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

Meijun Sun · Xiaoyu Wang · Jinchang Ren · Zheng Wang* · Zhou Yan* · Zhao Liming

*** Corresponding author. Email: wzheng@tju.edu.com**

Received: date/ Accepted

Structural abstract

Background/introduction

Different artists have their unique painting styles, which can be hardly recognized by ordinary people without essential 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, namely 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 information, semantic and brush stroke features. These features are learned from the RGB images and auxiliary Gray-Level Co-occurrence Matrix in an end-to-end manner to enhance the discriminative power of the features for the first time. Experimental results have demonstrated that our proposed model achieves significantly better classification performance than many state-of-the-art approaches.

Meijun Sun · Xiaoyu Wang · Zheng Wang

College of Intelligence and Computing, Tianjin University, Tianjin Key Lab of Machine Learning, Tianjin, China

* Zheng Wang is the corresponding author. Email: wzheng@tju.edu.com

Jinchang Ren

Department of Electronic and Electrical Engineering, University of Strathclyde

Zhou Yan · Zhao Liming

College of Management and Economics, Tianjin University

* Zhou Yan is the corresponding author. Email: zhouyantj@tju.edu.cn

Methods

In this paper, an end-to-end multi-task feature fusion method for Chinese painting classification is disclosed. The method includes: constructing a multi-task feature fusion network model, which is composed of two branches of top-level RGB image feature learning and low-level brush stroke feature learning. The top RGB image feature learning branch takes the original image of traditional Chinese painting as input, while the bottom brush feature learning branch takes the feature map of grayscale symbiosis matrix as input. They have learned different modal characteristics end-to-end. Multi kernel learning SVM is selected as the final classifier. The present invention solves the problem that the classification of Chinese paintings lacks of a large number of diverse training data and is prone to information loss and poor generalization ability.

Results

By adopting end-to-end multi-task feature fusion method, the proposed model improves the mAP by about 5% without bringing any additional computational cost on inference. When compared with the state-of-the-art classification method for cognitive Chinese painting, the proposed method achieves much improved results on our proposed datasets, without lowering its high efficiency.

Conclusions

The proposed approach has provided an effective solution for cognitive classification of Chinese painting, where the efficacy and efficiency of the approach have been fully validated.

1 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 Internet [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 is the lack of a large and diverse training data. To the best of our knowledge, there exist 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, therefore, 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 based on a shallow structure learning algorithm can extract certain image attributes, it is easy to suffer from information loss and poor generalization when working 20 with 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, in this paper, we first collect more than 5,000 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 deep learning-based painting classification approach. Since in Chinese art, especially ink painting, its texture [2] [3] carries the information about brush stroke, which has the capability to reflect the style difference among artists, and the Gray-Level Co-occurrence Matrix (GLCM) [4] is an alternative representation that can also fully capture the texture information. On the other hand, DenseNet [5][6] performs well in most image classification tasks. Therefore, we use the DenseNet as the backbone to design a novel multi-task feature fusion network, called MTFFNet, to cognitively represent the underlying abstract styles of the Chinese paintings. In MTFFNet, two branches take as input the image of painting and GLCM representation of painting, respectively, to learn different modalities features in an end-to-end manner. The fusion of these features enhances discriminative power of final description. [7] Then, to avoid the local extremum and over-fitting problem appeared in neural network, we choose SVM [8,9] instead of SoftMax as our final classifier to obtain a much better generalization ability. Finally, we evaluate the proposed MTFFNet approach 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:

i)We proposed an end-to-end multiple tasks architecture, namely MTFFNet, classify the traditional Chinese paintings, where both branches resort to the well-designed DenseNet as backbone to learn different features.

ii)It is the first time the GLCM modality is incorporated into deep learning-based painting classification framework for extracting the texture features, which is fused with the xxx features for classification.

iii) We have constructed a new near-Modern Chinese Paintings (MCPs) dataset, which 5000 traditional Chinese paintings with class-level annotations from 10 famous artists for validate the deep learning models.

iv)We used SVM to replace the original SoftMax structure, which can tackle the over-fitting of the network and improve the classification accuracy as validated in the comprehensive experiments.

The remaining of our paper is organized as follows. Section 2 gives a brief introduction to the related work. The architecture of the proposed MTFFNET is elaborated in Section 3. Experimental results are summarized in Section 4 along with the performance evaluation and comparison. Finally, some concluding remarks are drawn in Section 5.

2 Related work

The field of Chinese paintings classification has been studied for decades. [10][11] Jia et al. proposed a hybrid two-dimensional Multiresolution Hidden Markov Model (MHMM) method to classify black and white Chinese paintings. Berezhnoy et al. [134] designed an authenticity identification system using color and texture analysis technology. Jiang et al. [12] [33] proposed an algorithm to classify traditional Chinese paintings by realistic and freehand types via underlying features and SVM classifier. However, these methods exploited merely a small part of representation attributes of Chinese painting, which cannot completely reflect the underlying characteristic of paintings. Johnson et al. [13][35] introduced brush stroke analysis into artist identification techniques since different artist usually has their own brushwork style contained in their paintings. Shen et al. [14] considered the classification of western classical paintings using the classifier based on RBF neural network to extract local and global features, but its performances not convincing. Li et al. [15] used statistical methods to compare Van Gogh and his contemporaries to analyze many automatically extracted brush stroke. However, it is reported that the recognition results achieved by the methods discussed above are not satisfactory because of the limitation of expressiveness of the handcrafted features extracted to portray the content or brushwork information for paintings.

Recently, the successful application of deep learning technology in the tasks of handwritten digit recognition [16][17], human motion recognition[18][19], speech recognition[20], saliency detection[21][22][23],object detection[22], and ImageNet based[24][25] classification has shown its considerably powerful ability to carry out discriminative feature extraction. As a result, the trend of research on painting classification began to be shifted to deep learning-based approach. Karayev et al. [26] [36]proposed a deep convolutional neural networks model to perform the paintings style classification task. This framework took the photographic images as input to learn the remarkable features for identifying the paintings style. Experimental results showed that this method can achieve promising results. Recently, Gatys et al. [27] separated the style and content of the paintings using layered features of the deep convolutional neural network. This strategy of combination of different features, such as features from CNN, texture, and color features, performed particularly well in categorizing styles and artists [28][29][33].

In our work, we also choose deep learning-based approach for traditional Chinese paintings classification. Different from existing methods, we take advantage of the 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 image and Gray-Level Co-Occurrence Matrix of paintings to each of branch to yield different high-level texture features. The final fusion of these features improves the performance of Chinese paintings classification.

3Chinese Paintings classification via MTFFNet

3.1 Multi-task Feature Fusion Architecture

In this section, we will introduce the proposed Multi-task Feature Fusion (MTFFNet) architecture for Chinese paintings classification in much more detail. Fig. 1 gives the structure of our MTFFNet. the complete network mainly consists of two branches [30], namely RGB image feature learning and brush stroke feature learning, where both integrate the pervasive DenseNet as the important composition. The RGB image feature learning task takes the images of traditional Chinese painting as input to learn semantically high-level information to describe the characteristics of paintings from the perspective of RGB image. Since the Gray-Level Co-occurrence Matrix (GLCM) image is a commonly used representations to depict brush stroke contained in the paintings, so 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 map is fed into the learning branch. The output of each task branch is a 1024-dimensional vector that is the number of kernels of the last convolution layer. And it is worth pointing out that the same length of the learned features makes it convenient to fuse them together and we can also give different weights to these two kinds of output feature to emphasis their importance in classification task. At the end of the architecture, the Multi-class Multiple Kernel Learning [33] is used as the classifier to perform the final Chinese painting classification. Different from other previous work applying the multi-kernel method to image classification research, we obtain the classifier of multiple tasks instead of binary task, as a result, the final output of the classifier is in the form of vector.

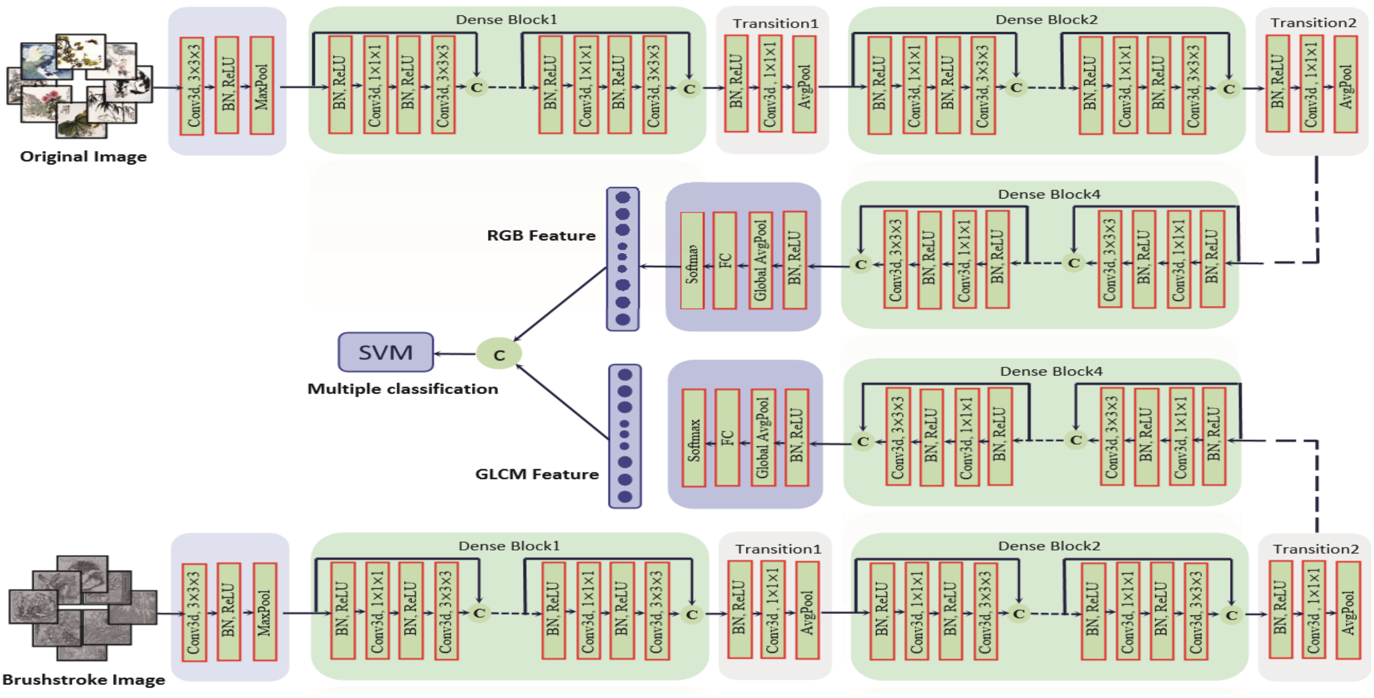

Figure 1: The overall network architecure of our Chinese painting classifier.

3.2 RGB Image Feature Learning Task

As aforementioned discussion, the RGB image feature learning task takes advantage of the original fine-art painting image as input to learn the high-level representation. The output of this task branch is a 1024-dimensional vector. In this paper, we use two versions of DenseNet, including DenseNet121 andDenseNet169. We follow the previous work of He et al. [38] to set building blocks, the number of blocks stacked, and the down-sampling stages.

3.3 Brush Stroke Information 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 Gray-Level Co-occurrence Matrix (GLCM)as input of the brush stroke information learning task. Here is the generation process of GLCM. Suppose there are $Nc \times Nc$ pixels in horizontal and vertical directions of the image, respectively. We quantize the grayscale of each pixel to Ng levels. Let $Lx = \{1,2..., Nc\}$, $Ly = \{1,2...Nc\}$,and $G = \{1,2...,Ng\}$ denote the horizontal space domain, the vertical space domain and the quantized intensity level set. $Lx \times Ly$ is the set of the image pixels ordered by row and column. Then the image function f can be expressed as such a function: Assigning each pixel one value G of Ng intensity levels

$$f : Lx \times Ly \rightarrow G. \tag{1}$$

The GLCM is defined as the probability of the occurrence of two pixels with a distance of d and direction of in the image domain within the range of $Lx \times Ly$:

$$P(\mathrm{i},\mathrm{j}|\mathrm{d},\theta) = \{[(k,l),(m,n)] \in (Lx * Ly)^2 \mid d,\theta,f(k,l)=i,f(m,n)=j\} \tag{2}$$

$\{x\}$ represents the valid pixel log. In this paper, we use the GLCM with distance d and horizontal direction as follows.

$$P(i,j|d,\theta) = \{[(k,l),(m,n)] \in (Lx * Ly)^2 \, |k-m=0, \; |l-n|=d, f(k,l)=i, f(m,n)=j\} \quad (3)$$

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 [29],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 CON will be:

$$CON = \sum_{n=0}^{Ng-1} n^2 \sum_{i=0}^{Ng-1} \sum_{j=0}^{Ng-1} P(i,j) \quad (4)$$

(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 has concentrated distribution, the ASM gets much larger value.

$$ASM = \sum_{i=0}^{Ng-1} \sum_{j=0}^{Ng-1} (P(i,j))^2 \quad (5)$$

(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 gets much larger value.

$$ENT = -\sum_{i=0}^{Ng-1} \sum_{j=0}^{Ng-1} P(i,j) \ln(P(i,j)) \quad (6)$$

(4) Homogeneity

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.

$$\mathrm{IDM} = \sum_{\mathrm{i}=0}^{\mathrm{Ng}-1} \sum_{\mathrm{j}=0}^{\mathrm{Ng}-1} \frac{\mathrm{P(i,j)}}{1+(\mathrm{i}-\mathrm{j})^2} \quad (7)$$

These four feature values reflect the information on grayscale distribution and texture thickness of the image from different aspect. Therefore, in Brush 160 Stroke Information 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. Finally, we feed the yielded texture image into the network branch using DenseNet as backbone. [39] The output of this task is also a 1024-dimensional vector.

3.4 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 [24],their learning rate for the training epoch p correlating to the current epoch i is set to be

$$\varepsilon_{\mathrm{i}} = 10^{-1-4\times\frac{\mathrm{i}-1}{\mathrm{p}-1}} \quad (8)$$

where 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 one [37]. It has already been shown that transfer learning works quite well from CNNs pre-trained on natural images to paintings [26] [27]. So, in order to overcome the limitation of the number of samples, we use DenseNet, which is pre-trained on the ImageNet dataset for our classification experiments and then is fine-tuned 175 with the our fine-art paintings dataset.

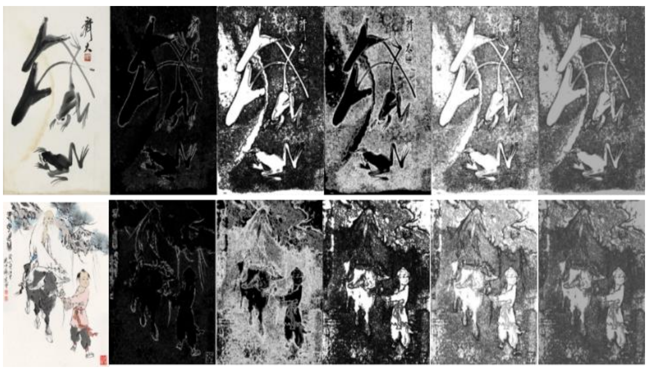

Figure 2: The two rows respectively are Baishi Qi and Zeng Fan's paingtings and their corresponding GLCM feature images. From left to right are an original painting, contrast, energy, entropy, homogeneity, and fused GLCM feature images.

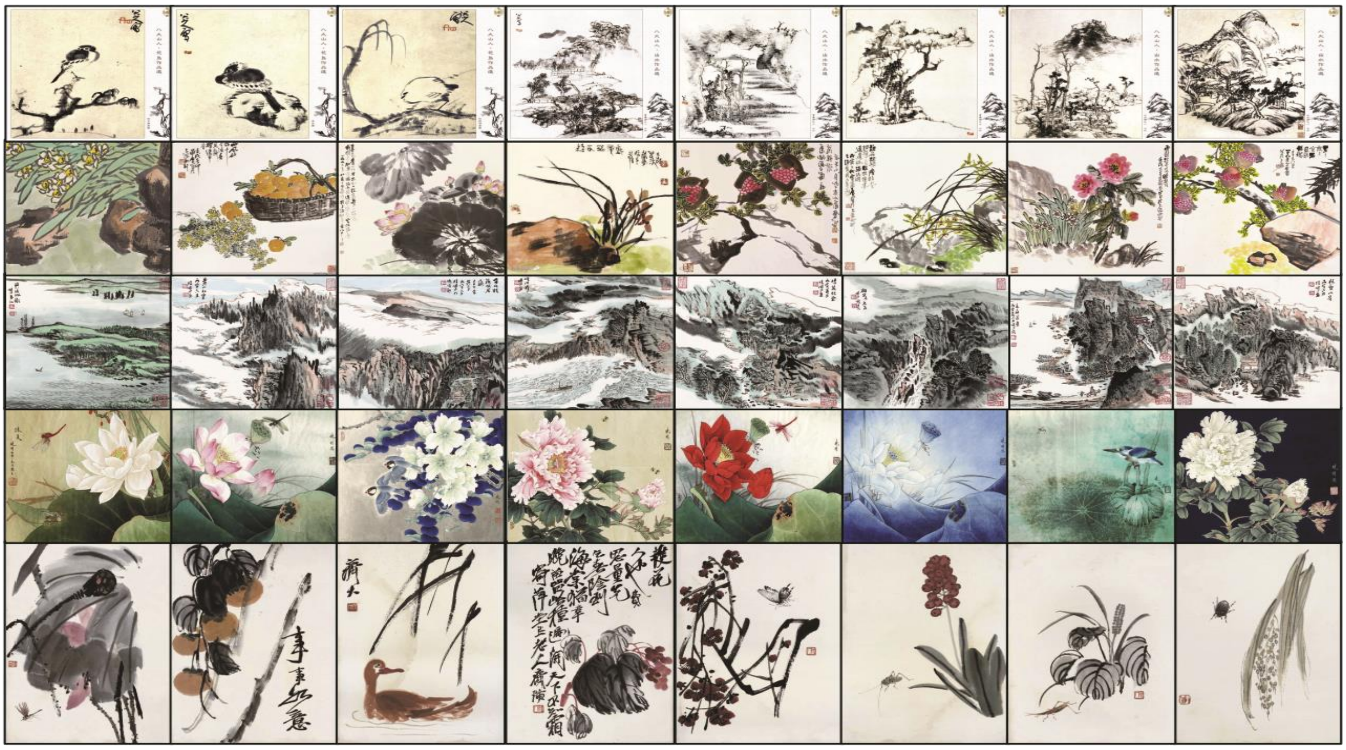

Figure 3: Illustration of five artists sample paintings that are randomly selected from the Dataset. Each row represents one of the five artists, including ZHU Da, CAO Jianlou, LU Yanshao, LI Xiaoming, and QI Baishi.

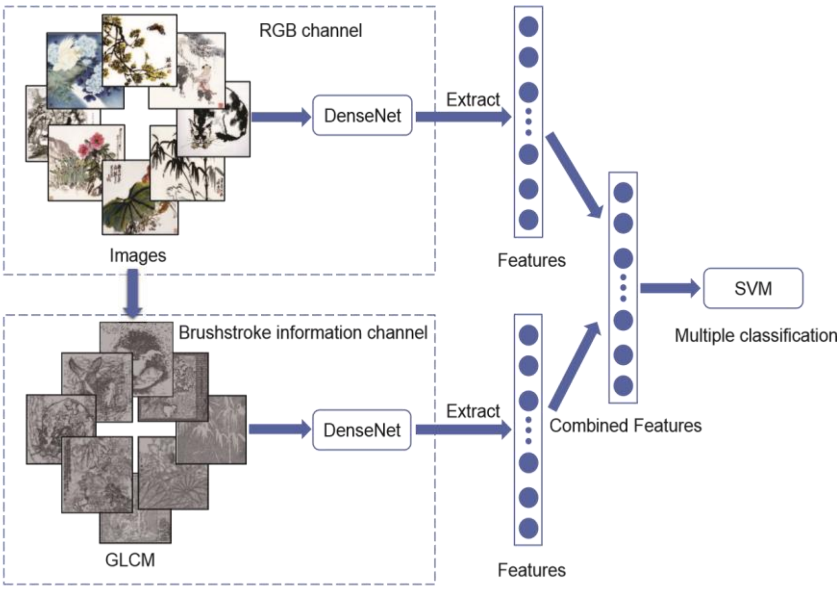

Figure 4: The structure of the proposed model.

4 Experiments

4.1 Dataset

In deep learning age, training data exerts an effect on the performance of 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.

4.2 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. 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) [35] 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 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 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. Fig.4 gives the overall structure of our proposed model.

4.3 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 method such as Saleh et al. [31], Tan et al. [40], Huang [41], Qian [42] and Sheng [43], 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 comparison 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.

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.

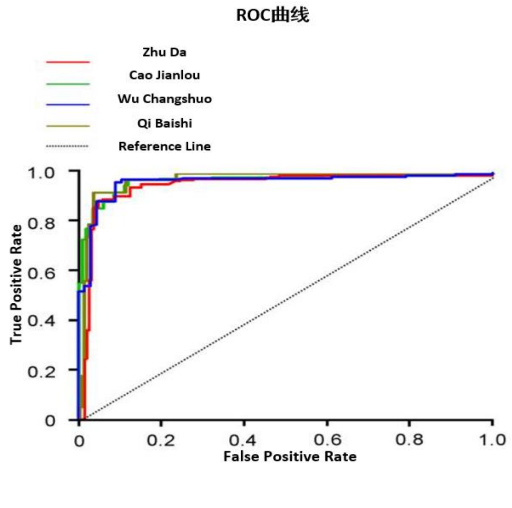

Figure 5：ROC curves for four painters of MTFFNet

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

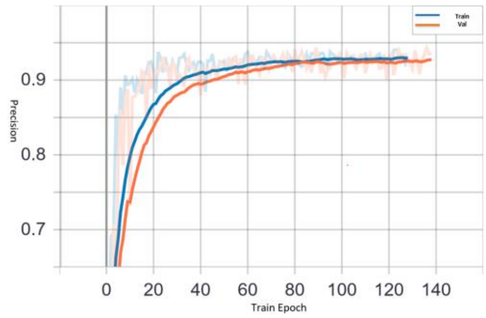

Figure 6: Training process of MTFFNet

Table 1: Comparison the classification accuracy of different methods.

Models	Accuracy
Saleh et al.	63.06
Tan et al. [40]	76.11
Huang X [41]	81.87
Qian W H.	82.15
Sheng J C. [43]	83.32
Sparse group	82.63
Decision Tree	65.52
SVM	74.17
MTFFNet	94.93

4.4 Ablation study

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

4.4.1 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 show 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.

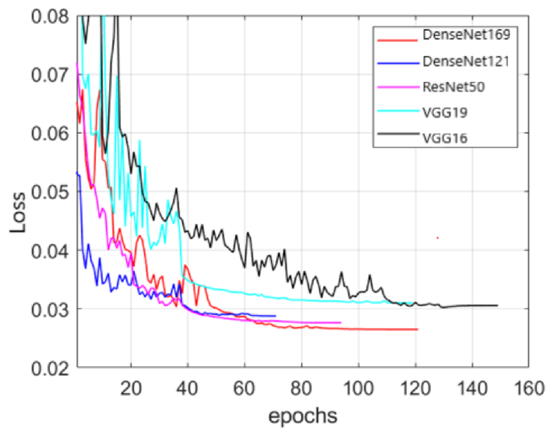

Figure 7: Training loss of MTFFNet with different backbones

Figure 8 gives the confusion matrix of different models on our dataset. It is shown that DenseNet get 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 demonstrate that DenseNet is capable of learning much more discriminative features than other networks, especially when dealing with images with similar appearance.

4.4.2 Analysis of the classifiers

An important question is whether we need the brush stroke input for Chinese paintings classification, and whether the SVM is better than SoftMax on that task. Table 3 include 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 2: Comparison of different networks on precision, recall, F1-score and Macro accuracy in the task of Chinese paintings 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
F1-score	0.92	0.90	0.91	0.93	0.94
Macro-accuracy	0.93	0.92	0.92	0.94	0.95

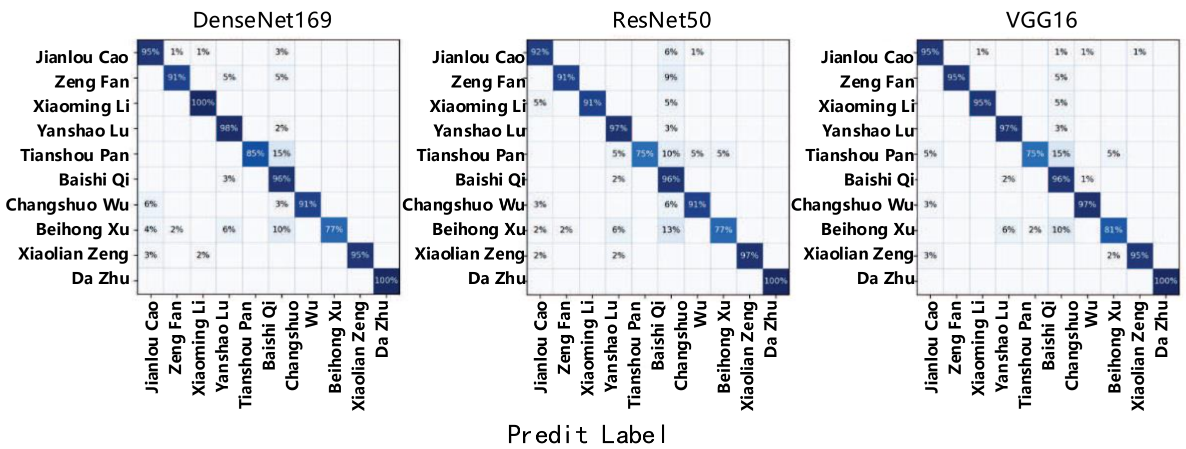

Figure 8: From left to right are visual representations of confusion matrix used by DenseNet169, ResNet50, and VGG16 as backbone for classification of Chinese paintings.

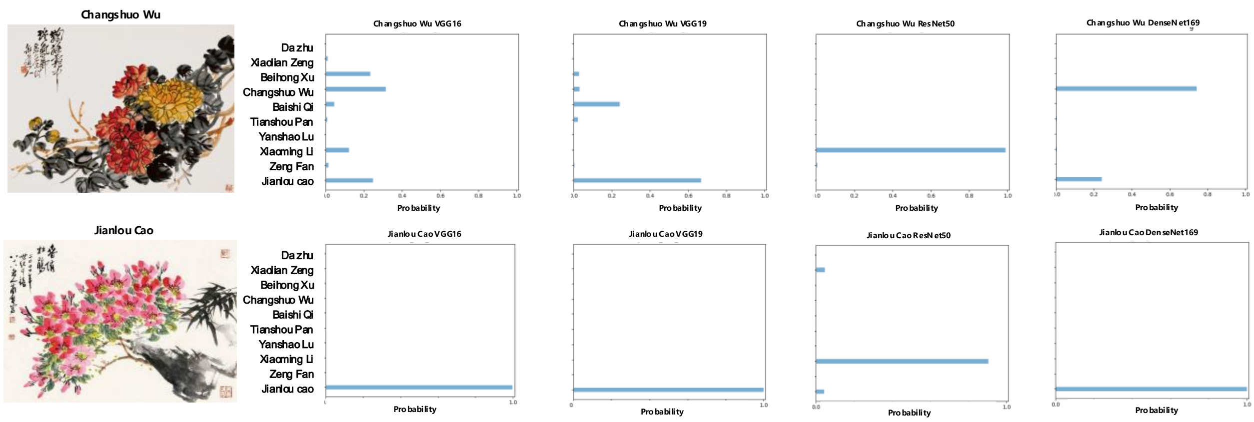

Figure 9: Network prediction histogram

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 paintings classification.

Top1 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:

i) 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, two-task VGG16 with SVM obtains 94% accuracy, which is just 1% less than DenseNet with SVM.

ii) Under the same experimental settings, both DenseNet169 with SVM and DenseNet169 with SoftMax yielded the best performance compared to other 260 networks, which further demonstrate that the underlying features learned by DenseNet is much more powerful and discriminative.

iii) As for the influence of brush stroke information on Chinese paintings 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 RGB image. The integration of brush stroke learning help improves the performance. The overall accuracy of the multi-task network is higher than 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%.

5 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, namely MTFFNet, for fine-art painting classification. By combination of two branches in the network, we can extract and fuse both the RGB features and the brush stroke information, where the Gray-Level Co-occurrence Matrix 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 others and ablation studies analysis have fully validated the efficacy of our proposed approach.

Although the network 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.

6 Acknowledgment

The authors wish to acknowledge the support for the research work from the National Natural Science Foundation, China under grant Nos. 61772360 and 61876125.

References

[1] G. Barequet, S. Har-Peled, Efficiently approximating the minimum-volume bounding box of a point set in three dimensions, Journal of Algorithms 38 (1) (2001) 91–109.

[2] Sherlock: A Semi-automatic Framework for Quiz Generation Using a Hybrid Semantic Similarity Measure[J]. Cognitive Computation, 2015, 7(6):667-679.

[3] Wang Z, Lin Z. Optimal Feature Selection for Learning-Based Algorithms for Sentiment Classification[J]. Cognitive Computation, 2020, 12(1):238-248.

[4] Xu X, Li Y, Wu Q M J. A Multiscale Hierarchical Threshold-Based Completed Local Entropy Binary Pattern for Texture Classification[J]. Cognitive Computation, 2019(8):1-14.

[5] Huang G, Liu Z, Maaten L V D, et al. Densely Connected Convolutional Networks[J]. Computer Era, 2017.

[6] Zhong G, Jiao W, Gao W, et al. Automatic Design of Deep Networks with Neural Blocks[J]. Cognitive Computation, 2019, 12(8).

[7] Zhang D, Han J, Zhao L, et al. Leveraging Prior-Knowledge for Weakly Supervised Object Detection Under a Collaborative Self-Paced Curriculum Learning Framework[J]. International Journal of Computer Vision, 2019(1).

[8] Manikandan J, Venkataramani B. Study and evaluation of a multi-class SVM classifier using diminishing learning technique[J]. Neurocomputing, 2010, 73(10-12):1676-1685.

[9] J. Ren, ANN vs. SVM: Which one performs better in classification of MCCs in mammogram imaging, Knowledge-Based Systems, Volume 26,2012, Pages 144-153,ISSN0950-7051, https://doi.org/10.1016/j. knosys. 2011.07.016.

[10] L. Jia, J. Z. Wang, Studying digital imagery of ancient paintings by mixtures of stochastic models, IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society 13 (3) (2004) 340–53.

[11] I. E. Berezhnoy, E. O. Postma, J. V. D. Herik, Computerized visual analysis of paintings.

[12] S. Jiang, Q. Huang, Q. Ye, G. Wen, An effective method to detect and categorize digitized traditional chinese paintings, Pattern Recognition Letters 27 (7) (2006) 734–746.

[13] C. R. Johnson, E. Hendriks, I. J. Berezhnoy, E. Brevdo, S. M. Hughes, I. Daubechies, J. Li, E. Postma, J. Z. Wang, Image processing for artist identification, Signal Processing Magazine IEEE 25 (4) (2008) 37–48.

[14] J. Shen, Stochastic modeling western paintings for effective classification, Pattern Recognition 42 (2) (2009) 293–301.

[15] L. Jia, Y. Lei, H. Ella, J. Z. Wang, Rhythmic brushstrokes distinguish van gogh from his contemporaries: findings via automated brushstroke extraction, IEEE Transactions on Pattern Analysis and Machine Intelligence 34 (6) (2012) 1159-76.

[16] D. C. Cirean, U. Meier, L. M. Gambardella, J. Schmidhuber, Deep, big, simple neural nets for handwritten digit recognition, Neural Computation 22 (12) (2010) 3207–3220.

[17] X. X. Niu, C. Y. Suen, A novel hybrid cnnsvm classifier for recognizing 325 handwritten digits, Pattern Recognition 45 (4) (2012) 1318–1325.

[18] J. Shuiwang, Y. Ming, Y. Kai, 3d convolutional neural networks for human action recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 35 (1) (2013) 221–231.

[19] Padfield N, Zabalza J, Zhao H, et al. EEG-Based Brain-Computer Interfaces Using Motor-Imagery: Techniques and Challenges[J]. Sensors, 2019, 19(6).

[20] O. Abdel-Hamid, A. R. Mohamed, H. Jiang, L. Deng, G. Penn, D. Yu, Convolutional neural networks for speech recognition, IEEE/ACM Transactions on Audio Speech and Language Processing 22 (10) (2014) 1533– 1545.

[21] Han J, Cheng G, Li Z, et al. A Unified Metric Learning-Based Framework for Co-saliency Detection[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2017:1-1.

[22] J. Han, H. Chen, N. Liu, C. Yan, X. Li, Cnns-based rgb-d saliency detection via cross-view transfer and multiview fusion, IEEE Transactions on Cybernetics 48 (11) (2018) 3171.

[23] Y. Yan, J Ren et al, Unsupervised image saliency detection with Gestalt-laws guided optimization and visual attention based refinement[J]. PATTERN RECOGNITION, 2018.

[24] Gong C, Junwei H, Peicheng Z, et al. Learning Rotation-Invariant and Fisher Discriminative Convolutional Neural Networks for Object Detection[J]. IEEE Transactions on Image Processing, 2018, 28:265-278.

[25] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks,

in: International Conference on Neural Information Processing Systems, 2012.

[26] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, F. F. Li, Imagenet: a largescale hierarchical image database, in: IEEE Conference on Computer Vision and Pattern Recognition, 2009.

[27] Sergey Karayev, Matthew Trentacoste, Helen Han, Aseem Agarwala, Trevor Darrell, Aaron Hertzmann, and Holger Winnemoeller. Recognizing Image Style. Proceedings of the British Machine Vision Conference. BMVA Press, September 2014.

[28] Paul A, Malathy C. An Innovative Approach for Automatic Genre-Based Fine Art Painting Classification[M]// Advanced Computational and Communication Paradigms. 2018.

[29] Bar Y, Levy N, Wolf L. Classification of Artistic Styles Using Binarized Features Derived from a Deep Neural Network[J]. 2014.

[30] J. Zabalza et al., "Novel Two-Dimensional Singular Spectrum Analysis for Effective Feature Extraction and Data Classification in Hyperspectral Imaging," in IEEE Transactions on Geoscience and Remote Sensing, vol. 53, no. 8, pp. 4418-4433, Aug. 2015, doi: 10.1109/TGRS.2015.2398468.

[31] B. Saleh, A. Elgammal, Large-scale classification of fine-art paintings: Learning the right metric on the right feature, Journal of Computer Science and Technology 25 (3) (2015) 595–605.

[32] J. Zabalza, J. Ren, M. Yang, Y. Zhang, J. Wang, S. Marshall, J. Han, Novel Folded-PCA for improved feature extraction and data reduction with hyperspectral imaging and SAR in remote sensing, ISPRS Journal of Photogrammetry and Remote Sensing, Volume 93, 2014, Pages 112-122.

[33] Yitian Xu, Rui Guo, Laisheng Wang. A Twin Multi-Class Classification Support Vector Machine[J]. Cognitive Computation, 2013.

[34] Xiangping H U. Multiple feature fusion via multiple kernel learning for image classification[J]. Computer Engineering and Applications, 2016.

[35] Sun G, Rong X, Zhang A, et al. Multi-Scale Mahalanobis Kernel-Based Support Vector Machine for Classification of High-Resolution Remote Sensing Images[J]. Cognitive Computation, 2019.

[36] Bian X, Zhang T, Zhang X, et al. Clustering-Based Extraction of Near Border Data Samples for Remote Sensing Image Classification[J]. Cognitive Computation, 2012.

[37] J. Tschannerl, J. Ren, P. Yuen, G. Sun, H. Zhao, Z. Yang, Z. Wang, S. Marshall, MIMR-DGSA: Unsupervised hyperspectral band selection based on information theory and a modified discrete gravitational search algorithm, Information Fusion, Volume 51, Pages 189-200, 2019.

[38] M. Borchani, G. Stamon, Texture features for image classification and retrieval, in: Multimedia Storage and Archiving Systems II, 1997.

[39] C. Hentschel, T. P. Wiradarma, H. Sack, Fine tuning cnns with scarce training data adapting imagenet to art epoch classification, in: IEEE International Conference on Image Processing, 2016.

[40] W. R. Tan, C. S. Chan, H. E. Aguirre, K. Tanaka, Ceci nest pas une pipe: A deep convolutional network for fine-art paintings classification, in: IEEE International Conference on Image Processing, 2016.

[41] X. Huang, S. H. Zhong, Z. Xiao, Fine-art painting classification via two channel deep residual network, in: Pacific Rim Conference on Multimedia, 2017.

[42] Qian W, Xu D, Guan Z, et al. Simulating Chalk Art Style Painting[J]. International Journal of Pattern Recognition and Artificial Intelligence, 2017, 31(12):1759026.1-1759026.20.

[43] Jia-Chuan S. Automatic Categorization of Traditional Chinese Paintings Based on Wavelet Transform[J]. Computer Science, 2014, 41(2):317-289.