Automatic Classification of Blue and White Porcelain Sherds Based on Data Augmentation and Feature Fusion

Yanzhe Liu, Bingxiang Liu, Jiajia Yu, Jingwen Xia & Canfei Luo

Published online: 24 Oct 2021.

Submit your article to this journal ↗

Article views: 935

View related articles ↗

View Crossmark data ↗

Taylor & Francis
Taylor & Francis Group

Automatic Classification of Blue and White Porcelain Sherds Based on Data Augmentation and Feature Fusion

Yanzhe Liu[a], Bingxiang Liu[a], Jiajia Yu[b], Jingwen Xia[a], and Canfei Luo[a]

[a]School of Information and Enginnering, Jingdezhen Ceramic University, Jingdezhen, China; [b]School of Electronic Information and Artificial Intelligence, Shaanxi University of Science & Technology, Xi'an, China

ABSTRACT

Many blue and white porcelain are unearthed in Jingdezhen every year. The patterns on the sherds have important research significance. At present, the classification of porcelain shards is mainly based on manual work, which has the disadvantages of large workload. The use of automatic classification methods also faces complex patterns and sample sizes. In order to solve these problems, this paper proposes a new automatic recognition method based on deep learning, including data preprocessing method combined with color segmentation algorithm, a new data augmentation method FCutMix for regions of interest, a new integration strategy and the redesigned deep network model FFCNet that integrates multiple features. After experiments, the data preprocessing method, feature fusion method and integration strategy proposed in the paper can effectively improve the performance of the model by removing redundant information and adding effective features. The FCutMix method can also obtain more accurate mixed samples than the traditional CutMix. The method proposed in this paper improves the accuracy of tasks in 14 categories from 71.7% to 83.2% in a dataset containing only 373 images of porcelain sherds. In the future, this research will further design the network structure and multi-level feature fusion.

Introduction

Jingdedzhen, one of the main porcelains productional area of China, where many porcelains are unearthed every year. Among them, the blue and white porcelain of the Yuan, Ming and Qing Dynasty is the most famous and precious. These blue and white porcelains use CoO as the pigment and are painted on the biscuit (unfired porcelain) with the technique of Chinese painting. These patterns are not only an important reference for the study of social economy and culture, but also an important basis in ceramic restoration.

The objective of this paper is a method to classify the archeological ceramic shards by incomplete patterns in the Yuan, Ming and Qing dynasties. Since the porcelain unearthed from archeological excavations is usually in sherds, it is

CONTACT Bingxiang Liu ✉ 18860363676@163.com 🏛 Jingdezhen Ceramic University, Jingdezhen, Jiangxi, China 333403

necessary to distinguish the sherds of different porcelains in advance and then splice them. The mixed porcelain shards are usually similar in texture and shape, so they can only be distinguished by the patterns on the shards. This task is usually carried out by experienced archeological experts, indeed as more and more shards are excavated, the low efficiency of manual classification cannot meet the demand, consequently it is necessary to propose an automated classification method. However, these patterns are not only complex in structure, but their integrity, which is often damaged. At the same time, because the patterns are drawn manually by workers, who are often affected by drawing techniques and other factors, making the same type of patterns look quite different. In addition, although many archeological ceramic shards have been unearthed in recent years, they are still far from enough as training data. These problems have made this work extremely challenging.

In the early days, some studies tried to use the texture and color features of ceramic to classify ceramics fragments. Such methods usually use Gabor filter to extract features, then use KNN or SVM to complete the classification (Abadi, Khoudeir, and Marchand 2012; Debroutelle et al. 2017; Hanzaei, Afshar, and Barazandeh 2017; Smith et al. 2010). This classification method relies on manual design with low accuracy and low generalization ability.

Neural network is widely used in various fields because of its powerful performance (Li et al. 2020), and the deep learning network model based on the LeNet model provides the possibility for image processing. Mukhoti, Dutta, and Sarkar (2020) completed the handwritten digit classification in bangla and hindi based on the LeNet model. In 2014, the rise of deep convolutional neural networks represented by VGG (Simonyan and Zisserman 2014) has brought image processing into a new stage of development. The subsequent appearance of ResNet (He et al. 2016) and GoogLeNet (Szegedy, Liu, and Jia et al. 2015) further established the dominant position of deep convolutional networks in image recognition. For example, Liu et al. (2020) proposed a new automatic end-to-end method based on deep learning for detection and counting oil palm trees from images obtained from unmanned aerial vehicle (UAV) drone.

Chetouani et al. (2018) tries to use deep convolutional networks to complete the classification of ceramics. However, in Chetouani's paper, there is no targeted design for the particularity of the ceramic image, resulting in poor results. Because the patterns on the fragments are more complicated and the features are difficult to extract, multi-feature fusion is a good idea. Thepade and Chaudhari (2020) used feature fusion to threshold aerial images and proved the effect of multi-level feature fusion. Taheri and Toygar (2019) fused a variety of features to estimate the age of the face, which proved the effectiveness of the feature fusion method in deep learning.

In addition, because the convolutional neural network needs to train a wealth of parameters, it is difficult to achieve the desired effect when the amount of training data is small, so data augmentation is also an indispensable and important work. At present, there are mainly the following methods of data augmentation. Image augmentation method based on basic image transformation, single-sample image augmentation method represented by CutOut (Chen et al. 2020; DeVries and Taylor 2017; Li, Li, and Long 2020; Singh, Yu, and Sarmasi et al. 2018; Zhong et al. 2020), and multi-sample image augmentation method represented by MixUp (Gong et al. 2021; Yun et al. 2019; Zhang, Cisse, and Dauphin et al. 2017). Also, there is an image augmentation method based on image generation represented by CycleGAN (Han, Murao, and Satoh et al. 2019; Shin et al. 2018; Zhu et al. 2017). However, the actual effect of this method remains to be verified due to the high complexity of the model.

This paper proposes a complete method from data extraction to automatic classification to facilitate the interpretation of this archeological ceramic heritage. Aiming at the problem of unclear feature importance, a data preprocessing method combined with color segmentation is proposed in fragment images; Aiming at the problem of insufficient data, both single-sample image augmentation and multi-sample image augmentation methods are used. And aiming at the problem of label and feature mismatch in multi-sample image augmentation, an image blend augmentation method FCutMix(Focus CutMix) based on interest region is proposed; In the recognition model, to deal with the problem that the features extracted by the convolutional network cannot represent the spatial correlation of the grayscale image, a network model FFCNet (Feature Fusion Ceramic Network) that integrates multiple features is proposed, which enriches the diversity of features; Finally, a new integrated learning strategy is proposed to further improve the performance of the model.

Automatic Recognition Method

This part introduces in detail three key methods to realize automatic fragment classification: data pre-processing, data augmentation, and classification network. Standardize data through data pre-processing methods to increase data availability and reduce the difficulty of model training. Using data augmentation methods greatly increases the number of data sets and improves the generalization ability of the model. A deep convolutional network is designed to extract image features and classify porcelain shards accordingly.

Data Pre-processing

In the process of extracting images, convolutional networks cannot distinguish between key features and useless features autonomously. For this study, the pattern is the key feature, and others are useless features. In order to improve the efficiency of network training, this research proposes a data preprocessing method combined with color segmentation algorithms, whose purpose is to remove or reduce the negative impact of useless features on network training. The data processing flow of this method is shown in Figure 1.

The core idea of this method is to extract the pattern area by using the prominent color feature of the blue and white pattern. The specific thinking is to convert the color space (from RGB to HSV) after down-sampled the original image and removing noise, then the pixels are marked in the range defined by the blue and white color distribution law.

Finally, the image is cropped by the smallest square which wraps all pixels, and the size is reset to a uniform value to complete the standardization. The standardized picture suppresses the influence of irrelevant information on the model and avoids the problem of picture distortion caused by directly changing the picture size. In addition, the size of the image cropping area is controlled by calculating the IoU when calculating the smallest square to improve the robustness of the algorithm in the research. The IoU calculation method in this article is represented by formula (1).

$$IoU = \frac{areaofoverlap}{areaofunion} = \frac{areaofparcel}{areaoforiginal} \tag{1}$$

Data Augmentation

The image augmentation consists of two parts in this research, the augmentation of the traditional image transformation method, and the augmentation of the FCutMix method. Among them, the FCutMix method from improved CutMix is the focus of this part.

Augmentation of the Traditional Image Transformation

The image augmentation method based on image transformation is not the focus of this article, but it is an indispensable part. This method achieves the purpose of generating new image samples by performing specific image transformation operations on the training image samples, and the label information of the new samples is consistent with the label information of the original image samples. The paradigm of basic image augmentation by transformation can be defined by the following formal description, where $X = \{x_j\}_{j=1}^{N}$ is used to describe the training set containing N samples, $Y = \{y_j\}_{j=1}^{N}$ is
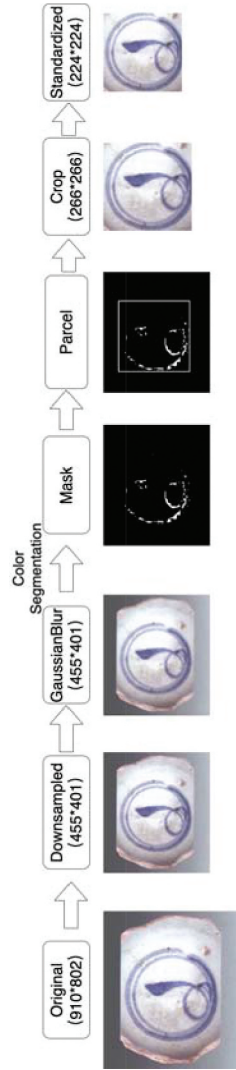
Figure 1. Data preprocessing.

the corresponding label set, $T(\cdot)$is the image transformation method, x'_j and y'_j respectively represent the sample after image transformation and the corresponding label.

$$x_{j'} = T(x_j)y_{j'} = y_j \qquad (2)$$

Sixteen traditional image transformation methods, such as bilinear interpolation, linear transformation, and inverted color channels, are grouped according to the transformation intensity and set different usage probabilities in this paper. The use probability of high transformation intensity should be set smaller, otherwise if the method with high transformation intensity is used frequently, that will reduce the image quality. Different transformation operations are randomly combined according to probability in the use process, so that the generated samples are different. This method increases the number of original samples from 373 to 3994. Figure 2 shows multiple new samples generated from one sample after data augmentation.

Augmentation of the FCutMix

The multi-sample image augmentation method represented by CutMix is currently the most used image augmentation method. In this method, patches are cut and pasted among training images where the ground truth labels are also mixed proportionally to the area of the patches. However, this method only relies on the area ratio and ignores the features in the image when recalculating the label. Even if the pattern feature is not included in the patches, the label of the sample will be introduced according to the size of the patch, which will lead to a mismatch between the sample and the label and mislead the model. This misleading sample is shown in Figure 4.

In order to solve this problem, this research improves the CutMix method and proposes the FCutMix method. Let $x \in \mathbb{R}^{w \times h \times c}$ and y denote the image and label. The goal of FCutMix is to generate a new sample$(\tilde{x}, \tilde{y})$by combining two samples (x_a, y_a) and (x_b, y_b). We define the combining operation as

$$\begin{cases} \bar{\bar{x}} \, M \cdot x_a + (1 - M) \cdot (N \cdot x_b) \\ \bar{\bar{y}} \, \lambda y_a + (1 - \lambda)y_b \end{cases} \qquad (3)$$

Where $M, N \in \{0, 1\}^{W \times H}$, M denotes a binary mask indicating where to drop out and fill in from two images, N is used to mark the area of interest. The combination ratio λ between two data points is sampled from the beta distribution Beta(1, 1), which means λ is sampled from the uniform distribution (0, 1).

This method divides the training samples into a base group and a feature group. The base group is a sample generated by augmentation of image transformation. The feature group is composed of pictures whose main content is the target feature. The samples in the base group are cut and pasted into
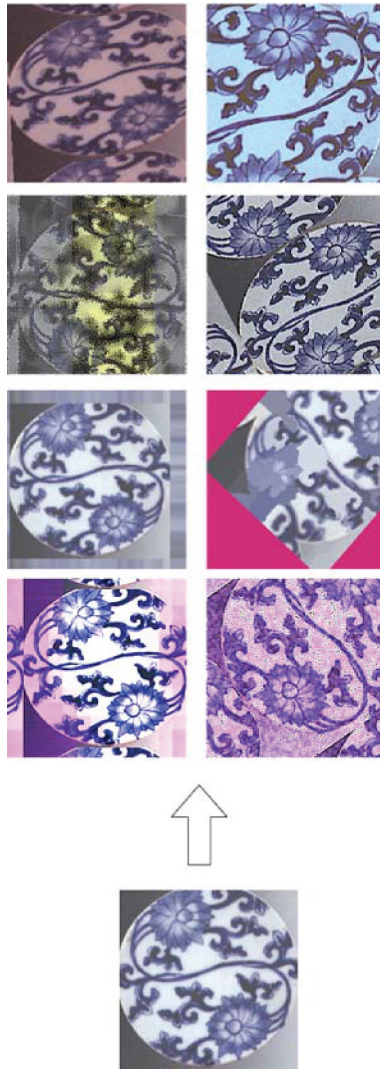
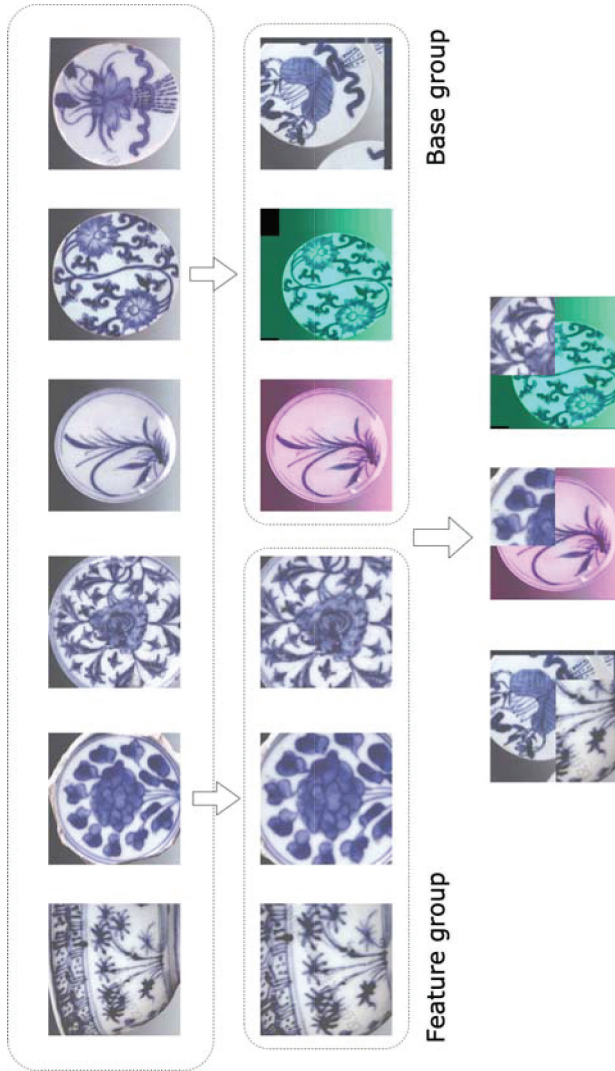**Figure 2.** Augmentation of the traditional image transformation.
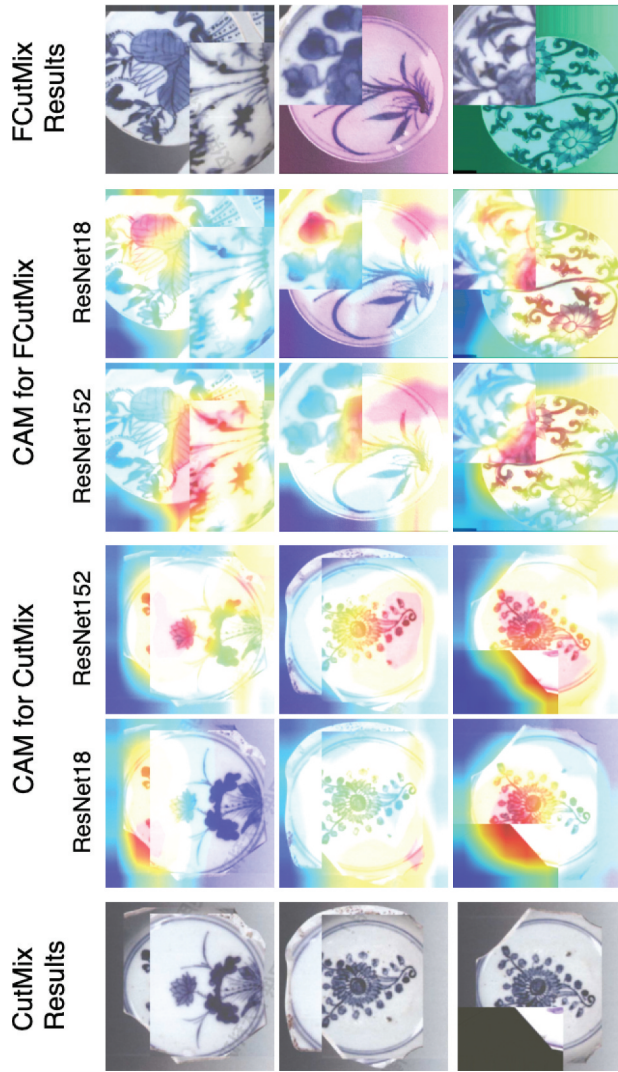
Figure 3. FCutMix result.

Figure 4. Class activation map.

the feature group. Among them, the feature group sample is obtained by cutting the pattern area, which including most of the feature pixels extracted by color segmentation. The purpose of this is to make the mixed image still contain the corresponding two features even when the cropping area is very small. The specific process is shown in Figure 3.

This paper visually compares the class activation map (CAM) (Zhou et al. 2016) of CutMix and FCutMix to clearly see the effect of augmentation method only. The result is shown in Figure 4, which shows that when the CutMix blend image is cropped to an area that does not contain a pattern, it will mislead the network's training direction, such as regarding the edge as an important feature, but FCutMix avoids this.

Classification Model

Convolutional network is the mainstream image classification model. The basic thinking is to use fully connected layer classification after extract image features through several layers of convolution operations. The features extracted by two residual networks of different depths at different stages are shown in Figure 5, which reveals when the convolutional network extracts features the resolution of the feature map decreases and the abstraction ability increases as the number of network layers increases.

But the results also show that in the task of convolutional network, the three-channel input for pictures makes the data redundant, and because the extracted feature map cannot show the spatial correlation of the gray image in the image, the texture feature is relatively rough. The gray-level co-occurrence matrix can show the gray-level spatial correlation in the image according to the gray-level relationship of neighboring pixels in space, and describe the image texture feature, which can make up for the lack of feature extraction by the convolutional network. However, when using this method to extract features, a large amount of other information in the picture will be lost, so the effect of using this method alone for classification is poor. According to the characteristics of the two methods to extract features, this paper designs a multi-feature network model FFCNet based on channel fusion, and the model structure diagram is shown in Figure 6. This channel-based fusion method can add new features based on retaining the original data, avoiding the loss of information. At the same time, the features are fused at the same precision, making the fusion more thorough. After feature fusion, it is the backbone network, which can be assumed by any deep convolutional network. After experimental verification, this feature fusion method can effectively improve the classification accuracy.
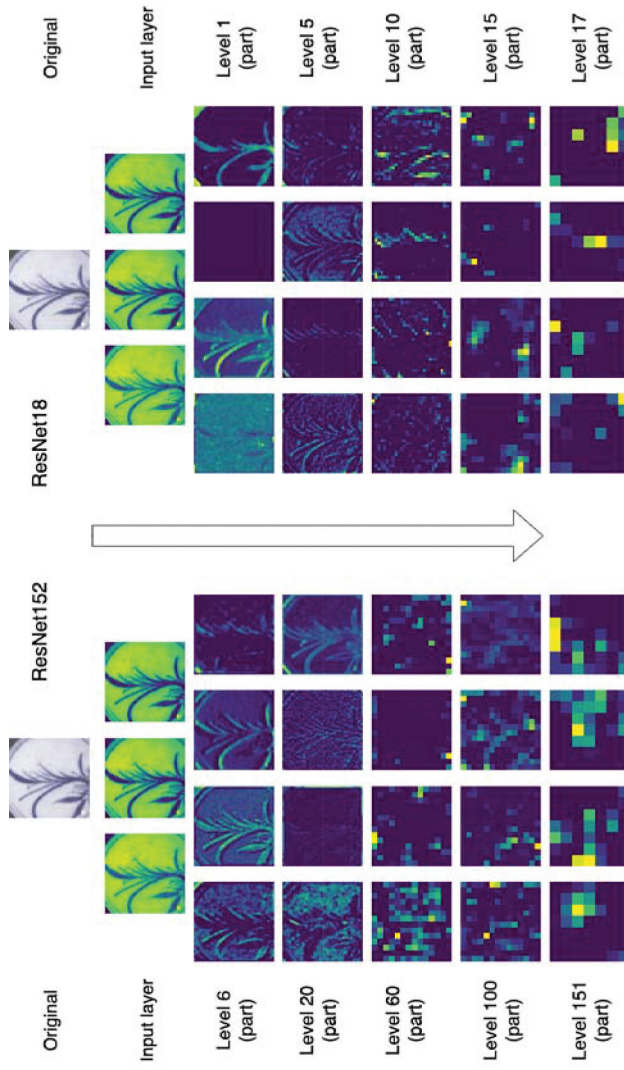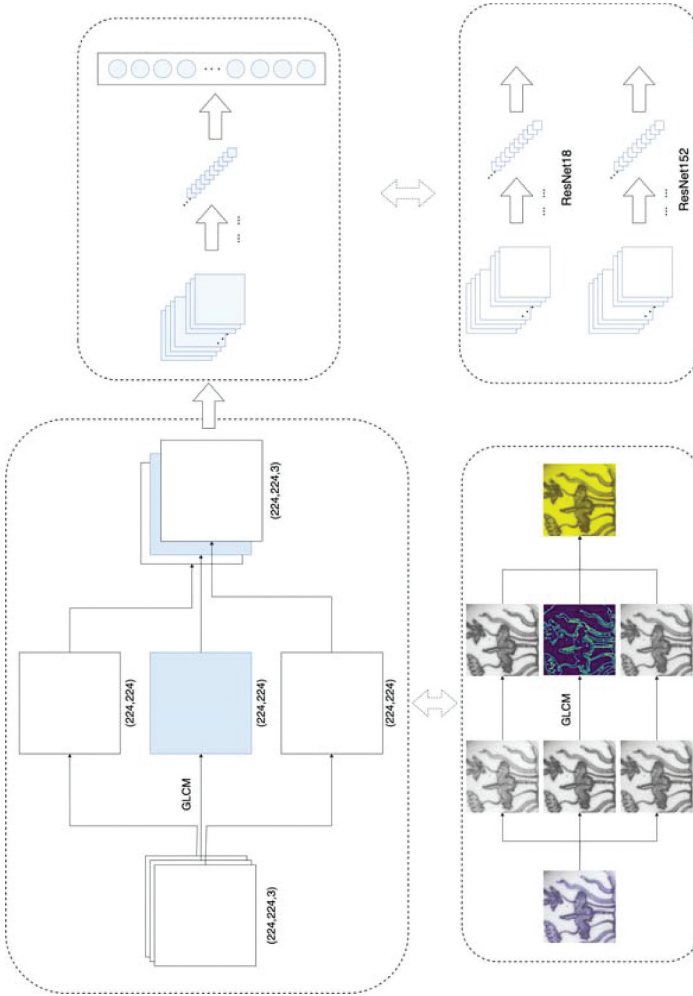
Figure 5. Feature map.

Figure 6. FFCNet structure diagram.

Integrated Learning

Since the features extracted by different methods have their own character-istics, this article uses integrated learning to fuse the classification results of different features. Traditional integrated learning usually averages one-hot encoding values or uses voting learning mechanisms, but these methods do not consider the different performance between different models. This paper proposes an integrated strategy of weighted average one-hot values and assigns corresponding weights according to the performance of the model in the training set, so that the model with higher accuracy has a higher priority when determining the result. After testing, this method can slightly improve the accuracy of the model and is better than other integration strategies. The calculation formula is as follows, where k_i is the weighted value determined according to the effect of the model in the training set.

$$\hat{y} = max\left(k_1 \begin{bmatrix} y_1^{[1]} \\ \vdots \\ y_n^{[1]} \end{bmatrix} + k_2 \begin{bmatrix} y_1^{[2]} \\ \vdots \\ y_n^{[2]} \end{bmatrix} + \cdots + k_m \begin{bmatrix} y_1^{[m]} \\ \vdots \\ y_n^{[m]} \end{bmatrix} \right) \tag{4}$$

Experimental Evaluation

The original data set used in this article is a total of 373 ceramic image samples of 14 types, and the theoretical data volume is 127,808 after data augmenta-tion. These samples are all blue and white porcelain archeological shards from the Yuan, Ming and Qing dynasties, collected from local cultural relics in Jingdezhen and a book (Jiangsu Provincial Study Society For Ancient Ceramics 2010). Figure 7 shows some of these specimens. All models are trained using the Adam optimizer with a learning rate of 10^{-5}, and the learning rate decays to $\frac{1}{10}$ of the current learning rate every 20 rounds. Weight decay is also used to avoid overfitting, and the weight decay rate is 10^{-4} (He et al. 2019). At the same time, all models use migration learning, the model parameters are initialized to the parameters obtained after a large amount of data pre-training, and the EarlyStopping strategy is also used to save the network parameters with the best accuracy.

This paper first uses ResNet152 to verify the effectiveness of the improved method proposed. The results are shown in Figure 8. The four figures respec-tively verify the effectiveness of data processing, augmentation of the tradi-tional image transformation, augmentation of the FCutMix and feature fusion through comparison. These improved methods increase the accuracy of ResNet152 from 71.7 to 79.5%.
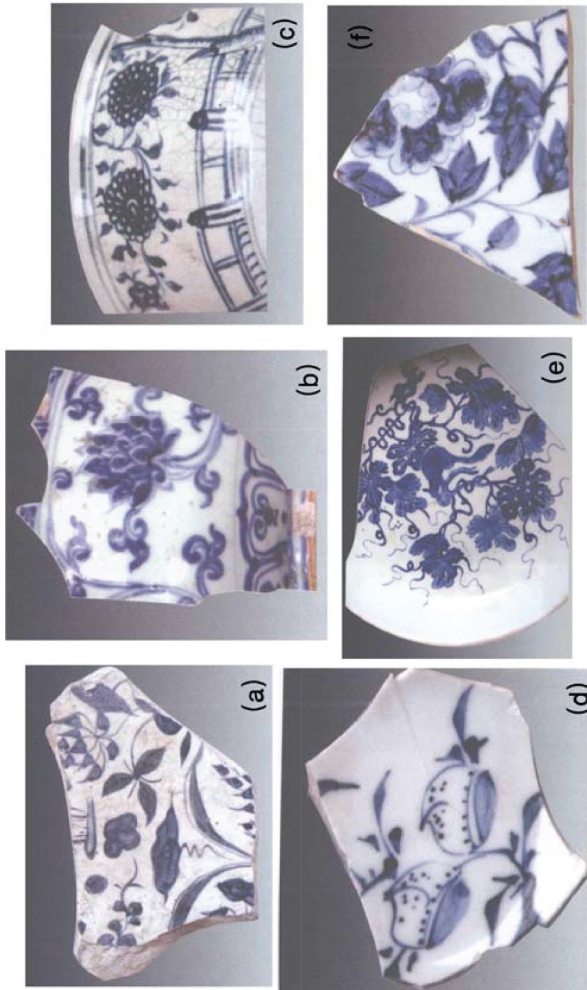
Figure 7. Examples of the archeological ceramic pieces: (a) lotus pattern, Yuan; b) decoration with interlock branch lotus, Ming; c) chrysanthemum pattern, Ming; d) peach grain, Qing; e) grape pattern, Qing; f) pattern with fold branch peony, Yuan.
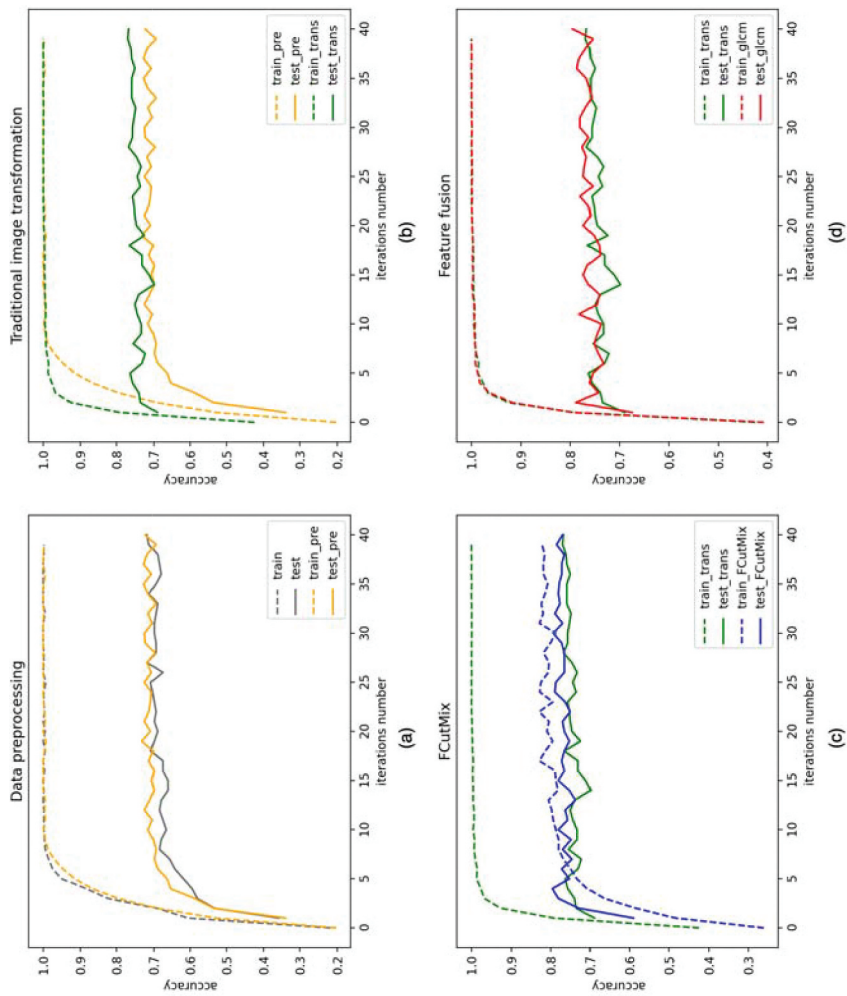

Figure 8. Accuracy comparison.

Table 1. Results of multiple tests.

|	Cut-Rate	Recall	Kappa	F$_1$ score	Accuracy	Std-acc	Mean-acc
Base	100%	0.690	0.641	0.684	0.679	0.015	0.698
		0.605	0.658	0.648	0.698		
		0.597	0.677	0.651	0.717		
	90%	0.543	0.650	0.609	0.693	0.050	0.634
	80%	0.519	0.613	0.581	0.660		
	70%	0.516	0.609	0.577	0.656		
	60%	0.485	0.565	0.542	0.613		
	50%	0.444	0.494	0.490	0.547		
FCutMix	100%	0.661	0.731	0.708	0.762	0.019	0.758
		0.707	0.702	0.720	0.734		
		0.654	0.749	0.711	0.778		
	90%	0.690	0.745	0.729	0.774	0.042	0.727
	80%	0.678	0.724	0.714	0.755		
	70%	0.628	0.691	0.668	0.731		
	60%	0.619	0.691	0.668	0.726		
	50%	0.514	0.605	0.574	0.651		
FFCNet	100%	0.678	0.757	0.728	0.785	0.007	0.795
		0.715	0.772	0.754	0.797		
		0.704	0.777	0.750	0.802		
	90%	0.688	0.748	0.729	0.776	0.017	0.745
	80%	0.637	0.715	0.687	0.746		
	70%	0.600	0.709	0.664	0.741		
	60%	0.640	0.702	0.683	0.733		
	50%	0.650	0.696	0.687	0.728		
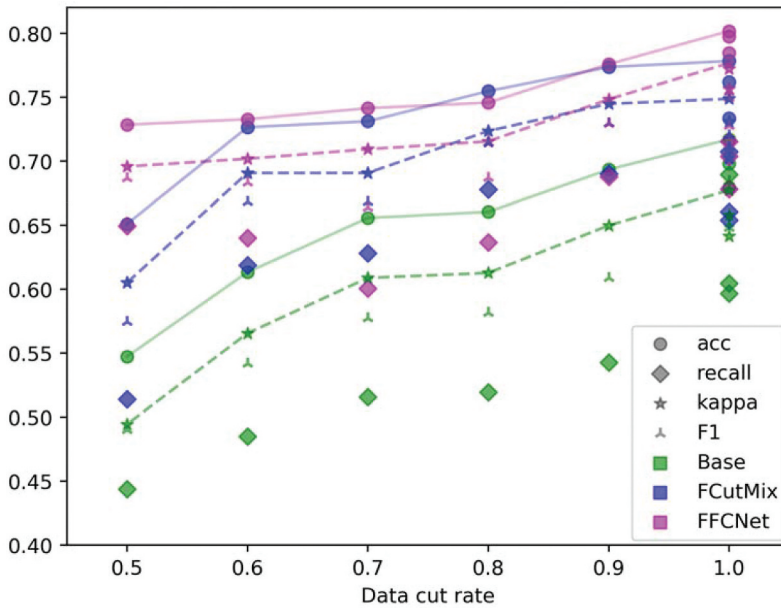
Figure 9. Results of multiple tests.
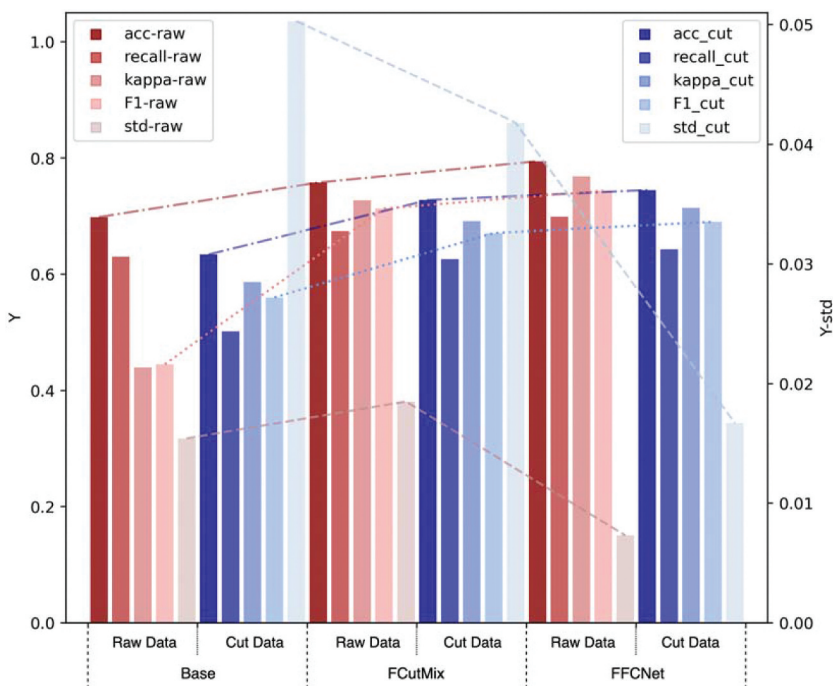
Figure 10. The mean value of test results.

Another experiment is designed to verify the credibility of the results and the generalization ability of the model. According to the standard that the verification set accounts for 30% of the total data volume, both sets of the training and the verfication are randomly divided, and the stability of the model is verified by multiple experiments. At the same time, in accordance with the characteristics of the study, the verification set is cut to 90%, 80%, 70%, 60% and 50% of the original respectively to simulate the different degrees of fragmentation of the tiles, and verify the generalization performance of the model. The results are shown in Table 1, Figures 9 and 10. Figure 9 shows a scatter plot of the results of multiple experiments. The mean value of the evaluation index of models in the raw data and the cut data is calculated, which is plot in Figure 10. The results show that the performance of all models increases as the degree of fragmentation decreases; Methods in this paper have greatly improved the accuracy and recall rate obtained in the experiments of

Table 2. Accuracy of each model.

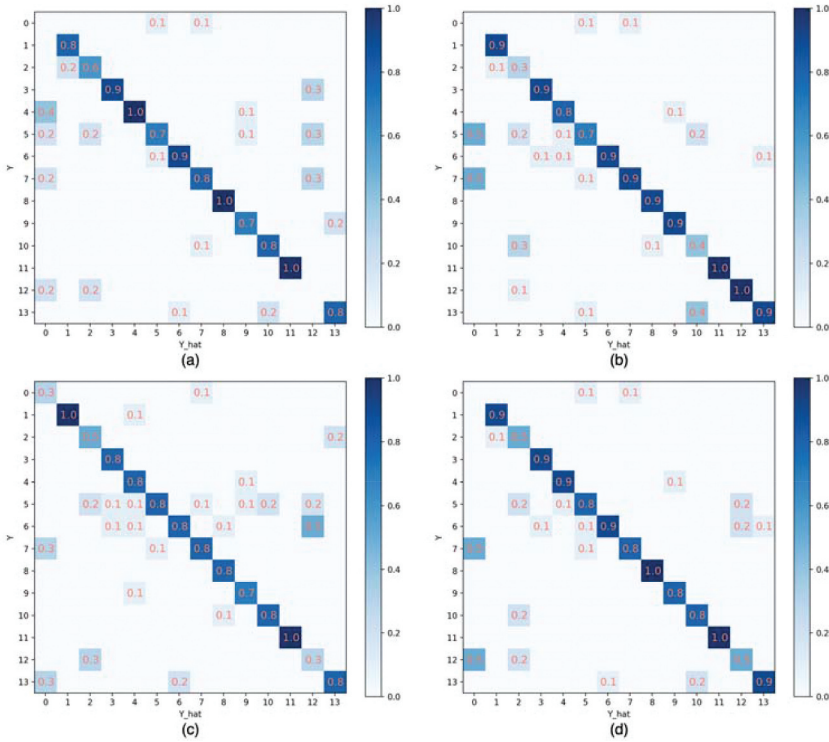|	Model 1	Model 2	Model 3	Integrated Learning (Voting strategy)	Integrated Learning (Average one-hot)	Integrated Learning (Weighted average one-hot)
Accuracy	79.31%	75.00%	80.17%	81.47%	81.90%	83.19%
Recall	66.76%	67.42%	70.36%	71.70%	72.88%	74.11%
F_1 Score	0.7249	0.7101	0.7495	0.7627	0.7713	0.7839
Kappa	0.7663	0.7176	0.7767	0.7905	0.7955	0.8103

Figure 11. Confusion matrix.

original fragments and cut fragments than the base model; Due to the random mixing of samples, the FCutMix method has a large variance in the experimental results, but the overall performance is still higher than the base model.

This paper uses the FCutMix augmentation method to train model 1 based on ResNet152 and uses ResNet18 and ResNet152 as the backbone network of FFCNet to train model 2 and model 3 respectively. The integrated strategy proposed in this paper is used to process the prediction results of the three models, in which weights are given to 1/3, 2/3, and 1 according to the difference in accuracy. The weight is set according to formula (5). $S(N_i, M)$ means the subscript of N_i in M, and M is arranged in ascending order of accuracy of N (the subscript starts from 1). The results are shown in Table 2, the integrated learning improves the accuracy by 3%, other indicators also prove that the method improves the performance of the model in an all-round way.

$$k_i = \frac{S(N_i, M)}{n}, M = sort\{N_1, N_2, \ldots, N_n\} \tag{5}$$

The confusion matrix is drawn according to the prediction results of the four models, as shown in Figure 11, where a, b, c, and d correspond to models 1, 2, 3 and the results obtained by integration learning. The figure shows such a clear result that the integrated learning strategy proposed in this paper can improve the accuracy rate as much as possible while ensuring the accuracy rate will not decrease, which is based on the current best result.

Conclusion

This paper is oriented to ceramic fragment classification tasks, proposes a multi-feature ceramic image classification model FFCNet that integrates GLCM texture features on the channel, which makes up for the shortcomings of convolutional networks that cannot show spatial correlation in images. Aiming at the problem of insufficient data and the shortcomings of traditional data augmentation methods, a data augmentation method FCutMix based on CutMix is proposed, which expands a large amount of data while ensuring the quality of samples. At the same time, a data preprocessing method based on color segmentation is proposed to eliminate irrelevant influencing factors in the picture. Finally, the integrated learning strategy is improved to further improve the accuracy. Experiments prove the effectiveness of each improved step and the better accuracy of this model. Future work will continue to research on data augmentation and feature fusion. Other data augmentation methods will be tried, such as some generative adversarial networks which are designed specially. Other feature fusion methods will also be considered, such as late fusion processing. In addition, the research will also be applied in related research such as ceramic repair.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

References

Abadi, M., M. Khoudeir, and S. Marchand. 2012. Gabor filter-based texture features to archaeological ceramic materials characterization. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 7340:333–42. doi:10.1007/978-3-642-31254-0_38.

Chen, P., S. Liu, H. Zhao and J. Jia. 2020. Gridmask data augmentation. arXiv Preprint arXiv:2001.04086.

Chetouani, A., T. Debroutelle, S. Treuillet, M. Exbrayat, and S. Jesset. 2018. Classification of ceramic shards based on convolutional neural network. Proceedings - International Conference on Image Processing, ICIP, Athens, Greece, 1038–42. doi:10.1109/ICIP.2018.8451728.

Debroutelle, T., S. Treuillet, A. Chetouani, M. Exbrayat, L. Martin, and S. Jesset. 2017. Automatic classification of ceramic sherds with relief motifs. *Journal of Electronic Imaging* 26 (2):23010. doi:10.1117/1.jei.26.2.023010.

DeVries, T., and G. W. Taylor. 2017. Improved regularization of convolutional neural networks with cutout. arXiv Preprint arXiv:1708.04552.

Gong, C., D. Wang, M. Li, V. Chandra, and Q. Liu. 2021. KeepAugment: A simple information-preserving data augmentation approach. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Online, 1055–64.

Han, C., K. Murao, S. Satoh and H. Nakayama. 2019. Learning more with less: Gan-based medical image augmentation. Medical Imaging Technology 37 (3):137–42

Hanzaei, S. H., A. Afshar, and F. Barazandeh. 2017. Automatic detection and classification of the ceramic tiles' surface defects. *Pattern Recognition* 66:174–89. doi:10.1016/j.patcog.2016.11.021.

He, K., X. Zhang, S. Ren and J. Sun. 2016. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, Las Vegas, USA, 770–78. doi:10.1109/CVPR.2016.90.

He, T., Z. Zhang, H. Zhang, Z. Zhang, J. Xie and M. Li. 2019. Bag of tricks for image classification with convolutional neural networks. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, USA, 558–67.

Jiangsu Provincial Study Society For Ancient Ceramics. 2010. *Flower and bird roll. Vol.1 of Chinese blue and white porcelain patterns*, ed. L. Shaobin, D. Wang, D. Zhou, and S. Xing, 13–183. NanJing, China: Southeast University Press.

Li, M., S. Lian, F. Wang, Y. Zhou, B. Chen, L.Guan and Y. Wu. 2020. Neural network modeling based double-population chaotic accelerated particle swarm optimization and diffusion theory for solubility prediction. Chemical Engineering Research and Design 2020, 155:98–107. doi:10.1016/j.cherd.2020.01.003.

Li, P., X. Li, and X. Long. 2020. FenceMask: A Data augmentation approach for pre-extracted image features. arXiv Preprint arXiv:2006.07877.

Liu, X., K. H. Ghazali, F. Han, and I. I. Mohamed. 2020. Automatic detection of oil palm tree from UAV images based on the deep learning method. *Applied Artificial Intelligence* 35 (1):13–24. doi:10.1080/08839514.2020.1831226.

Mukhoti, J., S. Dutta, and R. Sarkar. 2020. Handwritten digit classification in bangla and hindi using deep learning. *Applied Artificial Intelligence* 34 (14):1074–99. doi:10.1080/08839514.2020.1804228.

Shin, H. C., N. A. Tenenholtz, J. K. Rogers, C. G. Schwarz, M. L. Senjem, J. L. Gunter, K. P. Andriole and M. Michalski. 2018. Medical image synthesis for data augmentation and anonymization using generative adversarial networks. In International Workshop on Simulation and Synthesis in Medical Imaging, 1–11. doi:10.1007/978-3-030-00536-8_1.

Simonyan, K., and A. Zisserman 2014. Very deep convolutional networks for large-scale image recognition. *arXiv:1409.1556*.

Singh, K. K., H. Yu, A. Sarmasi, G. Pradeep and Y. J. Lee. 2018. Hide-and-seek: A data augmentation technique for weakly-supervised localization and beyond. arXiv Preprint arXiv:1811.02545.

Smith, P., D. Bespalov, S. Ali and P. Jeppson. 2010. Classification of archaeological ceramic fragments using texture and color descriptors. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, CVPRW, San Francisco, USA, 49–54. doi:10.1109/CVPRW.2010.5543523.

Szegedy, C., W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich. 2015. Going deeper with convolutions. Proceedings of the IEEE conference on computer vision and pattern recognition, Boston, USA, 1–9.

Taheri, S., and Ö. Toygar. 2019. Integrating feature extractors for the estimation of human facial age. *Applied Artificial Intelligence* 33 (5):379–98. doi:10.1080/08839514.2019.1577009.

Thepade, S. D., and P. R. Chaudhari. 2020. Land usage identification with fusion of thepade SBTC and sauvola thresholding features of aerial images using ensemble of machine learning algorithms. *Applied Artificial Intelligence* 35 (2):154–70. doi:10.1080/08839514.2020.1842627.

Yun, S., D. Han, S. J. Oh, S. Chun, J. Choe, and Y. Yoo. 2019. Cutmix: Regularization strategy to train strong classifiers with localizable features. Proceedings of the IEEE/CVF International Conference on Computer Vision, Long Beach, USA, 6023–32.

Zhang, H., M. Cisse, Y. N. Dauphin and D. Lopez-Paz. 2017. mixup: Beyond empirical risk minimization. arXiv Preprint arXiv:1710.09412.

Zhong, Z., L. Zheng, G. Kang, S. Li, Y. Yang. 2020. Random erasing data augmentation. *Proceedings of the AAAI Conference on Artificial Intelligence* 34 (7):13001–08. doi:10.1609/aaai.v34i07.7000.

Zhou, B., A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba. 2016. Learning deep features for discriminative localization. Proceedings of the IEEE conference on computer vision and pattern recognition, Las Vegas, USA, 2921–29.

Zhu, J. Y., T. Park, P. Isola, and A. A. Efros. 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. Proceedings of the IEEE international conference on computer vision, Honolulu, USA, 2223–32.