JIANG, W., WANG, X., REN, J., LI, S., SUN, M., WANG, Z. and JIN, J.S. 2021. MTFFNet: a multi-task feature fusion framework for Chinese painting classification. *Cognitive computation* [online], 13(5), pages 1287-1296. Available from: https://doi.org/10.1007/s12559-021-09896-9

MTFFNet: a multi-task feature fusion framework for Chinese painting classification.

JIANG, W., WANG, X., REN, J., LI, S., SUN, M., WANG, Z. and JIN, J.S.

2021

MTFFNet: a Multi-task Feature Fusion Framework for Chinese Painting Classification

Wei Jiang[1], Xiaoyu Wang[1], Jinchang Ren[2], Sen Li[1], Meijun Sun[1], Zheng Wang[1*], Jesse S. Jin[1]

[1] College of Intelligence and Computing, Tianjin University, Tianjin, China

[2] Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, UK

[*] Zheng Wang wzheng@tju.edu.com

Abstract
Different artists have their unique painting styles, which can be hardly recognized by ordinary people without professional knowledge. How to intelligently analyze such artistic styles via underlying features remains to be a challenging research problem. In this paper, we propose a novel multi-task feature fusion architecture (MTFFNet), for cognitive classification of traditional Chinese paintings. Specifically, by taking the full advantage of the pre-trained DenseNet as backbone, MTFFNet benefits from the fusion of two different types of feature information: semantic and brush stroke features. These features are learned from the RGB images and auxiliary gray-level co-occurrence matrix (GLCM) in an end-to-end manner, to enhance the discriminative power of the features for the first time. Through abundant experiments, our results demonstrate that our proposed model MTFFNet achieves significantly better classification performance than many state-of-the-art approaches. In this paper, an end-to-end multi-task feature fusion method for Chinese painting classification is proposed. We come up with a new model named MTFFNet, composed of two branches, in which one branch is top-level RGB feature learning and the other branch is low-level brush stroke feature learning. The semantic feature learning branch takes the original image of traditional Chinese painting as input, extracting the color and semantic information of the image, while the brush feature learning branch takes the GLCM feature map as input, extracting the texture and edge information of the image. Multi-kernel learning SVM (supporting vector machine) is selected as the final classifier. Evaluated by experiments, this method improves the accuracy of Chinese painting classification and enhances the generalization ability. By adopting the end-to-end multi-task feature fusion method, MTFFNet could extract more semantic features and texture information in the image. When compared with state-of-the-art classification method for Chinese painting, the proposed method achieves much higher accuracy on our proposed datasets, without lowering speed or efficiency. The proposed method provides an effective solution for cognitive classification of Chinese ink painting, where the accuracy and efficiency of the approach have been fully validated.

Keywords Multi-task feature fusion · Traditional Chinese paintings · Gray-level co-occurrence matrix

Introduction

As one of the most representative forms of ancient arts in China, traditional Chinese paintings have made significant contributions to the world's cultural heritage. Therefore, the protection of these paintings has become a problem to be addressed urgently. Fortunately, the recent advance of digital media and intelligent information processing technologies has provided an alternative way to digitize these antique priceless Chinese paintings and exhibit them on the Inter- net [1]. However, how to effectively manage and perform paintings classification in deep learning age is reported as a challenging problem due to the following reasons.

The first problem is the lack of a large and diverse painting dataset. To the best of our knowledge, there exists no suitable dataset designed for traditional Chinese paintings, which poses difficulty to transference of deep learning technology to this field. Second, due to the characteristics of Chinese art and similar techniques presented in many Chinese artists' artworks, it is usually difficult for researchers to fully extract the distinctive features from each painter's work and make an exact classification. Although most classical image classification algorithms based on manual feature extraction can obtain basic features, it is easy to suffer from information loss and poor generalization when working with 20 Chinese paintings. Third, Chinese painting mainly represents more abstract content via freehand brushwork style, which is quite different from realistic natural images. Therefore, more professional domain knowledge is needed for feature extraction from this kind of images.

To address these aforementioned problems, we first collect more than 5000 near-modern Chinese paintings (MCPs) from the

Eastern Jin Dynasty (A.D. 317) to now, which is provided by the Palace Museum and Tianjin Museum. We use class-level label to annotate these images. This newly established dataset makes it possible to implement Chinese painting classification based on deep learning. In Chinese art, especially ink painting, its texture [2] carries the information about brush stroke, which has the capability to reflect the style difference among artists, and the GLCM [3] is an alternative representation that can also fully capture the texture information. In addition, DenseNet [4] performs well in most image classification tasks. Therefore, we use DenseNet as the backbone to design a novel multi-task feature fusion network, called MTFFNet, to cognitively extract the underlying abstract styles of the Chinese paintings. In MTFFNet, we designed two branches to learn different pattern features end to end by taking the original image and its GLCM feature map as input, respectively. The fusion of these features enhances the discriminability of the final feature descriptor [5]. In order to avoid local extremum and overfitting problems existing in neural networks, we use SVM [6] rather than softmax as the final classifier to obtain better generalization ability. Finally, we evaluate the proposed MTFFNet on the benchmark dataset we constructed. The comprehensive experimental results show that our method can achieve promising accuracy, outperforming many state-of-the-art results on painting classification task.

Overall, the main contributions of this paper can be summarized as follows:

1. We proposed an end-to-end multi-task architecture (MTFFNet) to classify traditional Chinese paintings. Both branches use a well-designed DenseNet backbone network to learn different type of features of Chinese paintings.
2. It was the first time to introduce GLCM feature map to extract texture features, and integrate them with original semantic features from original images for traditional Chinese painting classification.
3. We have constructed a new near-modern Chinese painting (MCP) dataset, including 5000 traditional Chinese paintings with class-level annotations from 10 famous artists.
4. We used SVM instead of the original softmax structure for classification. Through comprehensive experiments, we proved that such design can tackle the over-fitting problem of the network, and improve the classification accuracy.

The remaining of our paper is organized as follows. In “Related Work,” we give a brief introduction of the related work. In “Chinese Painting Classification via MTFFNet,” the architecture of the proposed MTFFNet is elaborated in detail. In “Experiments,” we perform abundant experiments and compare the results with other mainstream model. Finally, in “Conclusion,” some concluding analyses are drawn.

Related Work

The field of Chinese paintings classification has been studied for decades. Wang et al. [6] proposed a hybrid two-dimensional multi-resolution hidden Markov model (MHMM) method to classify black and white Chinese paintings. Jiang et al. [7] proposed an algorithm to classify traditional Chinese paintings into realistic and freehand types via underlying features and SVM classifier. Johnson et al. [8] designed an authenticity identification system using color and texture analysis technology. However, these methods exploited merely a small part of representation attributes of Chinese painting, which cannot completely reflect the underlying characteristic of paintings. Li et al. [9] introduced brush stroke analysis into artist identification techniques, since different artists usually have their own brush stroke style contained in paintings. Niu and Suen et al. [10] considered the classification of western classical paintings using the classifier based on the RBF neural network to extract local and global features, but its performances are not convincing. Shi et al. [11] used statistical methods to compare Van Gogh and his contemporaries to analyze many automatically extracted brush stroke features. However, it is reported that the recognition results achieved by the methods discussed above are not satisfactory, because of the limitation of expressiveness for handcrafted features in depicting the content or brushwork information of paintings.

Recently, the successful application of deep learning technology in the tasks of handwritten digit recognition [12], human motion recognition [13], speech recognition [14], saliency detection [15], object detection [16], and ImageNet [17]-based classification has shown its considerably powerful ability in extracting discriminative feature. As a result, the trend of research on painting classification began to be shifted to deep learning-based approach. Krizhevsky et al. [18] proposed a deep convolutional neural network model to perform the painting style classification task. This framework took the photographic images as input, to learn the remarkable features for identifying the painting style. Experiment results showed that this method can achieve promising results. Recently, Deng et al. [19] separated the style and content of the paintings using multi-layer feature fusion of the deep convolutional neural network (CNN). This strategy of combining different features, such as multi-layer features from CNN, texture, and color features, performed particularly well in categorizing styles and artists.

In our work, we also choose the deep learning approach for traditional Chinese ink painting classification. Different from existing methods, we take advantage of DenseNet which shows promising results in classification tasks, as the key component of our multi-task architecture. Since brush stroke information generally is presented in the form of texture in the paintings, we also consider the artistic brush stroke by feeding both the original image and the GLCM feature map of paintings into each of the branch to yield different type features. The final fusion feature of these features improves the performance of Chinese paintings classification.

Chinese Painting Classification via MTFFNet

Multi-task Feature Fusion Architecture

In this section, we will introduce the proposed multi-task feature fusion network (MTFFNet) for Chinese painting classification in detail. Figure 1 gives the structure of our MTFFNet. The complete network mainly consists of two branches [20], including semantic feature learning and brush stroke feature learning, where both branches integrate pervasive DenseNet as the backbone network. The semantic feature learning task takes the images of traditional Chinese painting as input to learn semantically high-level information. Therefore, this branch describes the characteristics of paintings from the perspective of the original RGB image. The GLCM feature map is a commonly used representation to depict brush stroke contained in the paintings. In the brush stroke information learning task, we first generate four texture features based on the analysis of the painting, including contrast, energy, entropy, and homogeneity feature maps. Then, these four maps are linearly fused together according to respective weights assigned to them. Finally, the combinational feature map is fed into the brush stroke learning branch. The output of each task branch is a 1024-dimensional vector which is the number of kernels of the last convolution layer. It is worth noticing that the same length of the learned features makes it convenient to fuse them together. We can also give different weights to these two kinds of output feature to emphasize their importance in the classification task. At the end of the architecture, the multi-class multiple kernel learning [21] is used as the classifier to perform the final Chinese painting classification. Different from other previous work applying the multi-kernel method into image classification research, we obtain the classifier for multi-class instead of binary task. As a result, the final output of the classifier is a vector.

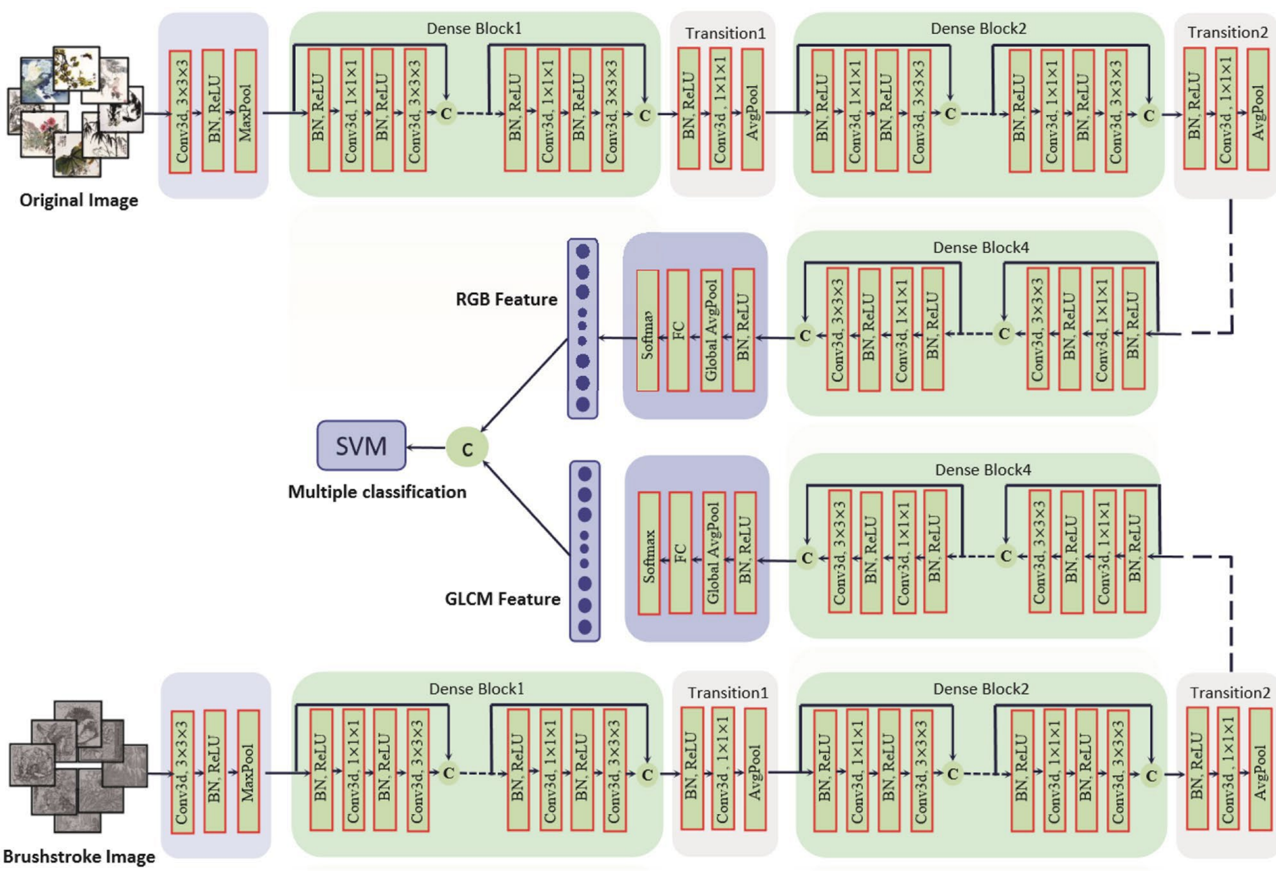


Fig. 1 The overall network architecture of our Chinese painting classifier

Semantic Feature Learning Task

As aforementioned discussion, the semantic feature learning task takes advantage of the original fine-art painting image as input to learn the high-level semantic representation.

With the layer-by-layer convolution, the semantic information contained in the convolutional layer becomes more and more abstract. In order to improve the ability of the network to express the target information, a multi-scale feature fusion strategy was introduced, which fused the multi-layer features of different scales together as the final feature. At present, most networks adopt the method of multi-layer feature fusion to obtain more abundant image information. In these networks, the recently proposed DenseNet has the best performance.

DenseNet establishes dense connections between all the front and rear layers. One of the advantages of DenseNet is the feature reuse, through the connection of features across the channel. These characteristics allow DenseNet to achieve better performance than other networks with fewer parameters and calculation costs. In order to obtain rich semantic information in the picture and ensure efficiency, DenseNet is adopted as the backbone network in our semantic feature learning task.

The output of this task branch is a 1024-dimensional vector. In this paper, we use two versions of DenseNet, including DenseNet121 and DenseNet169. We follow the previous work of Zabalza et al. [22] to set building blocks, the number of blocks stacked, and the down-sampling stages.

Brush Stroke Feature Learning Task

The brush stroke, as a fundamental part of paintings, carries information about texture, which plays an important role in painting analysis and classification. In order to extract texture information, we use the GLCM feature map as input of the brush stroke feature learning task.

GLCM provides the information about image gray direction, interval, and change amplitude. And the co-occurrence matrix is used to calculate corresponding eigenvalues, which can reflect the texture information of the image. In literature [21], 14 texture feature parameters based on GLCM were proposed. Here, we only extract four of these parameters with much stronger descriptive ability, including contrast, energy, entropy, and homogeneity.

1. Contrast

The contrast reflects the sharpness of the image and the depth of the texture. The gray difference means that the more pixel pairs with large contrast, the bigger Contrast (CON) will be Eq. 1:

$$CON = \sum_{n=0}^{Ng-1} n^2 \sum_{i=0}^{Ng-1} \sum_{j=0}^{Ng-1} P(i,j) \tag{1}$$

2. Energy

Energy is the sum of squares of all the elements in GLCM and reflects the evenness of gray distribution and the thickness of texture. When the elements in the co-occurrence matrix have concentrated distribution, the ASM gets a much larger value (Eq. 2).

$$ASM = \sum_{i=0}^{Ng-1} \sum_{j=0}^{Ng-1} (P(i,j))^2 \tag{2}$$

3. Entropy

Entropy is a measure of image information, representing the non-uniformity or complexity of image texture. When the elements in the co-occurrence matrix have dispersed distribution, entropy (ENT) gets a much larger value (Eq.3).

$$ENT = -\sum_{i=0}^{Ng-1} \sum_{j=0}^{Ng-1} P(i,j)\ln(P(i,j)) \tag{3}$$

4. Homegeneity

Both the homogeneity and local changes of image texture are reflected by the following formula. The large value indicates that there is no change between different areas of the image texture and the local area is quite uniform (Eq. 4).

$$H = \sum_{i=0}^{Ng-1}\sum_{j=0}^{Ng-1}\frac{P(i,j)}{1+(i-j)^2} \tag{4}$$

These four feature values reflect the information on gray-scale distribution and texture thickness of the image from different aspects. Therefore, in the brush stroke feature learning task branch, we first generate these four values of painting to form four kinds of texture feature images. Then, we linearly fuse these four images with different weights into a combinational one (Fig. 2). Finally, we feed the yielded texture image into the network branch using DenseNet as back-bone [23].The output of this task is also a 1024-dimensional vector.

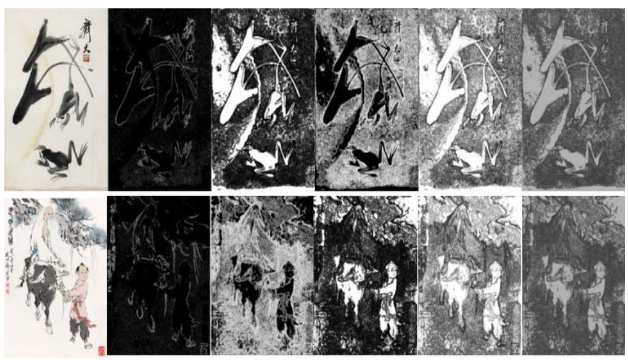


Fig. 2 The two rows respectively are Baishi Qi and Zeng Fan's paintings and their corresponding GLCM feature images. From left to right are an original painting, contrast, energy, entropy, homogeneity, and fused GLCM feature images

Implementation

We use deep learning framework TensorFlow and Keras implement our architecture. The MTFFNet is trained using the stochastic gradient descent (SGD) with a batch size of 64 images. By following the setting of the AlexNet, their learning rate for the training epoch p correlating to the current epoch i is set to be Eq. 5:

$$\varepsilon_i = 10^{-1-4\times\frac{i-1}{p-1}} \tag{5}$$

were p is a positive integer to ensure that the model is convergent. In our experiments, p is set to be 50.

We use LIBSVM Toolbox to implement the SVM classifier and use the Gaussian kernel and the grid optimization to find the optimal value C in the parameter space [2-10:1000] with a step of size 1. It has already been shown that transfer learning works quite well from CNNs pre-trained on natural images to paintings [24]. In order to overcome the limitation of the number of samples and obtain generalization ability, we use DenseNet which is pre-trained on the ImageNet dataset for our classification experiments, and then is fine-tuned with our fine-art painting dataset.

Experiments

Dataset

In deep learning age, training data exerts an effect on the performance of the classification model. Since there exists no suitable dataset containing large-scale and diversity Chinese painting images, we decided to construct a dataset to evaluate our proposed model. We collected almost 5000 traditional Chinese paintings showing different artistic styles from 10 famous Chinese artists, Jianlou Cao, Zeng Fan, Xiaoming Li, Yanshao Lu, Tianshou Pan, Baishi Qi, Changshuo Wu, Beihong Xu, Xiaolian Zeng, and Da Zhu, which are provided by the Palace Museum and Tianjin Museum. We label these images manually with class-level annotations to match our classification task. We use 8 data augmentation methods to expand the dataset and divide them into a training set, verification set, and test set according to the proportion 7:1:2 (Fig. 3).

Fig. 3 Illustration of five artists' sample paintings that are randomly selected from the Dataset. Each row represents one of the five artists, including ZHU Da, CAO Jianlou, LU Yanshao, LI Xiaoming, and QI Baishi

Experimental Setting

From the previous discussion, our proposed architecture mainly consists of two parts. The first part is a multi-task feature extraction module. Two different types of features, the main RGB information feature vector and brush stroke information feature vector, are extracted from the input images of paintings [25]. To ensure better results, we pre-process the input by data augmentation strategies and perform normalization and dimensionality reduction operations on feature data. That is, these two kinds of features are first normalized between 0 and 1. To speed up the algorithm and reduce redundancy, then, principal component analysis (PCA) [23] is used to reduce the dimension to 100. The second part of our model is classification module, including feature fusion stage and classification stage. After obtaining the RGB learning-based and brush stroke learning-based feature, we fuse them using add function. And unlike the traditional convolution neural network, here we use the SVM with a radial basis function kernel instead of the fully connected layers as classifier because SVM [26] is more suitable for our Chinese painting classification task. For the setting of SVM in our experiment, the best parameter C and G is selected by five-fold cross-validation. We train our model in an end-to-end manner. Precision and recall were used to measure the performance of the classification method. Figure 4 gives the overall structure of our proposed model.

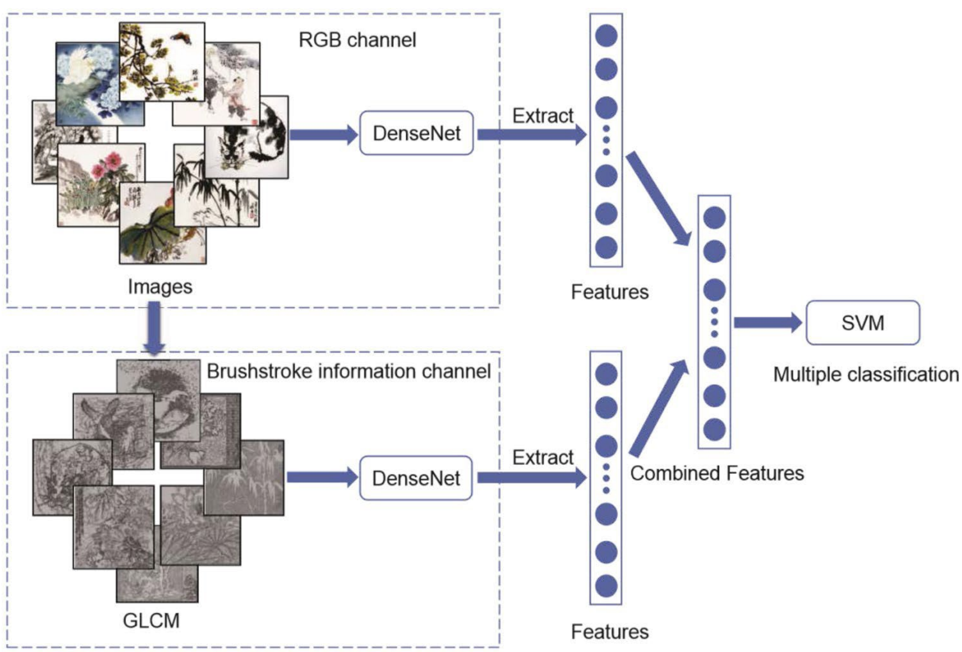
Table 1 Comparison the classification accuracy of different methods

Models	Accuracy (%)
Saleh and Elgammal [27]	63.06
Tan et al. [28]	76.11
Huang et al. [29]	81.87
Qian et al. [30]	82.15
Jia [31]	83.32
Simon et al. [32]	82.63
Utgoff et al. [33]	65.52
Li [9]	74.17
MTFFNet	94.93

Experimental Result

In this section, we present an evaluation of our proposed model with a comparison to the state-of-the-art methods, containing deep learning–based methods such as Saleh and Elgammal [27], Tan et al. [28], Huang et al. [29], Qian et al. [30], and Jia [31], and traditional machine learning-based methods, Sparse group LASSO, Decision Tree C4.5, and SVM. The experiments are conducted on the Chinese paintings dataset we built. Table 1 shows the com- parison experimental results of 8 selected methods. The proposed MTFFNet model significantly outperforms the selected methods, achieving 94.93% classification accuracy. In the process of comparison, we used the same dataset as in the previous experiment, and 10 times random sampling is also conducted to get the result. Compared with previous works for the classification of Chinese painting, MTFFNet performs very well.

Fig. 4 The structure of the proposed model



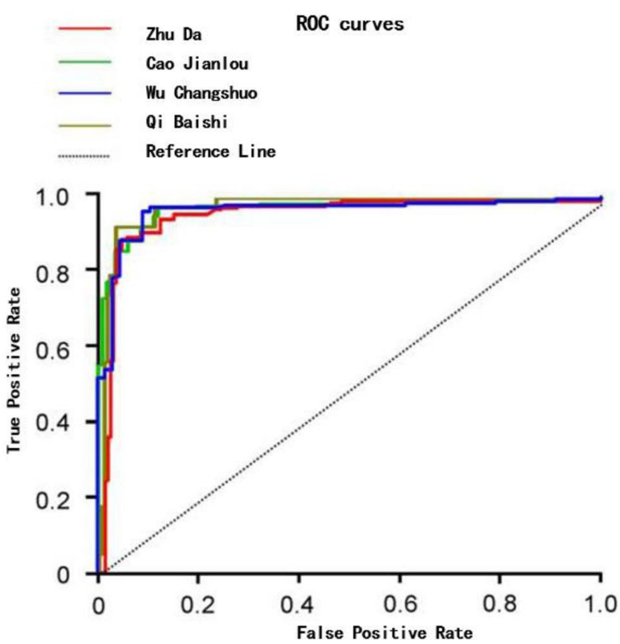


Fig. 5 ROC curves for four painters of MTFFNet

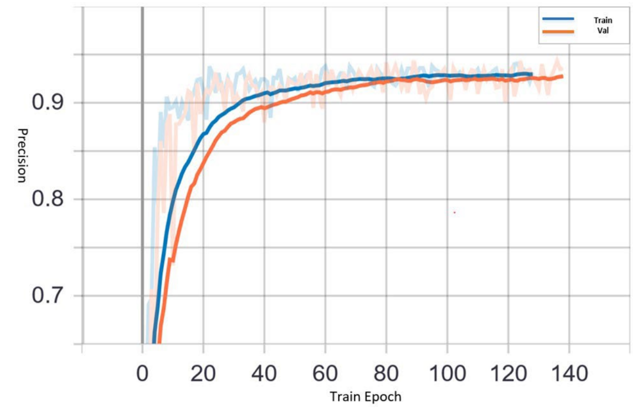


Fig. 6 Training process of MTFFNet

As shown in Fig. 5, we show the ROC curves of four painters with similar styles. The AUC values of MTFFNet for the three painters all exceed 0.90, indicating that the method proposed in this paper has good generalization ability.

In order to make the training process clearer and prove that our method has indeed achieved good training effects, we visualized the training process, as shown in Fig. 6.

Ablation study

To analyze our design choices and provide more insights, we carry out comprehensive ablation experiments on our dataset.

DenseNet or Other Network as Backbone?

Here, we choose DenseNet, ResNet, and VGG with different layers to do the classification task of fine-art paintings to further show their feature extraction ability. And all these deep learning models are pre-trained with ImageNet to obtain higher accuracy. Table 2 shows the precision, recall, F1 score, and macro accuracy achieved by these approaches evaluated for the task of Chinese painting classification.

We visualized the loss during the training process as shown in Fig. 7. Compared with other backbones used for Chinese painting classification, the loss of DenseNet is smoother and easier to converge.

Figure 8 gives the confusion matrix of different models on our dataset. It is shown that DenseNet gets the much higher accuracy on each category than other networks. In this experiment, we observed that some of Wu's paintings are similar to Cao's, so we plan to test these two kinds of similar paintings further as follows.

Figure 9 shows the prediction results in the form of histogram. The first row are Jianlou Cao's painting and corresponding network prediction histograms, and the second row are Changshuo Wu's painting and corresponding network prediction histograms. This experiment further demonstrates that DenseNet is capable of learning much more discriminative features than other networks, especially when dealing with images with similar appearance.

Table 2 Comparison of different networks on precision, recall, F1-score, and macro-accuracy in the task of Chinese painting classification

Performance	Network				
	VGG16	VGG19	ResNet50	DenseNet121	DenseNet169
Precision	0.93	0.92	0.93	0.94	0.95
Recall	0.92	0.88	0.91	0.93	0.94
Fl-score	0.92	0.90	0.91	0.93	0.94
Macro-accuracy	0.93	0.92	0.92	0.94	0.95

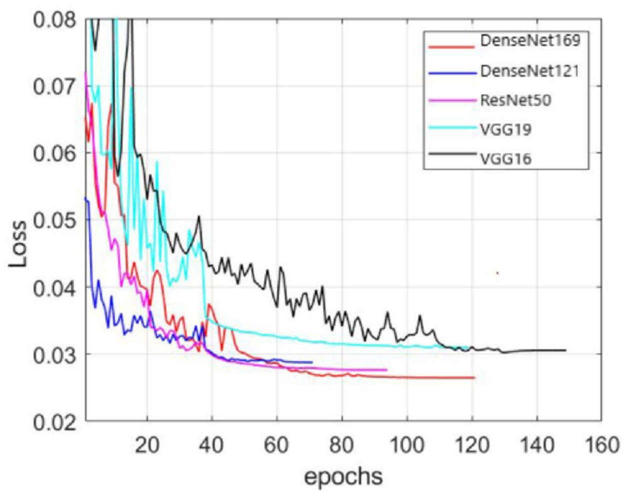


Fig. 7 Training loss of MTFFNet with different backbones

Analysis of the Classifiers

An important question is whether we need the brush stroke input for Chinese painting classification, and whether the SVM is better than softmax on that task. Table 3 includes a direct comparison between our proposed multi-task architecture with different networks as backbone and with different 250 classifiers against the exact same counterparts without using brush stroke input. The networks we selected are VGG16, ResNet50, and DenseNet169.

Table 3 Top 1 accuracy comparison of different networks on two- channel and RGB channels. Besides, we also compare the precision of the same network with softmax or SVM as a classifier in the task of Chinese painting classification

Top 1 accuracy		Stroke channel	RGB channel	Combined channel
VGG16	+ SVM	0.82	0.93	0.94
	+ Softmax	0.76	0.81	0.83
VGG19	+ SVM	0.83	0.90	0.92
	+ Softmax	0.75	0.78	0.80
ResNet50	+ SVM	0.86	0.92	0.93
	+ Softmax	0.78	0.81	0.82
DenseNet169	+ SVM	0.89	0.94	0.96
	+ Softmax	0.85	0.92	0.93

From Table 3, we can draw the following conclusions.

1. In either two-task cases or single task cases, combination with SVM clearly achieves higher accuracy than that with softmax, which means that SVM plays an important role in painting classification tasks. For example, VGG16 with SVM obtains 94% accuracy in two tasks, which is just 1% less than DenseNet with SVM.
2. Under the same experimental settings, both DenseNet169 with SVM and DenseNet169 with softmax yielded the best performance compared to other networks, which further demonstrate that the underlying features learned by DenseNet is much more powerful and discriminative.
3. As for the influence of brush stroke information on Chinese painting classification, it can be reported that the overall accuracy of the multi-task network integrating stroke learning is higher than the single-task network only using the RGB image. The integration of brush stroke learning helps improve the performance. The overall accuracy of the multi-task network is higher than the single-task network, which indicates that the brush stroke information channel proposed in this paper plays a certain role with the overall improvement of about 2%.

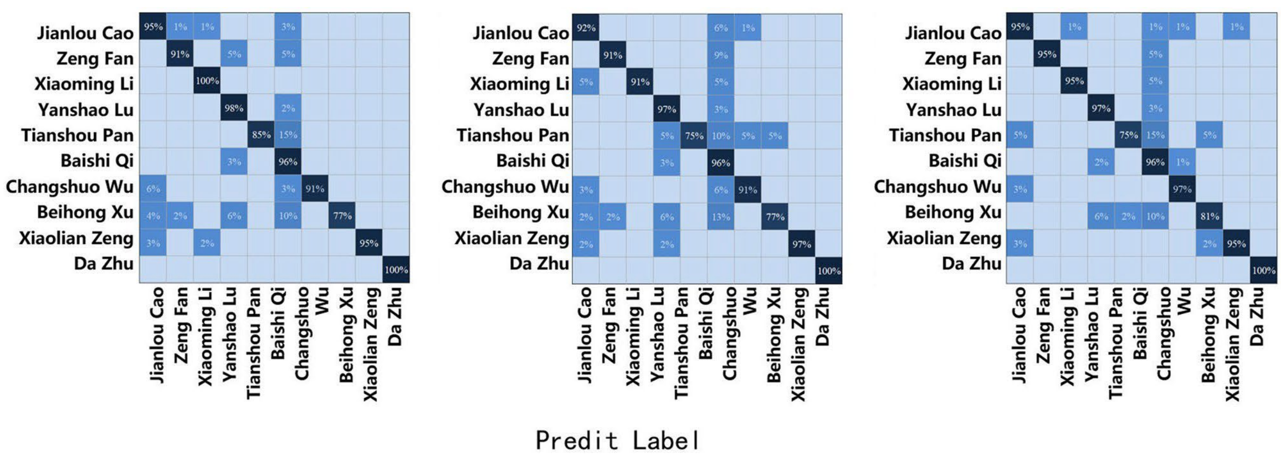


Fig. 8 From left to right are visual representations of confusion matrix used by DenseNet169, ResNet50, and VGG16 as backbone for classification of Chinese paintings

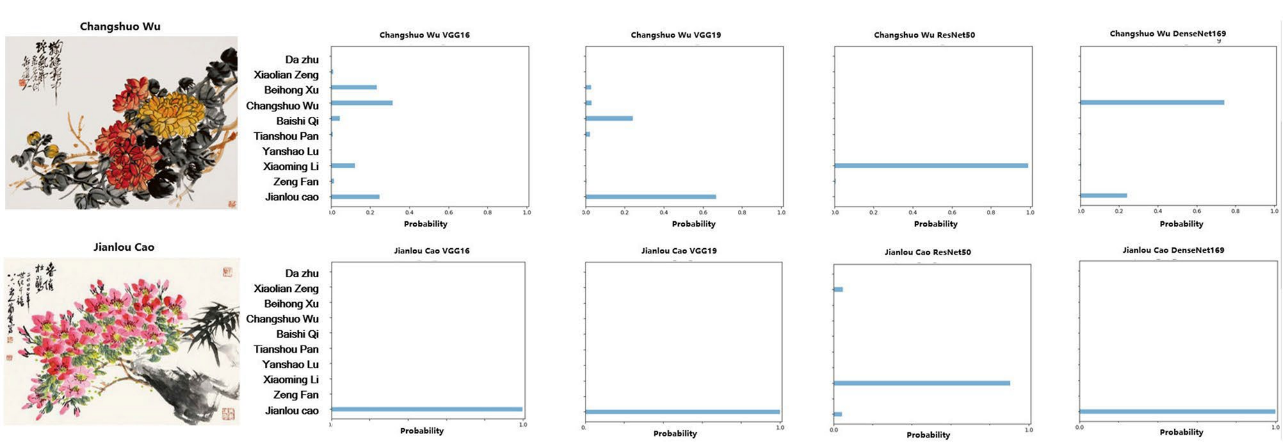


Fig. 9 Network prediction histogram

Conclusion

When classifying the traditional Chinese paintings, brush stroke is an important and powerful feature for understanding the textured pattern of the paintings. However, there are few reported works that take the brush stroke into account for classification of Chinese paintings. In this paper, we proposed an end-to-end multi-task feature fusion network named MTFFNet, for fine-art painting classification. By combination of two branches in the network, we can extract and fuse both the semantic features and the brush stroke information, where the GLCM feature map is found to be an effective input for measuring the texture feature. Eventually, the SVM also demonstrates superior performance for classification of the extracted features. Comprehensive experiments, including comparison to other models and ablation studies analysis, have fully validated the efficiency of our proposed approach.

Although MTFFNet proposed in this paper can effectively improve the classification accuracy of Chinese paintings, the running time of the algorithm will be affected by the independent feature extraction of the two branches. Therefore, a future improvement direction is how to reduce the running time of the model.

Funding This research work was financially supported by the National Natural Science Foundation, China, under grant nos. 61772360 and 61876125.

Declarations

Conflict of Interest The authors declare no competing interests.

References

1. Zabalza J, Ren J, Zheng J, et al. Novel segmented stacked autoencoder for effective dimensionality reduction and feature extraction in hyperspectral imaging. Neurocomputing. 2016;185:1–10.
2. Zhang D, Han J, Zhao L, et al. Leveraging prior-knowledge for weakly supervised object detection under a collaborative self-paced curriculum learning framework. Int J Comput Vision. 2019;1:4–10.
3. Ren J, ANN vs. SVM: Which one performs better in classification of MCCs in mammogram imaging. Knowledge-Based Systems. 2012;26:144–153.
4. Huang G, Liu Z, Weinberger K, Maaten L. Densely connected convolutional networks . arXiv preprint arXiv. 2016;1608(06993):1–4.
5. Li J, Wang J. Studying digital imagery of ancient paintings by mixtures of stochastic models. IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society. 2004;13(3):340–353.
6. Wang JZ. Image processing for artist identification - computerized analysis of Vincent van Gogh's painting brushstrokes. Nuclear instruments & methods in physics research. 2008;122(4):650–656.
7. Jiang S, Huang Q, Ye Q, et al. An effective method to detect and categorize digitized traditional Chinese paintings. Pattern Recogn Lett. 2006;27(7):734–746.
8. Johnson CR, Hendriks E, Berezhnoy IJ, et al. Image processing for artist identification. Signal Processing Magazine IEEE. 2008;25(4):37–48.
9. Li J. Rhythmic brushstrokes distinguish van Gogh from his contemporaries: findings via automated brushstroke extraction. IEEE Trans Pattern Anal Mach Intell. 2012;34(6):1159–1176.
10. Niu XX, Suen CY. A novel hybrid CNN–SVM classifier for recognizing handwritten digits. Pattern Recogn. 2012;45(4):1318–1325.
11. Shi L, Zhang Y, Cheng J, Lu H. Skeleton-based action recognition with multi-stream adaptive graph convolutional networks. IEEE. 2019;29:3247–3257.
12. Padfield N, Zabalza J, Zhao H, et al. EEG-based brain-computer interfaces using motor-imagery: techniques and challenges[J]. Sensors. 2019;19(6):1–5.
13. Abdel-Hamid O, Mohamed AR, Jiang H, et al. Convolutional neural networks for speech recognition. IEEE/ACM Transactions on Audio Speech & Language Processing. 2014;22(10):1533–1545.
14. Han J, Cheng G, Li Z, et al. A unified metric learning-based framework for co-saliency detection. IEEE Trans Circuits Syst Video Technol. 2018;28(10):2473–83.
15. Leibe B, Matas J, Sebe N, et al. [Lecture Notes in Computer Science] Computer Vision -ECCV 2016 Volume 9912. Saliency Detection via Combining Region-Level and Pixel-Level Predictions with CNNs https://doi.org/10.1007/978-3-319-46484- 8(Chapter49):809-825.
16. Niu M, Song K, Huang L, et al. Unsupervised saliency detection of rail surface defects using stereoscopic images. IEEE Transactions on Industrial Informatics. PP(99):1–1.
17. Gong C, Junwei H, Peicheng Z, et al. Learning rotation-invariant and Fisher discriminative convolutional neural networks for object detection. IEEE Trans Image Process. 2018;28:265–278.
18. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems 25 (NIPS'2012). 2012:1–5.
19. Deng J, Dong W, Socher R, et al. ImageNet: A large-scale hierarchical image database. Proc of IEEE Computer Vision & Pattern Recognition. 2009;1:248–255.
20. Thomas L. The classification of style in painting. Dissertation Abstracts International, Volume: 66–09, Section: B, page: 4908. Advisers: Charles Tappe, 2008.
21. Saleh B, Elgammal A. Large-scale classification of fine-art paintings: learning the right metric on the right feature. J Comput Sci Technol. 2015;25(3):595–605.
22. Zabalza J, Ren J, Zheng J, et al. Novel two-dimensional singular spectrum analysis for effective feature extraction and data classification in hyperspectral imaging. IEEE Transactions on Geoence and Remote Sensing. 2015;53(8):1–16.
23. Zabalza J, Ren J, Yang M, et al. Novel Folded-PCA for improved feature extraction and data reduction with hyperspectral imaging and SAR in remote sensing. IsprJournal of Photogrammetry & Remote Sensing. 2014;93(7):112–122.
24. Zamani F, Jamzad M. A feature fusion based localized multiple kernel learning system for real world image classification. Eurasip J Image Video Process. 2017;2017(1):75–8.
25. Liu P, Guo JM, Chamnongthai K, et al. Fusion of color histogram and LBP-based features for texture image retrieval and classification. Information ences. 2017;390:95–111.
26. Niazmardi S, Demir B, Bruzzone L, et al. Multiple kernel learning for remote sensing image classification. IEEE Transactions on Geoence & Remote Sensing. 2018;1:1–19.
27. Saleh B, Elgammal A. Large-scale classification of fine-art paintings: learning the right metric on the right feature. J Comput Sci Technol. 2015;25(3):595–605.
28. Tan W, Chan C, Aguirre H, et al. Ceci n'est pas une pipe: A deep convolutional network for fine-art paintings classification. 2016 IEEE International Conference on Image Processing (ICIP). 2016;3703–3707.
29. Huang X, Zhong SH, Xiao Z. Fine-art painting classification via two-channel deep residual network. Pacific Rim Conference on Multimedia. Springer, Cham. 2017(2):1–5.
30. Qian W, Xu D, Guan Z, et al. Simulating chalk art style painting. International Journal of Pattern Recognition and Artificial Intelligence. 2017;31(12):1759026.1–20.
31. Jia S. Automatic categorization of traditional Chinese paintings based on wavelet transform. Comput Sci. 2014;41(2):317–319.
32. Simon N, Friedman J, Hastie T, et al. A sparse-group Lasso. J Comput Graph Stat. 2013;22(2):231–245.
33. Utgoff PE, Berkman NC, Clouse JA. Decision tree induction based on efficient tree restructuring[J]. Mach Learn. 1997;29(1):5–44.