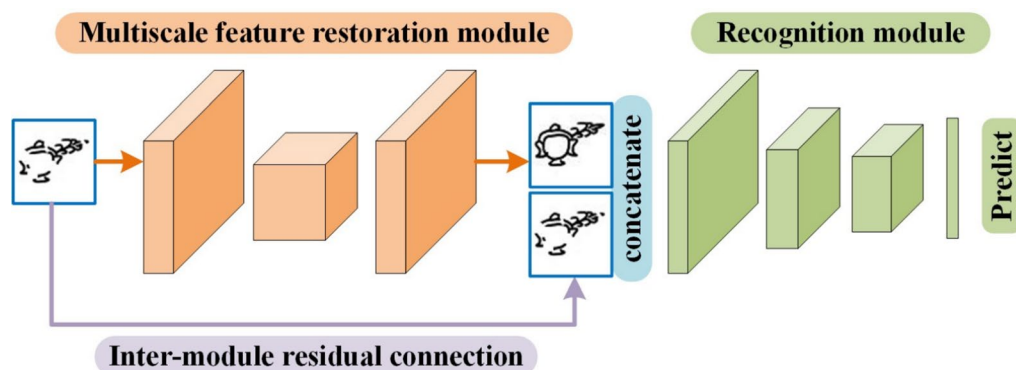
Fig. 2 Brief overview of the proposed DB2RNet. It can be seen that the incomplete input and its restoration version are concatenated, which is a residual connection structure (inter-module residual connection). And this structure can boost the performance of the recognition module. The orange blocks, green blocks and purple line represent brief structure of the Multiscale feature restoration module, Recognition module and Inter-module residual connection, separately, as well as their employed position. Detailed structure of the Multiscale feature restoration module, Inter-module residual connection and Recognition module are illustrated in sections Multiscale feature restoration module and Sect. Inter-module residual connection and recognition module

4) Experiments are conducted on four different ratios incomplete Dongba character datasets, CASIA-HWDB1.1 and OBC306 datasets. Experimental results demonstrate the superiority and robustness of our DB2RNet both in incomplete Dongba character recognition and restoration.

Methods

In this study, we mainly focus on two significant problems. On the one hand, for the reason that previous end-to-end image restoration-to-recognition methods only fed the restoration image into the recognition module, overlooking the importance of the initial incomplete image. And thus leads to poor performance of the recognition module. Therefore, we propose a inter-module residual connection method. By concatenating the initial incomplete Dongba character and its restoration version, better recognition results will be accomplished. On the other hand, incomplete handwritten Dongba characters are short of texture and contextual information, and are commonly known to be ambiguous to be restored. Therefore, a novel multiscale feature block is adopted in the restoration module, which can make use of the different size output feature maps of different encoders, and can fuse these feature maps to get the Dongba character features which contain multiscale information.

In general, an incomplete handwritten Dongba character image recognition by multiscale feature restoration method (DB2RNet) is proposed, which can not only solve the problems of the previous end-to-end image restoration-to-recognition methods, but also solve the short of texture and contextual information problem.

The architecture of DB2RNet

The proposed DB2RNet is mainly composed of two components, i.e., restoration module and recognition module, as shown in Fig. 2. The DB2RNet adopts an incomplete

Dongba character as input, and restores a corresponding complete Dongba character and outputs its recognition result. To be specific, the restoration module is used to restore complete Dongba character based on the original incomplete input character. The recognition module is designed to recognize the original incomplete input character by concatenating the original input and the output of the restoration module. All detailed structures of the restoration module and recognition module will be given in the following sections.

Multiscale feature restoration module

The restoration module plays a significant role in our DB2RNet. Better restoration quality results in less difficulties in recognizing Dongba character. The architecture of the restoration module is shown in Fig. 3.

Due to the lack of texture and contextual information in incomplete handwritten Dongba characters, their restoration is often ambiguous. Consequently, extracting more detailed and global information is crucial for restoring both texture and contextual details. Experimental evidence suggests that deeper layer features encompass more global and semantic information, while shallower layer features contain both spatial and finer details [21, 31].

Therefore, a Multiscale Feature Block (MFB) (illustrated in Fig. 4) is introduced. This block leverages the varying sizes of output feature maps from different encoders and merges these three feature maps to extract Dongba character features that encompass multiscale information. By providing decoders with rich texture and contextual information, MFB enables the transmission of more effective restoration features, ultimately leading to improved restoration effects and recognition outcomes. The various output feature maps from the three encoders can be computed as follows:
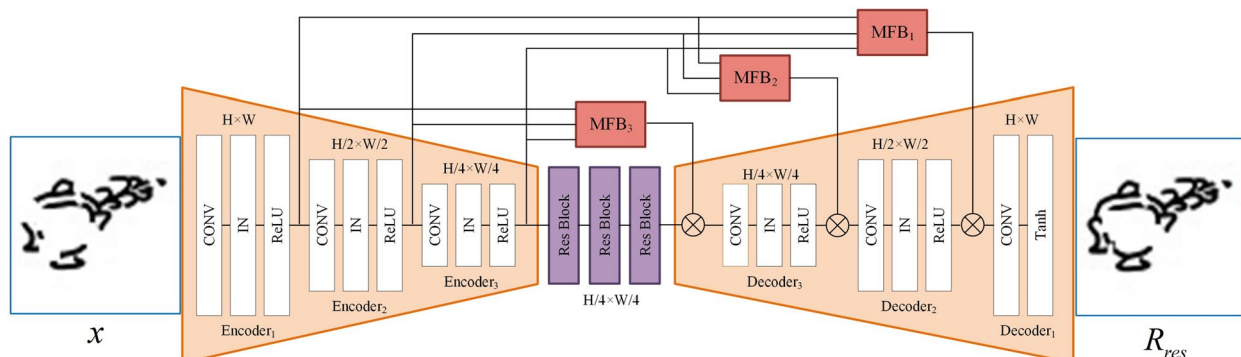

Fig. 3 Overview of the multiscale feature restoration module of the proposed DB2RNet. MFB_i ($i \in \{1, 2, 3\}$) indicate multiscale feature block shown in Fig. 4. IN represents instance normalization
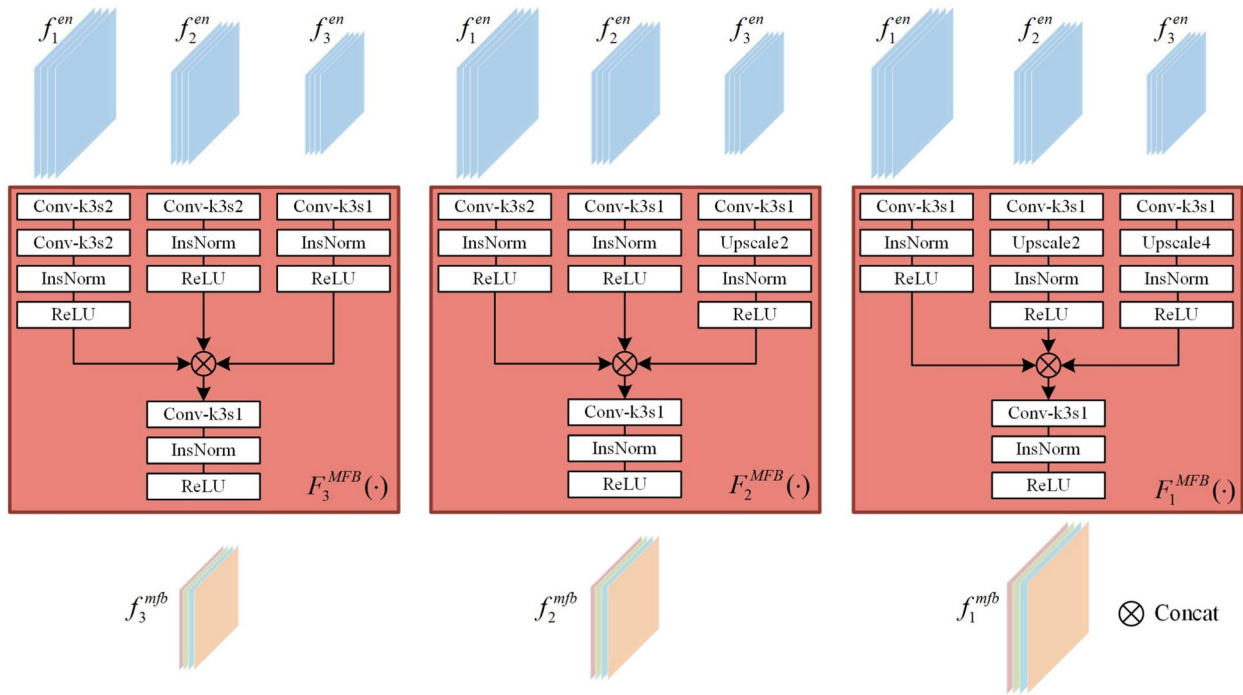
Fig. 4 Overview of the multiscale feature block in Fig. 3. "Conv-kxsy" represents a convolutional kernel size of x and a stride of y. "Upscalez" indicates that the scale of the feature map is enlarged by a factor of z compared to its current size. "InsNorm" is the Instance Normalization. "ReLU" is the ReLU activation function

$$\begin{aligned} f_1^{en} &= F_1^{en}(x) \\ f_2^{en} &= F_2^{en}(f_1^{en}) \\ f_3^{en} &= F_3^{en}(f_2^{en}) \end{aligned} \tag{1}$$

where x is the input incomplete Dongba character and f_i^{en} ($i \in \{1, 2, 3\}$) refers to the output of Encoder$_i$ ($i \in \{1, 2, 3\}$), and $F_i^{en}(x)$ ($i \in \{1, 2, 3\}$) is the function of Encoder$_i$.

Specifically, the MFB proposed in this paper integrates features of varying sizes from three encoder layers to capture rich multiscale information for Dongba characters. Larger feature maps typically encode finer details such as corners and edges, while smaller feature maps capture more global semantic information like overall shape. Each MFB$_i$ ($i \in \{1, 2, 3\}$) combines the output features from these three encoders, resulting in multiscale fusion features that encompass both detailed and global information.

Due to the varying feature sizes outputted by the encoder layers, we have designed three distinct MFB$_i$ ($i \in \{1, 2, 3\}$) to ensure that the fused features align with the feature size of the corresponding decoder layer. Each MFB$_i$ receives three differently-sized feature maps from the encoders as inputs and produces fused features. This process can be expressed as follows:

$$f_i^{mfb} = F_i^{MFB}(f_1^{en}, f_2^{en}, f_3^{en}) \tag{2}$$

where f_i^{mfb} is the output of MFB$_i$ ($i \in \{3, 2, 1\}$), and f_1^{en}, f_2^{en}, f_3^{en} are computed by Eq. (1).

Here, we delve into the detailed calculation process of MFB$_3$ as an illustrative example, as depicted in the left structure of Fig. 4. Initially, we adjust the sizes of all features f_1^{en}, f_2^{en} and f_3^{en} to match the target size of f_3^{en}. Specifically, f_1^{en} undergoes two convolutional layers with a kernel size of 3 and a stride of 2, resulting in a size reduction to 1/4. We apply instance normalization and the ReLU activation function after the last convolutional layer. Similarly, f_2^{en} passes through a single convolutional layer with a kernel size of 3 and a stride of 2, halving its size. On the other hand, f_3^{en} remains unchanged after passing through a convolutional layer with a kernel size of 3 and a stride of 1. Next, we concatenate the transformed features. Finally, we obtain the output f_3^{mfb} by applying a convolutional layer with a kernel size of 3 and a stride of 1 to the concatenated features. As with the previous layers, we employ instance normalization and the ReLU activation function. The calculation processes for MFB$_2$ and MFB$_1$ are analogous to that of MFB$_3$ but differ in the specific operations used for feature size transformation. For instance, in some cases, we might employ a

combination of a convolutional layer with a kernel size of 3 and a stride of 1 followed by an upsampling operation. By following this process, we ensure that the fused features not only align with the corresponding decoder layer's feature size but also encapsulate multiscale information crucial for effective Dongba character restoration.

The multiscale feature block (MFB) proposed in this paper is designed to fuse the output features of the three encoders, enabling the extraction of both detailed and global information. This fusion process aims to restore texture and contextual information effectively, crucial for Dongba character image restoration and recognition. By leveraging multiscale features, the MFB ensures a more comprehensive understanding of the input image, thus providing more effective information transmission during the restoration process. This enhanced information flow is pivotal in facilitating accurate recognition of Dongba characters.

The detailed structure and hyperparameters of the proposed restoration module are shown in Table 1. We use Input s &c and Output s &c to represent the input and output size and channel of the feature map of current layer. Ref p is padding number before current Conv and Up s indicates up-sample. K, S, and P denote the kernel size, stride, and padding of operators, respectively. Norm indicates whether Instance normalization layer is used after each convolution layer. Activation indicates the nonlinear function after the layer. At the DeConv1 layer, we use Tanh activation to map each pix of the output feature map to [0, 1].

Inter-module residual connection and recognition module

Although previous methods [25–27] have made progress to solve image restoration and recognition problem. They have overlooked the importance of the original incomplete image in the recognition module, and they only fed the restoration image into the recognition model, leading to inadequate transmission of feature information and poor performance of the recognition model. It has been widely demonstrated that residual connection can strengthen feature propagation [28–30]. Therefore, to solve this problem, we propose a novel inter-module residual connection, which can assist the recognition module in achieving better performance.

Specifically, the input incomplete Dongba character is firstly fed into the restoration module, and a restoration version of the input character is obtained, which can be computed as:

$$R_{res} = F_{res}(x) \tag{3}$$

where x is the input incomplete Dongba character, and R_{res} refers to the restored character from the restoration module $F_{res}(\cdot)$. Then, the original incomplete character x needs to be firstly concatenated with the restored character R_{res} in the channel dimension, and then the concatenated feature map is fed to the recognition module, which can be computed as:

$$y = F_{rec}(\text{Concat}(x, R_{res})) \tag{4}$$

where $\text{Concat}(\cdot)$ stands for the operation of concatenating along the channel dimension, and y refers to the predicted label of the input incomplete Dongba character x from the recognition module $F_{rec}(\cdot)$.

Detailed structure and hyperparameters of the recognition module is shown in Table 2. The recognition module structure is inspired by the basic network of MAAN [21] and it should be noted that any recognition models can be employed in our DB2RNet. The recognition model with better performance will also perform better in our DB2RNet. It should also be noted that the In c = 2 in the Conv1 of Table 2 represents the concatenated feature map of the original incomplete character and the final output of the restoration module and this strategy is implemented by the inter-module residual connection.

Loss functions

Loss functions are significant factor in training a end-to-end model which can simultaneously handle incomplete Dongba character image recognition and restoration tasks. We utilize the following loss functions for training the proposed end-to-end DB2RNet.

MAE loss: With the help of restoration module, we restore a complete version of the Dongba character

Table 1 Detailed structure and hyperparameters of the proposed restoration module

Layer	In s	Out s	In c	Out c	Ref p	Up s	K	S	P	Norm	Activation
Encoder1	64	64	1	64	3	–	7	1	0	IN	ReLU
Encoder2	64	32	64	128	–	–	4	2	1	IN	ReLU
Encoder3	32	16	128	256	–	–	4	2	1	IN	ReLU
Decoder3	16	32	256	128	–	2	3	1	1	IN	ReLU
Decoder2	32	64	128	64	–	2	3	1	1	IN	ReLU
Decoder1	64	64	64	1	3	–	7	1	0	–	Tanh

Table 2 Detailed structure and hyperparameters of the the recognition module. Meanings of the abbreviations are the same as Table 1

Layer	In s	Out s	K, S, P, In c, Out c
Conv1	64	64	$\begin{bmatrix} 3,1,1,\mathbf{2},64 \\ 3,1,1,64,64 \end{bmatrix} \times 2$
Pooling1	64	32	2×2 max pool, stride 2
Conv2	32	32	$\begin{bmatrix} 3,1,1,128,128 \\ 3,1,1,128,128 \end{bmatrix} \times 2$
Pooling2	32	16	2×2 max pool, stride 2
Conv3	16	16	$\begin{bmatrix} 3,1,1,256,256 \\ 3,1,1,256,256 \end{bmatrix} \times 2$
Pooling3	16	8	2×2 max pool, stride 2
Conv4	8	8	$\begin{bmatrix} 3,1,1,512,512 \\ 3,1,1,512,512 \end{bmatrix} \times 2$
Pooling4	8	4	2×2 max pool, stride 2
Conv5	4	4	$\begin{bmatrix} 3,1,1,512,512 \\ 3,1,1,512,512 \end{bmatrix} \times 2$
Pooling5	4	1	Global average pool
Fully-connected layer, Softmax			

R_{res}. To enforce the restoration module generating the expected complete Dongba character, we utilize the Mean Absolute Error (MAE) of the difference between the ground truth character and the generated character as loss function, which can be calculate by:

$$\mathcal{L}_{MAE} = \|R_{res} - R\|_1 \tag{5}$$

where R_{res} is the restored character from the restoration module. R represents the ground truth Dongba character.

SSIM loss: However, MAE loss is based on pixel-by-pixel comparisons, without considering human visual perception, let alone human aesthetics. Therefore, the quality of restoration character needs to be further improved. Therefore, we also adopt SSIM loss to ensure the realistic property, which can be calculate by:

$$\mathcal{L}_{SSIM} = 1 - SSIM(R_{res}, R) \tag{6}$$

SSIM is an increasing function with a range of [0, 1], therefore, 1-*SSIM* can be used as a loss function.

Recognition loss: The input into the recognition module is restored Dongba character with its corresponding incomplete version. To be more specific, the restored and incomplete Dongba character are firstly concatenated in the channel dimension, and then the concatenated feature maps are fed to the recognition module. Just like the original incomplete Dongba character input plays a significant role in the restoration module, it also has substantial impact on the recognition module. The recognition loss is Softmax loss:

$$\mathcal{L}_{REC} = - \sum_{i=1}^{K} y_i \log(p_i) \tag{7}$$

where K denotes total categories of dataset, y_i stands for the label, and p_i is the predicted probability.

Total loss: The above defined three loss functions play diverse duties for the whole DB2RNet. Specifically, the restoration loss and the adversarial loss aim to enhance the quality of the restored Dongba character, and the recognition loss guarantee learning significant features for recognition. We employ the three functions in a weighted fashion to acquire the final loss function, which can be computed as:

$$\mathcal{L}_{TOTAL} = \alpha \mathcal{L}_{MAE} + \beta \mathcal{L}_{SSIM} + \gamma \mathcal{L}_{REC} \tag{8}$$

where α, β, γ are hyperparameters for weighting different loss functions. We empirically set $\alpha = 0.5$, $\beta = 0.35$, $\gamma = 0.15$, which give good performance.

Datasets and evaluation metrics

Datasets: All the previous methods focus on complete/perfect handwritten Dongba character recognition and no Dongba character restoration methods have been proposed [5, 17, 18, 21]. However, lots of incomplete Dongba character exist in Dongba ancient heritage documents. Therefore, we construct incomplete Dongba character dataset to enable research on incomplete Dongba character recognition and restoration.

We construct incomplete Dongba character dataset based on the complete Dongba character database [21], which consists of 445,273 images with 1,404 classes. Dongba character instances are shown in the first row of Fig. 5. Mask images of different ratios are firstly proposed [24] and most image inpainting methods adopt these mask images. To make our method more credible and convincing, we use 4 categories of masks with different hole-to-image area ratios of the irregular mask [24]. Four different mask ratio images: (0.01, 0.1], (0.1, 0.2], (0.2, 0.3] and (0.3, 0.4] are shown in the second, fourth, sixth and eighth rows of Fig. 5, respectively. The third, fifth, seventh and ninth rows represent four different ratios of the incomplete Dongba characters. The incomplete Dongba character are generated by element-wise addition of the ground truth images and mask images. That is, the white area of the mask image obscures the corresponding ground truth image area and the black area of the mask image do not change the corresponding ground truth image area. It can be easily seen that the larger level incomplete characters loss more information, and they are more difficult to restore and recognize. For the reason that all Dongba characters in Dongba character database and all masks in mask database are different
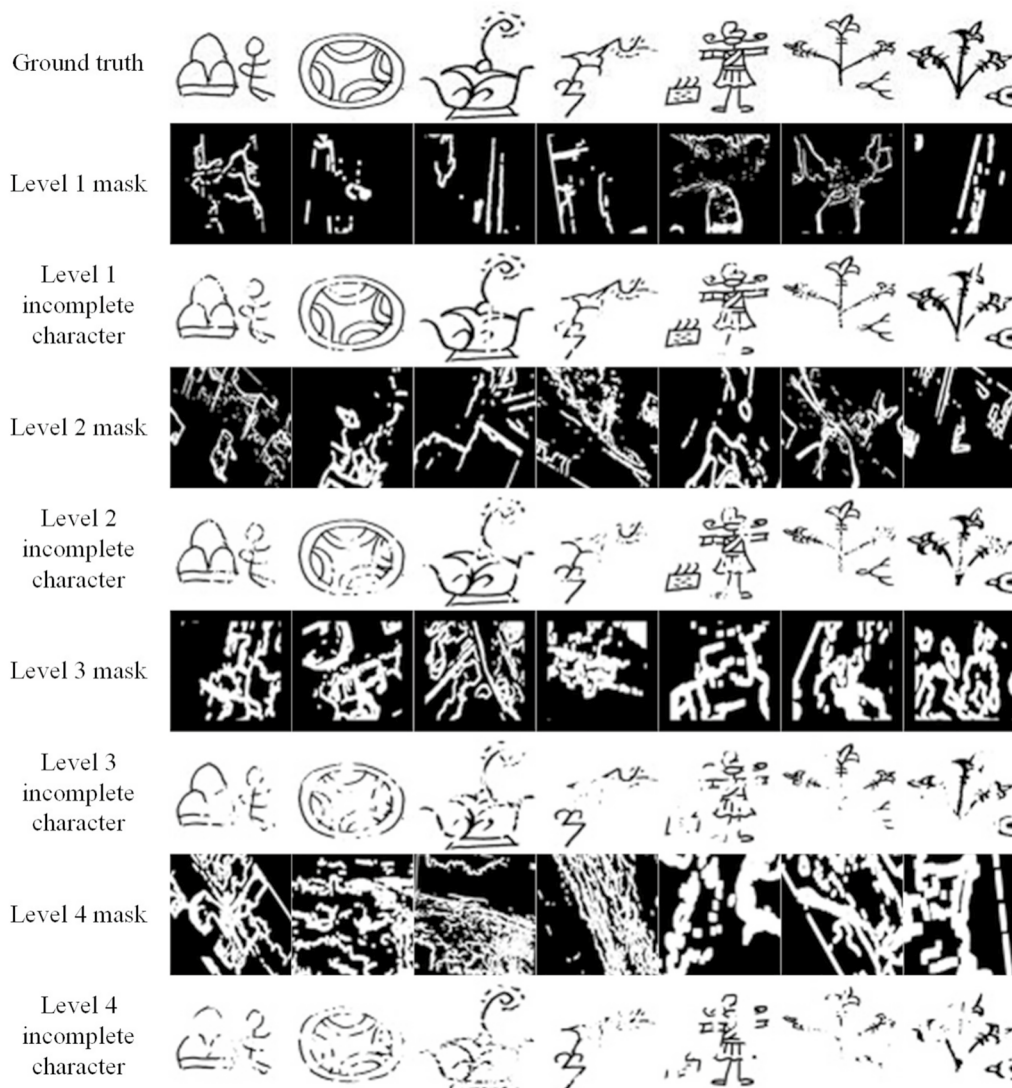
Fig. 5 Examples of complete handwritten Dongba characters [21] (the first row), four different mask ratio images [24] (the second, fourth, sixth and eighth rows), and our constructed four different mask ratio incomplete Dongba characters (the third, fifth, seventh and ninth rows)

from each other, the generated incomplete Dongba characters must be different. Therefore, all the experiment results are convincing.

Evaluation metrics: We utilize recognition accuracy as the measurement to evaluate the recognition performance. What's more, to evaluate the restoration performance, the peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) index are employed to test each complete/perfect Dongba character with its corresponding incomplete version and restored version. If the PSNR, SSIM and accuracy of restored Dongba character are higher than that of the original incomplete Dongba character and the direct recognition accuracy, it means that the proposed DB2RNet both improves the image

quality and also benefits recognition of the incomplete Dongba character.

On the one hand, PSNR is an error-sensitive image quality evaluation index, which reflects the pixel-level differences between the images before and after restoration, which can be computed as:

$$PSNR = 10 \cdot \lg \left(\frac{MAX_I^2}{MSE} \right) \tag{9}$$

where *MSE* stands for mean squared error, which represents the average squared difference between two images. MAX_I indicates the maximum numerical value of the image colors.

Table 3 The performance of our proposed topology with different baseline networks on the four different mask level testing datasets

Mask level dataset	Baseline networks	Accuracy	PSNR	SSIM
Level 1 incomplete Dongba character	Original incomplete character	–	33.0841	0.9805
	Direct recognition	97.14%	–	-
	No inter-module residual connection	96.27%	33.1572	0.9835
	No multiscale feature block	97.84%	32.8954	0.9813
	DB2RNet(Ours)	**98.25**%	**33.9187**	**0.9872**
Level 2 incomplete Dongba character	Original incomplete character	–	24.2027	0.9465
	Direct recognition	92.85%	–	–
	No inter-module residual connection	89.24%	25.3192	0.9647
	No multiscale feature block	94.50%	23.7284	0.9554
	DB2RNet(Ours)	**94.69**%	**26.7415**	**0.9688**
Level 3 incomplete Dongba character	Original incomplete character	–	19.7016	0.8991
	Direct recognition	88.84%	–	–
	No inter-module residual connection	88.59%	22.2379	0.9386
	No multiscale feature block	89.82%	21.7369	0.9334
	DB2RNet(Ours)	**90.75**%	**22.5452**	**0.9420**
Level 4 incomplete Dongba character	Original incomplete character	–	18.4215	0.8688
	Direct recognition	81.61%	–	–
	No inter-module residual connection	80.63%	20.4858	0.9112
	No multiscale feature block	83.37%	20.2383	0.9076
	DB2RNet(Ours)	**84.29**%	**20.6328**	**0.9139**

The best results are shown in bold

SSIM, on the other hand, is used to measure the similarity in brightness, contrast, and structure between two images, which can be computed as:

$$SSIM = \frac{(2\mu_x\mu_y + c_1)(\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \qquad (10)$$

where μ_x and μ_y represent the mean values of images x and y, respectively. σ_x and σ_y stand for the standard deviations of images x and y, respectively. σ_{xy} is the covariance between images x and y. c_1 and c_2 are default constants used to avoid computational errors when the denominator in the formula approaches zero.

Experimental setup

The proposed DB2RNet model is carried out by employing the PyTorch framework. We set the batch size as 128, and set the total epoch as 40. All models are trained by the optimizer Adam with from scratch. The decay of momentum of gradient is set as 0.5 and 0.9. The initial learning rate is set as 0.001 in the first 20 epochs, then the learning rate of each epoch becomes 0.9 times than that of the previous epoch.

We divide the training set and the testing set in a ratio of eight to two. After we train the proposed DB2RNet on the training set, the recognition accuracy of 1404 classes of incomplete Dongba characters will be tested on the testing set, Meanwhile, the PSNR and SSIM are compared to evaluate the restoration module.

Results and discussion

Ablation studies

In this section, the ablation experiments are conducted on our four different levels incomplete character datasets to validate the effectiveness of all components of DB2R-Net. The performance of four baseline networks are compared by removing one of the components.

(1) Direct recognition. This network is a simple recognition module without restoration module. By this mean, the imperfect Dongba character is directly fed into the recognition module for prediction. For the reason that direct recognition means that we only perform recognition task with our recognition module, and no restoration module is used. Therefore, no PSNR and SSIM values are shown in the corresponding cells. The empty cells represent that this network will not generate corresponding values.

(2) No inter-module residual connection. This network only feeds the restoration character of the restoration module into the recognition module. In this way, the original incomplete character is not concatenated with the restoration character. This network also means that
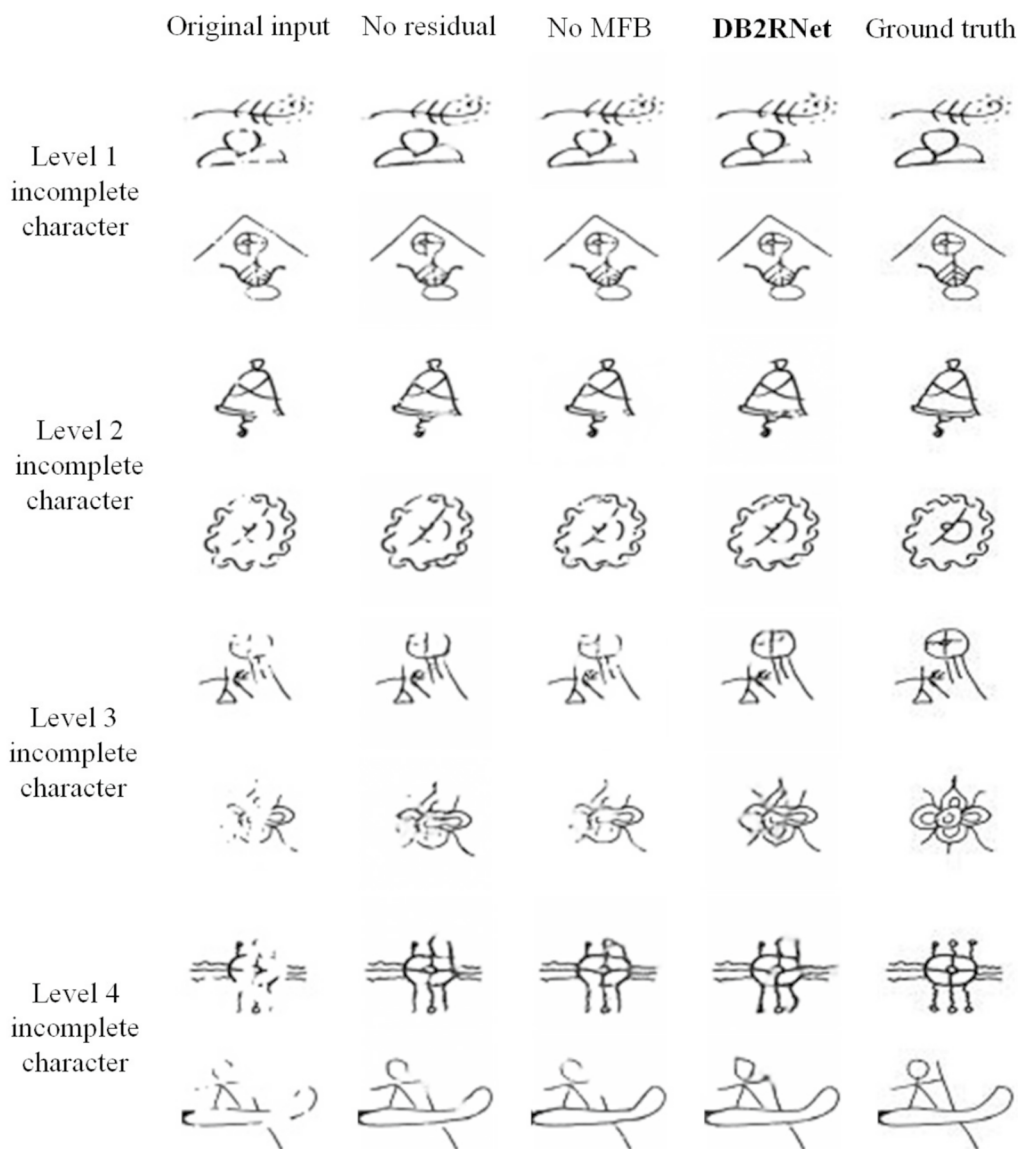
Fig. 6 Results of ablation studies presenting the contributions of individual components of our method. Original input, No residual, No MFB and DB2RNet correspond to Original incomplete character, No inter-module residual connection, multiscale feature block and DB2RNet (Ours method) in Table 3, respectively

the "Inter-module residual connection" in Fig. 2 is not used.

(3) No multiscale feature block. This network only feeds the decoder with the previous convolutional layer output. In this way, no multiscale features can be obtained, and the information transmission from the encoder to the decoder will be lost.

(4) The whole DB2RNet (ours method). This network is our proposed whole DB2RNet.

Results of ablation studies for DB2RNet and other baseline networks are given in Table 3. Accuracy, PSNR and SSIM are used as measurements. These results effectively illustrate the excellence performance of our proposed DB2RNet. It should be noted that for original incomplete character, we directly calculate the PSNR and SSIM values between the original incomplete characters and their corresponding complete characters, without feeding them into the model. Therefore, no accuracy metric appeared in the corresponding cells. From Table 3, several observations could be drawn:

Table 4 Evaluation results of excellent image recognition models and Dongba character recognition model on four different incomplete Dongba character testing sets.

Mask level datasets	Models (source)	Accuracy	
		Direct recognition	**Employed in DB2RNet**
Level 1 incomplete Dongba character	ResNet50 (CVPR2016) [28]	96.79%	97.99%
	MobileNetV1 (ArXiv2017) [32]	94.63%	95.59%
	EfficientNet (ICML2019) [33]	96.46%	97.61%
	RepVGG (CVPR2021) [34]	93.74%	94.84%
	ResNeSt50 (CVPR2022) [35]	96.06%	97.23%
	MAAN (ESWA2023) [21]	97.35%	98.35%
	Recognition module (Ours)	**97.14**%	98.25%
Level 2 incomplete Dongba character	ResNet50 (CVPR2016) [28]	92.03%	93.92%
	MobileNetV1 (ArXiv2017) [32]	88.94%	90.77%
	EfficientNet (ICML2019) [33]	91.79%	93.67%
	RepVGG (CVPR2021) [34]	88.48%	90.27%
	ResNeSt50 (CVPR2022) [35]	91.69%	93.55%
	MAAN (ESWA2023) [21]	93.32%	95.12%
	Recognition module (Ours)	92.85%	94.69%
Level 3 incomplete Dongba character	ResNet50 (CVPR2016) [28]	87.24%	89.19%
	MobileNetV1 (ArXiv2017) [32]	82.69%	84.67%
	EfficientNet (ICML2019) [33]	86.15%	88.05%
	RepVGG (CVPR2021) [34]	80.23%	82.19%
	ResNeSt50 (CVPR2022) [35]	85.74%	87.67%
	MAAN (ESWA2023) [21]	88.97%	90.86%
	Recognition module (Ours)	**88.84**%	**90.75**%
Level 4 incomplete Dongba character	ResNet50 (CVPR2016) [28]	78.34%	80.56%
	MobileNetV1 (ArXiv2017) [32]	71.88%	74.37%
	EfficientNet (ICML2019) [33]	77.64%	80.12%
	RepVGG (CVPR2021) [34]	69.85%	72.49%
	ResNeSt50 (CVPR2022) [35]	77.01%	79.59%
	MAAN (ESWA2023) [21]	82.01%	84.95%
	Recognition module (Ours)	**81.61**%	**84.29**%

Our methods are shown in bold

(1) We find that direct recognition gets a lower accuracy because of the low-quality and texture and contextual feature loss of the input incomplete Dongba character.

(2) According to the results no inter-module residual connection, we can know that the network gets the lowest accuracy, even lower than direct recognition. It can be concluded that the concatenation of restored character and original character in the recognition module is significant for recognition. It demonstrates that the inter-module residual connection makes great sense in end-to-end restoration-to-recognition network.

(3) For the network without multiscale feature block, the value of PSNR and SSIM get the lowest value. The value of PSNR even lower than original incomplete character in Level 1 and Level 2 datasets. It may be concluded that multiscale feature block in restoration module strategy has significant impact on the restored image quality.

(4) Compare all mask level datasets, we find that our proposed DB2RNet can achieve the best values and all evaluation indicators are higher than without one of the strategies. It can be concluded that the DB2RNet can handle different mask level Dongba character datasets with good robustness.

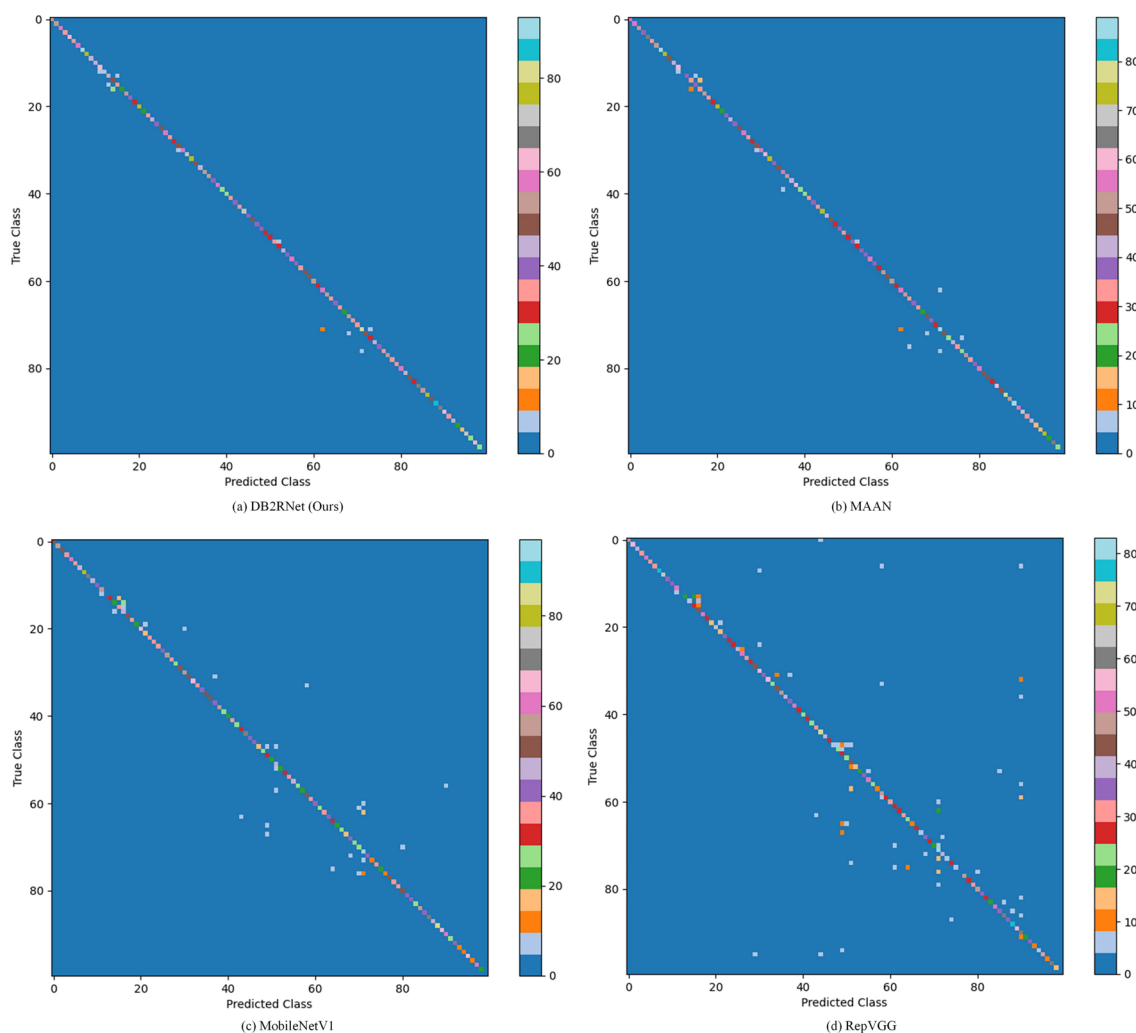(5) Combine all of the above strategies, the DB2RNet achieves the highest recognition accuracy, restora-

Fig. 7 Confusion matrices of four different models on randomly selected one hundred level 4 incomplete Dongba character classes

tion PSNR and SSIM, proving combination of above strategies obtain further improvements. The combination of all strategies demonstrates their complementary ability exceeds employing either alone. The results also prove the effectiveness of all strategies in the DB2RNet.

The results in Table 3 demonstrates that all strategies contribute to the final excellent performance of DB2RNet. For the qualitative visual analysis of the effectiveness of our proposed DB2RNet, we show the incomplete Dongba character restoration results in Fig. 6. In particular, multiscale feature block has much more effect on the restoration rather than the inter-module residual connection. Therefore, the restored Dongba character of no multiscale feature block in restoration has more broken lines and intermediate completed sketches. And the restored Dongba character of

no inter-module residual connection is closely to the best DB2RNet.

Comparison with state-of-the-art methods

Recently, lots of state-of-the-art recognition neural networks have been designed. In order to illustrate the superiority of our method, a good deal of experiments are conducted to make comparison between these excellent recognition methods and our proposed DB2RNet. What's more, we also compare our methods with different excellent end-to-end image restoration-to recognition models to validate the superiority of our proposed DB2RNet.

Comparison with recognition models

In order to reveal the superiority of the proposed DB2RNet in this paper, we conduct sufficient comparative experiments with state-of-the-art deep neural

Table 5 Detailed structure and hyperparameters of the baseline network of MAAN [21]

Layer	In s	Out s	K, S, P, In c, Out c
Conv1	64	64	$\begin{bmatrix} 3,1,1,\mathbf{2},64 \\ 3,1,1,64,64 \end{bmatrix} \times 2$
Pooling1	64	32	2×2 max pool, stride 2
Conv2	32	32	$\begin{bmatrix} 3,1,1,128,128 \\ 3,1,1,128,128 \end{bmatrix} \times 2$
Pooling2	32	16	2×2 max pool, stride 2
Conv3	16	16	$\begin{bmatrix} 3,1,1,256,256 \\ 3,1,1,256,256 \end{bmatrix} \times 2$
Pooling3	16	8	2×2 max pool, stride 2
Conv4	8	8	$\begin{bmatrix} 3,1,1,512,512 \\ 3,1,1,512,512 \end{bmatrix} \times 2$
Pooling4	8	4	2×2 max pool, stride 2
Conv5	4	4	$\begin{bmatrix} 3,1,1,1024,1024 \\ 3,1,1,1024,1024 \end{bmatrix} \times 2$
Pooling5	4	1	Global average pool
Fully-connected layer, Softmax			

Meanings of the abbreviations are the same as Table 1

network methods in the field of image recognition, such as ResNet50 [28], MobileNetV1 [32], EfficientNet [33], RepVGG [34] and ResNeSt50 [35]. At the same time, we also compare our DB2RNet with MAAN [21], which is the state-of-the-art model of perfect/complete Dongba character recognition. In this section, Accuracy is used as the measurement. Sufficient experimental results are illustrated in Table 4. Direct recognition accuracy means that the model is directly used to recognize the incomplete Dongba character. Employed in DB2RNet means that our recognition module is replaced with the model and therefore the model is employed in DB2RNet. It should be noted that our recognition module employed in DB2RNet is DB2RNet itself.

As is shown in Table 4, the DB2RNet (our recognition module employed in DB2RNet) proposed in our paper achieves competitive recognition accuracy, achieving a accuracy of 98.25% in Level 1 incomplete Dongba character. For example, compared with the best representative image recognition method ResNet50 [28] (direct recognition), DB2RNet obtains the improvement of the performance by 1.46% of accuracy. At the same time, our DB2RNet also exceeds the best perfect/complete Dongba character recognition model MAAN [21] (direct recognition) with a recognition rate of 0.9%. The above results fully verify the superiority of the propose DB2RNet. These experiments results also demonstrate that our end-to-end restoration-to-recognition has better recognition accuracy than previous image recognition methods.

As is illustrated in Table 4, the accuracy of all models decreases with the increasing irregular mask ratios of incomplete Dongba character. This is because these characters have different types of incompletion situations and larger irregular incompletion usually represents less information, which are difficult to be modeled. Nevertheless, our proposed DB2RNet remains the state-of-the-art recognition accuracy. What's more, the gap between the recognition accuracy of our DB2RNet (recognition module employed in DB2RNet) and that of other models is also increasing. Take ResNeSt50 [35] (direct recognition) as an example, the recognition accuracy gap increases from 2.19% (Level 1 incomplete Dongba character dataset) to 7.28% (Level 4 incomplete Dongba character dataset). These results demonstrate that the DB2RNet has good robustness and applicability.

As can be seen from Table 4, the recognition module in our DB2RNet (direct recognition) get a lower recognition accuracy than MAAN [21] (direct recognition), a minimum number of 0.13% in Level 3 incomplete Dongba character dataset and a maximum number of 0.47% in Level 2 incomplete Dongba character dataset. Table 2 and Table 5 show detailed structure and hyperparameters of our recognition module and the baseline network of MAAN [21]. By comparing Tables 2 and 5, we can find that the number of channels of Conv5 in Table 5 is twice that of Conv5 in Table 2, which means that Table 5 will consume more memory. The recognition module of the DB2RNet saves memory compared to MAAN and only requires a minimal recognition accuracy cost. Therefore, we adopt our recognition module rather than MAAN. It should be noted that any recognition models can replace our recognition module and can be employed in our DB2RNet. If memory is not a concern, our recognition module can be replaced with any CNN network that offers better recognition accuracy. It can be seen from Table 4 that better direct recognition accuracy models lead to better recognition accuracy when it is employed in our DB2RNet. In addition, in order to intuitively analyze the recognition results on the testing set, we also make confusion matrices for our DB2RNet, MAAN, MobileNetV1 and RepVGG, as shown in Fig. 7. To better show the differences of confusion matrices, we randomly selected one hundred level 4 incomplete Dongba character classes. It is obvious that our DB2RNet have discriminative effects for incomplete Dongba character recognition.

All these results demonstrate our proposed DB2RNet not only can achieve the state-of-the-art recognition accuracy but also can handle different incomplete level Dongba character datasets with good robustness. All the experimental results also demonstrate that incomplete handwritten Dongba characters is a challenging task that

Table 6 Evaluation results of our proposed DB2RNet with other state-of-the-art end-to-end restoration-to-recognition methods on four different ratios incomplete Dongba character testing sets

Mask level dataset	Baseline networks	Accuracy	PSNR	SSIM
Level 1 incomplete Dongba character	Original input	–	33.0841	0.9805
	Direct recognition	97.14%	–	–
	Joint-Caps-Net [27]	96.13%	33.1746	0.9834
	ADFRNet [25]	96.34%	33.3523	0.9846
	SketchGAN [26]	96.87%	33.7364	0.9853
	CPN [36]	96.56%	33.7154	0.9849
	DB2RNet(Ours)	98.25%	**33.9187**	**0.9872**
Level 2 incomplete Dongba character	Original input	–	24.2027	0.9465
	Direct recognition	92.85%	–	–
	Joint-Caps-Net [27]	88.97%	25.5428	0.9652
	ADFRNet [25]	90.46%	25.8435	0.9667
	SketchGAN [26]	92.28%	26.3528	0.9673
	CPN [36]	91.77%	26.1759	0.9670
	DB2RNet(Ours)	**94.69**%	**26.7415**	**0.9688**
Level 3 incomplete Dongba character	Original input	–	19.7016	0.8991
	Direct recognition	88.84%	–	–
	Joint-Caps-Net [27]	88.47%	22.2719	0.9391
	ADFRNet [25]	88.59%	22.3148	0.9398
	SketchGAN [26]	88.81%	22.3550	0.9401
	CPN [36]	88.63%	22.3246	0.9400
	DB2RNet(Ours)	**90.75**%	**22.5452**	**0.9420**
Level 4 incomplete Dongba character	Original input	–	18.4215	0.8688
	Direct recognition	81.61%	–	–
	Joint-Caps-Net [27]	80.25%	20.4926	0.9117
	ADFRNet [25]	80.78%	20.5013	0.9119
	SketchGAN [26]	81.53%	20.5107	0.9122
	CPN [36]	81.15%	20.5087	0.9122
	DB2RNet(Ours)	**84.29**%	**20.6328**	**0.9139**

The best results are shown in bold

previous image recognition methods cannot handle it well and our DB2RNet is an effective method.

Comparison with end-to-end restoration-to-recognition models

In order to certify the advantage of the proposed end-to-end restoration-to-recognition DB2RNet, we compare our DB2RNet with the state-of-the-art restoration-to-recognition network of vehicle logo image Joint-Caps-Net [27], sketch image ADFRNet [25] and SketchGAN [26] and ancient Chinese character [36]. We can fully and vigorously prove the advantages of the proposed DB2RNet. Table 6 shows the quantitative evaluation of our method.

It can be observed from Table 6 that the proposed DB2RNet outperforms any other state-of-the-art end-to-end image restoration-to-recognition models. Although the SketchGAN obtains better results than the other

models, our proposed DB2RNet outperforms it by a accuracy of 1.38% (Level 1 incomplete Dongba character) to 2.76% (Level 4 incomplete Dongba character). It can also be observed that our proposed DB2RNet is superior to the other models in Dongba character restoration in terms of restoration performance (higher PSNR and higher SSIM). Meanwhile, we can also find that these three modules even get lower accuracy than our direct recognition module, the reason may be that our recognition module (inspired by the state-of-the-art perfect Dongba character recognition method MAAN) has better performance and may be that they have overlooked the significance of the inter-module residual connection and the original incomplete Dongba character input not being transmitted to the recognition module in there methods.

Furthermore, we also show the restored images of the models mentioned above in Fig. 8. Compared with
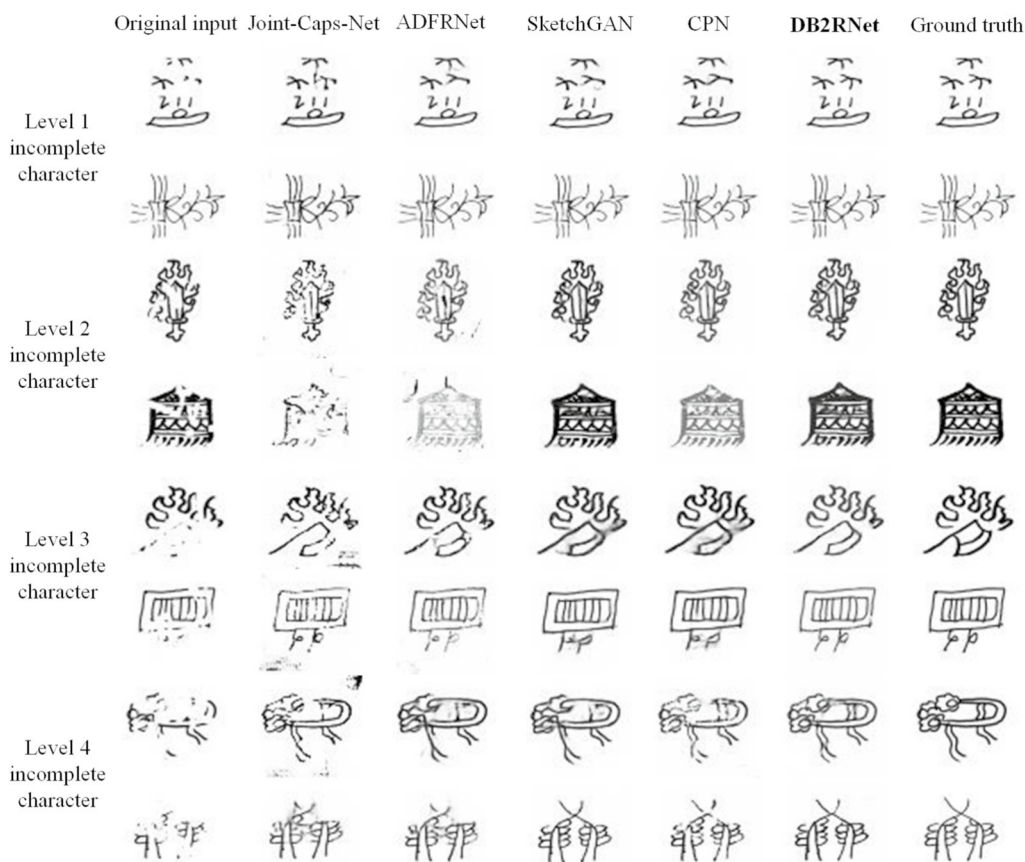
Fig. 8 Comparison with state-of-the-art end-to-end restoration-to-recognition models on four different ratios incomplete Dongba character datasets

the above four methods, our DB2RNet produces better results in terms of incomplete Dongba character restoration, which confirms the superiority of the proposed DB2RNet. In conclusion, by comparing with recently state-of-the-art end-to-end restoration-to-recognition models quantitatively and qualitatively, it can be concluded that our DB2RNet outperforms the current state-of-the-art end-to-end restoration-to-recognition models.

Performance on open sets

In order to further demonstrate the proposed DB2R-Net generalization and robustness, we conduct extensive experiments on Chinese character dataset CASIA-HWDB1.1 [4] and Oracle character dataset OBC306 [37]. On one hand, we make comparisons with state-of-the-art image recognition models which used in section Comparison with recognition models to reveal the superiority of the proposed DB2RNet, and the experimental results are shown in Table 7. On the other hand, to certify the advantage of the proposed end-to-end restoration-to-recognition DB2RNet, we compare our

DB2RNet with the state-of-the-art restoration-to-recognition networks which used in section Comparison with end-to-end restoration-to-recognition models and the experimental results are shown in Table 8.

It can be seen from Table 7 that better direct recognition accuracy models lead to better recognition accuracy when it is employed in our DB2RNet, no matter this strategy is used in CASIA-HWDB1.1 or OBC306 open sets. Both the direct recognition accuracy of our recognition module and when it is employed in DB2RNet obtain competitive results. Although they get a lower recognition accuracy than MAAN, our recognition module saves memory compared to MAAN and only requires a minimal recognition accuracy cost as we concluded in section Comparison with recognition models.

According to the results of Table 8, it may be concluded that the proposed DB2RNet has achieved higher accuracy, PSNR and SSIM values than other state-of-the-art end-to-end restoration-to-recognition methods. Therefore, these results can demonstrate that our proposed DB2RNet can outperform state-of-the-art methods effect on other open sets like CASIA-HWDB1.1 and OBC306.

Table 7 Evaluation results of excellent image recognition models on four different incomplete CASIA-HWDB1.1 [4] and OBC306 [37] open sets

Mask	Models (source)	Accuracy of CASIA-HWDB1.1 [4] & (OBC306 [37])			
		Direct recognition		Employed in DB2RNet	
Level 1	ResNet50 (CVPR2016) [28]	96.25%	[86.35%]	97.84%	[87.24%]
	MobileNetV1 (ArXiv2017) [32]	91.24%	[83.56%]	93.17%	[84.62%]
	EfficientNet (ICML2019) [33]	95.73%	[86.07%]	97.05%	[87.11%]
	RepVGG (CVPR2021) [34]	90.38%	[82.95%]	91.85%	[84.00%]
	ResNeSt50 (CVPR2022) [35]	95.26%	[85.29%]	96.54%	[86.42%]
	MAAN (ESWA2023) [21]	97.57%	[87.94%]	98.64%	[88.55%]
	Recognition module (Ours)	**97.18%**	**[87.06%]**	**98.21%**	**[88.14%]**
Level 2	ResNet50 (CVPR2016) [28]	94.23%	[83.49%]	95.47%	[84.75%]
	MobileNetV1 (ArXiv2017) [32]	90.07%	[81.20%]	91.52%	[83.08%]
	EfficientNet (ICML2019) [33]	93.61%	[83.09%]	95.04%	[83.96%]
	RepVGG (CVPR2021) [34]	88.38%	[80.83%]	89.67%	[82.79%]
	ResNeSt50 (CVPR2022) [35]	93.18%	[82.46%]	94.55%	[83.68%]
	MAAN (ESWA2023) [21]	95.97%	[85.93%]	97.35%	[86.57%]
	Recognition module (Ours)	**95.29%**	**[84.13%]**	**96.31%**	**[85.71%]**
Level 3	ResNet50 (CVPR2016) [28]	91.58%	[81.23%]	93.30%	[82.25%]
	MobileNetV1 (ArXiv2017) [32]	86.80%	[76.25%]	88.46%	[77.99%]
	EfficientNet (ICML2019) [33]	91.05%	[80.12%]	92.97%	[82.09%]
	RepVGG (CVPR2021) [34]	85.27%	[75.95%]	87.67%	[77.67%]
	ResNeSt50 (CVPR2022) [35]	90.83%	[79.10%]	92.14%	[81.95%]
	MAAN (ESWA2023) [21]	94.05%	[82.67%]	95.93%	[84.16%]
	Recognition module (Ours)	**92.94%**	**[81.94%]**	**94.61%**	**[83.58%]**
Level 4	ResNet50 (CVPR2016) [28]	88.15%	[73.58%]	90.32%	[75.95%]
	MobileNetV1 (ArXiv2017) [32]	84.56%	[67.75%]	86.34%	[70.14%]
	EfficientNet (ICML2019) [33]	87.69%	[72.95%]	89.29%	[75.81%]
	RepVGG (CVPR2021) [34]	82.94%	[65.85%]	85.70%	[68.32%]
	ResNeSt50 (CVPR2022) [35]	87.03%	[72.71%]	88.95%	[75.50%]
	MAAN (ESWA2023) [21]	90.23%	[76.39%]	92.54%	[77.92%]
	Recognition module (Ours)	**89.42%**	**[75.67%]**	**91.86%**	**[77.82%]**

Results of OBC306 are shown in "[ ]". Our methods are shown in bold

The experimental results of our DB2RNet are significantly higher than those of other state-of-the-art methods, and the gap between them is even wider than what is shown in Table 6.

Furthermore, we also show the restored Chinese character and Oracle character of the models in Fig. 9. It can be seen that our DB2RNet produces better results in terms of both incomplete Chinese character and Oracle character restoration, which confirms the superiority of the proposed DB2RNet. In conclusion, by comparing with recently these models quantitatively and qualitatively on open sets, the credibility, robustness and generality of our proposed DB2RNet are proved.

Conclusion

In this paper, a novel end-to-end network for incomplete handwritten Dongba character recognition by multiscale feature restoration (DB2RNet) is proposed, and the restored characters can be used to assist boosting the performance of Dongba character recognition. We first generate four incomplete Dongba character

Table 8 Evaluation results of our proposed DB2RNet with other state-of-the-art end-to-end restoration-to-recognition methods on four different ratios CASIA-HWDB1.1 [4] and OBC306 [37] open sets

Mask	Baseline networks	CASIA-HWDB1.1 [4]			OBC306 [37]		
		Accuracy	PSNR	SSIM	Accuracy	PSNR	SSIM
Level 1	Original input	–	28.4357	0.9677	–	18.7426	0.8703
	Direct recognition	97.18%	–	–	87.06%	–	–
	Joint-Caps-Net [27]	92.71%	29.8596	0.9687	84.78%	26.4257	0.9281
	ADFRNet [25]	94.76%	30.2564	0.9808	85.46%	28.7657	0.9493
	SketchGAN [26]	96.76%	33.7815	0.9896	87.35%	31.3429	0.9668
	CPN [36]	95.37%	32.6352	0.9870	87.03%	30.6652	0.9611
	DB2RNet(Ours)	**98.21%**	**35.4827**	**0.9930**	**88.14%**	**31.7295**	**0.9728**
Level 2	Original input	–			–	13.6524	0.7401
	Direct recognition	95.29%	–	–	84.13%	–	–
	Joint-Caps-Net [27]	90.89%	24.5623	0.9402	82.98%	24.3605	0.8896
	ADFRNet [25]	92.99%	26.8513	0.9598	83.69%	25.4682	0.9038
	SketchGAN [26]	94.88%	28.3892	0.9729	84.49%	26.8657	0.9296
	CPN [36]	93.46%	27.2753	0.9658	84.19%	26.4816	0.9220
	DB2RNet(Ours)	**96.31%**	**30.3125**	**0.9818**	**85.71%**	**27.3292**	**0.9371**
Level 3	Original input	–	18.9652	0.8515	–	11.1254	0.6258
	Direct recognition	92.94%	–	–	81.94%	–	–
	Joint-Caps-Net [27]	89.74%	22.4410	0.9091	78.18%	21.9568	0.8351
	ADFRNet [25]	91.62%	24.0192	0.9277	80.57%	22.4690	0.8427
	SketchGAN [26]	93.29%	25.3416	0.9490	81.91%	23.7605	0.8819
	CPN [36]	92.31%	24.1608	0.9350	81.27%	23.5428	0.8737
	DB2RNet(Ours)	**94.61%**	**27.2610**	**0.9654**	**83.58%**	**24.0188**	**0.8941**
Level 4	Original input	–	17.3593	0.7901	–	9.5343	0.5370
	Direct recognition	89.42%	–	–	75.67%	–	–
	Joint-Caps-Net [27]	85.69%	20.5183	0.8659	71.16%	20.6297	0.7973
	ADFRNet [25]	87.26%	21.7918	0.8851	72.96%	20.8642	0.8027
	SketchGAN [26]	90.94%	23.2168	0.9273	75.65%	21.3225	0.8252
	CPN [36]	89.45%	22.8608	0.8949	74.53%	21.1634	0.8176
	DB2RNet(Ours)	**91.86%**	**24.7923**	**0.9422**	**77.82%**	**21.9686**	**0.8410**

The best results are shown in bold

datasets with four different incompleteness levels. Then, the incomplete Dongba character is first restored by the proposed restoration module, and then sent to the recognition module. The incomplete Dongba character is concatenated in restoration and recognition module in a inter-module manner, which can assist improve the performance of the proposed DB2RNet. In addition, novel multiscale feature blocks are introduced, which can provide more effective texture and contextual information transmission for Dongba character image restoration, and thus yielding better restoration effects and better recognition results. Extensive experiments results confirm the effectiveness of the components of the proposed DB2RNet, and demonstrate that it achieves the state-of-the-art performance on incomplete Dongba character recognition task. By comparing with recently state-of-the-art models quantitatively and qualitatively on CASIA-HWDB1.1 and OBC306 open sets, the credibility, robustness and generality of our proposed DB2RNet are proved.

Fig. 9 Comparison with state-of-the-art end-to-end restoration-to-recognition models on four different ratios incomplete Chinese character [4] and Oracle character OBC306 [37] datasets

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no conflict of interest.

References
1. Wang J, Wang S, Chen S. Research on the digital protection strategies of intangible cultural heritage. Softw Guide. 2011;10(8):49–51.
2. Wu G, Ding C, Xu X, Wang N. Intelligent recognition on Dongba manuscripts hieroglyphs. J Electron Meas Instrum. 2016;30(11):1774–9.
3. Guo H, Jy Zhao, Xn Li. Preprocessing method for NaXi pictographs character recognition using wavelet transform. Int J Digit Content Technol Appl. 2010;4(3):117–31.
4. Liu CL, Yin F, Wang DH, Wang QF. Online and offline handwritten Chinese character recognition: benchmarking on new databases. Pattern Recognit. 2013;46(1):155–62.
5. Xu X, Jiang Z, Wu G, Wang H, Wang N. Identification method of Dongba pictograph based on topological characteristic and projection method. J Electron Meas Instrum. 2017;31(1):150–4.
6. Da M, Zhao Jy, Suo G, Guo H. Online handwritten Naxi pictograph digits recognition system using coarse grid. In: International Workshop on Computer Science for Environmental Engineering and EcoInformatics. Springer; 2011. p. 390–396.

7. Tong S, Jianjun Z, Wensi L, Yunmu W, Yifei X, Zhijian Z, et al. Research on recognition of Dongba script by a combination of HOG feature extraction and support vector machine. J Nanjing Univ. 2020;56(6):870–6.
8. Zhang J, Lai Z, Kong H, Shen L. Robust Twin Bounded Support Vector Classifier With Manifold Regularization. IEEE Trans Cybern. 2023;53(8):5135–50.
9. Meng T, Huang R, Lu Y, Liu H, Ren J, Zhao G, et al. Highly sensitive terahertz non-destructive testing technology for stone relics deterioration prediction using SVM-based machine learning models. Herit Sci. 2021;9:1–9.
10. El-Askary NS, Salem MAM, Roushdy MI. Features processing for Random Forest optimization in lung nodule localization. Expert Syst Appl. 2022;193: 116489.
11. Sun Z, Wang G, Li P, Wang H, Zhang M, Liang X. An improved random forest based on the classification accuracy and correlation measurement of decision trees. Expert Syst Appl. 2024;237: 121549.
12. Vincent SSM, Duraipandian N. Detection and prevention of sinkhole attacks in MANETS based routing protocol using hybrid AdaBoost-Random forest algorithm. Expert Syst Appl. 2024;249: 123765.
13. Wang Y, Pang W, Jiao Z. An adaptive mutual K-nearest neighbors clustering algorithm based on maximizing mutual information. Pattern Recognit. 2023;137: 109273.
14. Ali A, Hamraz M, Gul N, Khan DM, Aldahmani S, Khan Z. A k nearest neighbour ensemble via extended neighbourhood rule and feature subsets. Pattern Recognit. 2023;142: 109641.
15. Huang Y, Chen D, Wang H, Wang L. Gender recognition of Guanyin in China based on VGGNet. Herit Sci. 2022;10(1):93.
16. Jin X, Wang X, Xue C. Nondestructive characterization and artificial intelligence recognition of acoustic identifiers of ancient ceramics. Herit Sci. 2023;11(1):144.
17. Luo H, Xu D, Yang B, Zhang H. Multi-scale Feature Fusion Based Dongba Character Recognition. In: 2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE). IEEE; 2020. p. 1571–1575.
18. Hua R, Xu X. Intelligent classification on images of Dongba ancient books. J Eng. 2019;2019(23):9039–42.
19. Luo Y, Bi X, Wu L, Li X. Dongba pictographs recognition based on improved residual learning. CAAI Trans Intell Syst. 2022;17(1):79–87.
20. Xie Y, Dong J. Research on Dongba hieroglyph recognition using ResNet network. Comput Era. 2021;343(1):6–10.
21. Luo Y, Sun Y, Bi X. Multiple attentional aggregation network for handwritten Dongba character recognition. Expert Syst Appl. 2023;213: 118865.
22. Li L. Naxi pictographs and transcription characters dictionary. 1st ed. Kunming: Yunnan Nationalities Publishing House; 2001.
23. Fang G, He Z. Naxi pictograph character chart. 1st ed. Kunming: Yunnan People's Publishing House; 1981.
24. Liu G, Reda FA, Shih KJ, Wang TC, Tao A, Catanzaro B. Image inpainting for irregular holes using partial convolutions. In: Proceedings of the European conference on computer vision (ECCV); 2018. p. 85–100.
25. Wan J, Zhang K, Li H, Chan AB. Angular-driven feedback restoration networks for imperfect sketch recognition. IEEE Trans Image Process. 2021;30:5085–95.
26. Liu F, Deng X, Lai YK, Liu YJ, Ma C, Wang H. Sketchgan: Joint sketch completion and recognition with generative adversarial network. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2019. p. 5830–5839.
27. Chen R, Mihaylova L, Zhu H, Bouaynaya NC. A deep learning framework for joint image restoration and recognition. Circuits Syst Signal Process. 2020;39:1561–80.
28. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. p. 770–778.
29. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017. p. 4700–4708.
30. Huang G, Liu S, Van der Maaten L, Weinberger KQ. Condensenet: An efficient densenet using learned group convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2018. p. 2752–2761.
31. Lin TY, Dollár P, Girshick R, He K, Hariharan B, Belongie S. Feature pyramid networks for object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017. p. 2117–2125.
32. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861. 2017; https://doi.org/10.48550/arXiv.1704.04861.
33. Tan M, Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In: International conference on machine learning. PMLR; 2019. p. 6105–6114.
34. Ding X, Zhang X, Ma N, Han J, Ding G, Sun J. Repvgg: Making vgg-style convnets great again. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition; 2021. p. 13733–13742.
35. Zhang H, Wu C, Zhang Z, Zhu Y, Lin H, Zhang Z, et al. Resnest: Split-attention networks. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition; 2022. p. 2736–2746.
36. Xu Y, Zhang XY, Zhang Z, Liu CL. Large-scale continual learning for ancient Chinese character recognition. Pattern Recognit. 2024;150: 110283.
37. Huang S, Wang H, Liu Y, Shi X, Jin L. OBC306: A large-scale oracle bone character recognition dataset. In: 2019 International Conference on Document Analysis and Recognition (ICDAR). IEEE; 2019. p. 681–688.

Publisher's Note