Classification of style in fine-art paintings using transfer learning and weighted image patches

Catherine Sandoval Rodriguez, Margaret Lech and Elena Pirogova
School of Engineering
RMIT University
Melborune, Australia
s3465033@student.rmit.edu.au, {margaret.lech, elena.pirogova}@rmit.edu.au

***Abstract*—With the ongoing expansion of digitized artworks, the automated analysis and categorization of fine art paintings have become a rapidly growing research field. However, due to the implicit subjectivity and nuances separating different artistic movements, numerical art analysis implies significant challenges. This paper describes a new efficient method that improves the classification accuracy of fine-art paintings compared to the existing baseline methods. The proposed approach is based on transfer learning and classification of sub-regions or patches of the painting. A weighted sum of the individual-patch classification outcomes gives the final stylistic label of the analyzed painting. The patch weights are optimized to maximize the overall style recognition accuracy. Experimental validation based on two standard art classification datasets and six different pre-trained convolutional neural network (CNN) models (AlexNet, VGG-16, VGG-19, GoogLeNet, ResNet-50 and Inceptionv3) shows that the proposed approach outperforms the baseline techniques and offers low computational and data costs.**

Keywords—Fine-art style classification, Convolutional neural networks, Transfer learning, Image analysis, Image classification

I. Introduction

In recent years, the analysis and classification of fine art paintings have been steadily gaining more researcher attention due to the continuous expansion of digitized artworks. Libraries, museums, galleries, art centers and art organizations have digitized their collections to encourage public interest in art and to facilitate access to exclusive artworks and masterpieces. As the Internet fades geographical barriers, online galleries are promoting exhibitions of paintings and conduct marketing of artworks. These activities create a demand for automatic indexation, identification and classification of digitized art.

In art, a style is usually defined according to the artistic movement [1]. Painting style classification is typically carried out by art historians and curators based on the relation between physical characteristics (light, line, color, texture, shape, space, etc.), subjective attributes, and historical periods [2]. However, in many cases significant stylistic variations can be observed, such as smooth transition between artistic movements over time, style differences between paintings made by the same author, unique personal characteristics that do not belong to any style or an artistic period, artworks with elements that belong to multiple styles, influence of one artist on others, diverse interpretation of abstract and sub-real elements [3]. Therefore, the process of labeling a painting with a unique style is a challenging task, even for an expert. Different approaches have been applied with the aim to automatically classify paintings according to their artistic movement. Classical approaches have addressed the task of style classification based on the selection or identification of the appropriated features and classification algorithms [4-6]. In more modern deep learning (DL) approaches, artistic style recognition is achieved by training convolutional neural network (CNN) models, where relevant features are automatically extracted from images by the network itself [8-17]. The vast majority of style classification studies based on DL techniques apply the concept of transfer learning from the domain of object classification to the domain of painting style categorization. Within this scheme, a CNN model is pre-trained on a very large dataset of natural images, and only short training (fine-tuning) on a new relatively small image dataset representing different artistic styles is needed to adapt the network to a new classification task. The advantage is that lengthy and data-costly fresh training can be avoided without a significant loss of classification accuracy.

The current study describes and validates a new fine art classification method that uses transfer learning and optimized weighted image-patch classification. The proposed approach is tested on two standard image classification databases. A comparison with the existing state-of-the-art techniques is presented.

The remainder of this paper is organized as follows: Section II presents previous relevant research works. The proposed methodology is described in Section III. Section IV describes the experiments and results. Finally, Section V gives the conclusion.

II. Related work

A. Classical approaches

Early studies addressed the problem of painting classification using classical machine learning techniques that were based on a large set of local features, which were extracted from image datasets that contained a relatively small number of artworks. Shamir et al. [4] classified 3 artistic movements over 513 images by performing several image transforms and a weighted nearest neighbor technique. Arora and Elgammal [5] proposed the style classification of 7 artistic movements with a labeled training dataset of 490 paintings by means of local level features extraction techniques. These techniques included the color scale-invariant feature transform (CSIFT) and the opponent-SIFT (OSIFT) algorithm. The features were classified using the support vector machine (SVM) algorithm. The same SVM classifier was implemented by Khan et al. [6] to categorize multiple types of global and local features such as: the local binary patterns (LBP), color LBP, GIST, pyramid of histograms of orientation gradients (PHOG), scale-invariant feature transform (SIFT) and the histogram of oriented gradients (HOG) parameters. It has been demonstrated that

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

978-1-5386-5602-0/18/$31.00 ©2018 IEEE

using only low-level features was insufficient to efficiently distinguish between different stylistic classes.

B. Deep learning approaches

Recent developments have addressed the problem of painting classification through transfer learning, which was implemented either via feature extraction or fine-tuning of different CNN architectures that were pre-trained using the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) dataset of more than 1.2 million natural images of objects with 1,000 categories [7]. The first large-scale study was carried out by Karayev et al. [8] and created a publicly available collection of digital fine art paintings to recognize a larger variety of visual styles. A linear SVM classifier was trained with features generated by the fine-tuned CNN model. In a similar way, CNN models were implemented as feature extractors and linear classifiers in [9-11]. Tan et al. [12] presented a large-scale style classification of fine art paintings by fine-tuning of a pre-trained CNN model. It was found that CNN fine-tuned models outperform the freshly trained CNN models, as well as methods using linear classifiers to categorize features extracted from CNNs. Sun et al. [13] proposed a CNN architecture with two pathways that extracted object features and texture features to address the style classification problem. The object pathway was given by a standard CNN architecture, and the texture pathway was calculated by the gram-matrices of intermediate features in the object pathway. Florea et al. [14] compared the use of the color, structure and topographic descriptors, and adapted boosted SVMs against CNN fine-tuned models. One of the most recent painting classification studies by Elgammal et al. [15] implemented style classification by transfer learning and fine-tuning of three CNN models. In addition to style categorization, interpretability of learned representations of the CNN and chronology of the paintings was examined through dimensionality reductions methods such as principal component analysis (PCA).

C. Using image patches or sections

Digitized image paintings have different sizes which are typically larger than the fixed input image size required by pre-trained CNN models. Therefore, it is necessary to resize the images to the required dimensions before passing them as inputs to CNNs. However, in the context of art, the image re-sizing process can introduce significant distortions, and loss of important details such as texture, brush strokes, or other characteristics that may be essential for automatic fine art analysis [15]. To address this issue, it was proposed to use image patches (sub-images) extracted from the original image [16] [17]. A deep multibranch neural network was applied by Bianco et al. [16] for the task of style recognition using a dataset of 2338 paintings with 13 different styles. Three branches processed the input image split into three random patches taken at two different resolutions: the CNN input size and two times the CNN input size. Painting analysis with patches was explored by Folego et al. [17]. A Binary classification of Vincent van Gogh's paintings, based on a small dataset of 333 paintings, was performed by dividing each painting into non-overlapping patches and classifying each patch individually to determine the final decision by the patch with the highest probability score. The study used a CNN model as a feature extractor and the SVM classifier.

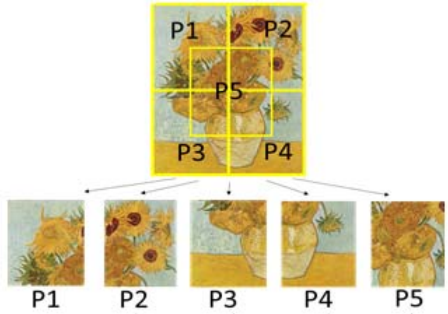

Fig. 1. Generation of image patches P1-P5.

III. METHOD

The current study introduces a new approach to the automatic categorization of fine-art paintings. In general terms, the approach divides each input image into five sub-images (patches) that have fixed locations within the image array. The patches are used to fine-tune a pre-trained CNN structure modified in a way that gives classification outcomes separately for each of the five types of patches. The inference process derives the final classification result for a given input image using a weighted sum of the five CNN outcomes, with each outcome given for a different type of patch. The weights, denoting the importance of each patch to the final decision, are derived using a numerical optimization procedure.

To give more details, the proposed method consisted of three main steps. In the first step, the original images were re-sized to the double of the input size required by the CNN model. From each image, five patches were extracted. As shown in Fig. 1, the first four patches represented image corners (upper right P1, upper left P2, lower right P3, lower left P4), and the fifth patch was the central patch (P5). There was 25% of overlap between the central patch and each of the corner patches. Thus, from each re-sized image, five sub-images were generated and used as the inputs to the CNN.

In the second step, a pre-trained CNN model was fine-tuned on the patch-image data. To fine-tune the CNN, the three outer layers, fully connected layer, softmax and the classification output layer were removed from the pre-trained CNN, and replaced by their new modified versions. The size of the new fully-connected layer was determined by the number of classes (i.e. number of different artistic styles) of the input data. The softmax layer was designed to return a matrix of probabilities of each class, while the classification layer was made to assign to each of the five input patches one of the mutually exclusive stylistic classes. Given this modification to the standard CNN structure, the network had multiple outputs that allowed it to give separate classification outcomes C1-C5 for each type of patch (see Figs. 2 and 3). In other words, the five patches were classified independently by the same interconnected network structure.

In the third step, the final stylistic class of the input painting was determined by calculating a weighted linear combination of the individual-patch classification outcomes $C_{n,k}$ (n=1,…,5) using (1).

$$CT_k = \sum_{n=1}^{5} w_n C_{n,k} \qquad (1)$$

978-1-5386-5602-0/18/$31.00 ©2018 IEEE

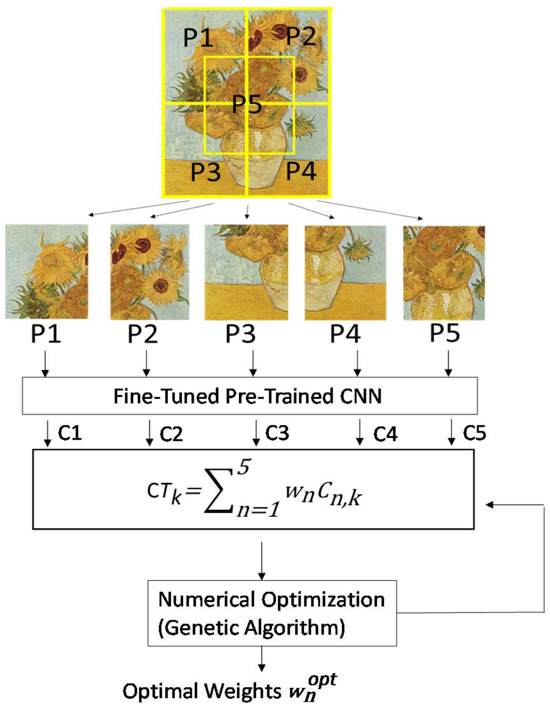

Fig. 2. The optimization process to derive the optimal weight values for each patch using outputs from a fine-tuned CNN. The CNN was designed and trained to provide separate style recognition outcomes for each of the five patches.

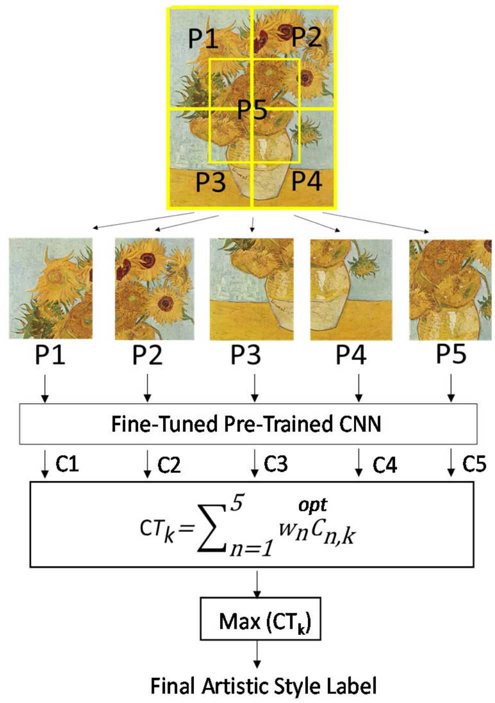

Fig. 3. The inference (classification) process to derive the final prediction, based on the weighted linear combination of classification outcomes for each of the five patches. The optimal weight values for each patch were determined using an optimization procedure (GA).

Where CT_k denotes the total probability vector of the k-th image, n is the patch type number, and w_n are the weights for each of the five patches. As shown in Fig. 2, the optimal weight values for each patch were derived using an optimization procedure. It applied a genetic algorithm (GA) search which was iteratively changing the weights until the algorithm converged to the maximum value of the classification accuracy. The classification accuracy was calculated as the quotient between the number of correct predictions and the number of total predictions. The constraints given in (2) and (3) were defined for the optimization procedure to ensure that none of the patch weights exceeded the value of 1, and the sum of all weights was equal to 1.

$$0 \le w_n \le 1 \tag{2}$$

$$\sum_{n=1}^{M} w_n = 1 \tag{3}$$

Given the optimal weight values, the inference procedure was conducted as shown in Fig. 3. The final stylistic class label for the k-th image was given as the class corresponding to the maximum probability value of the vector CT_k.

IV. Experiments and Results

A. Datasets

The proposed method was validated experimentally using two datasets of fine art images. The first dataset (dataset 1) was composed of images representing the Australian Aboriginal style, as well as images representing five styles from the commonly used Wikiart dataset [18]. Only classes containing sufficient numbers of representative images to match the number of Aboriginal-style images were selected from the Wikiart dataset. These styles included: Expressionism, Impressionism, Post Impressionism, Realism and Romanticism. While the Australian Aboriginal art pictures were professionally labeled, the Wikiart collection had to undergo a manual inspection and revision to ensure labeling correctness. This was due to the fact that, the Wikiart images were labeled by general public volunteers, and some pictures were either not depicting fine art paintings or had a very poor quality. These cases were discarded. In the final dataset 1 used in the experiments described here, each one of the six stylistic classes was represented by 5145 images, giving a total of 30870 images.

The second dataset (dataset 2) included the original publicly available Pandora 18K dataset [19], which contained pictures of paintings representing 18 different styles, and the Australian Aboriginal style paintings added to dataset 2 by the authors. The Paintings Dataset for Recognizing the Art movement (Pandora 18K) was created by Florea et al [14, 20], and provided a very high-quality dataset that was subjected to a rigorous style-labeling process done by art experts. Dataset 2 included in total 19320 images representing 19 different artistic styles. Fig. 4 shows the percentage distribution of different artistic styles in dataset 2.

B. Experimental setup

In the style classification experiments, a three-fold cross validation scheme was adopted. The reported results are given as the average classification accuracy over the three folds. In the weight optimization procedure based on the GA algorithm, a five-fold cross validation scheme was adopted. In both cases, 80% of the data was selected to train the CNN models, and the remaining 20% was used to test the system performance.

978-1-5386-5602-0/18/$31.00 ©2018 IEEE

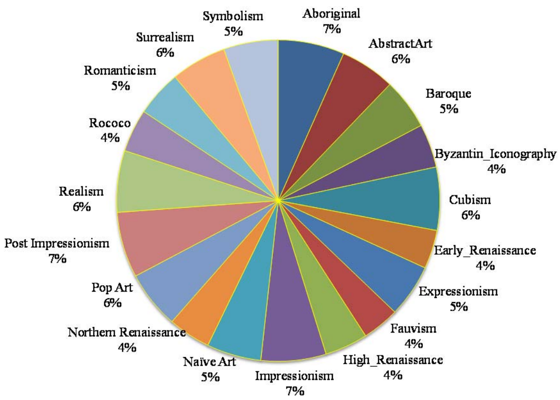

Fig. 4. Dataset 2 - percentage distribution by style (Pandora18K dataset plus Australian Aboriginal art style).

The proposed method was evaluated using six alternative pre-trained CNN models: AlexNet [21], VGG-16 [22], VGG-19 [22], GoogLeNet [23], ResNet-50 [24] and Inceptionv3 [25]. All of these networks were pre-trained on a very large number of images from the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) dataset. Table I shows the input image size required by each model, the number of layers, and the names of the three outer layers that were replaced to fine tune each of the tested CNN models.

C. *Results and discussion*

For each dataset, and for each of the six pre-trained CNN models, a four-step experiment was conducted.

Step 1 (*Baseline*). A baseline test was performed using whole (uncropped) original images as inputs to the CNN. This implies that each whole image had to be re-sized to fit into the dimensions required by each of the six different CNNs. Every CNN was fine-tuned with the original number of images included in the training dataset. The average classification accuracy is shown in Figs. 5 and 6 (blue bars).

Step 2 (*5-patches*). In this test, the classification was performed separately for each of the 5 types of image patches. The patches were generated from whole images that had been re-sized to twice the CNN input, so each patch had the same size as the CNN input. No connection between patches and the original images was maintained; it was a picture-independent style classification task. Since each of the original images generated 5 patches, the CNNs were trained using five times as many images than in Step 1. The final accuracy (red bars in Figs. 5 and 6) was calculated as the average of individual classification results over all patches.

Step 3 (*Average*). In the third test, the class probability values obtained for each patch in Step 2 were used to determine the average classification value for each image of the dataset. This was done by grouping the values obtained for the respective five patches for a given image (picture dependent) task and calculating the average classification accuracy across all images (green bars in Figs. 5 and 6).

Step 4 (*Optimal weights*). Finally, in the fourth test the proposed approach was tested by repeating the above Steps 2-3 but in Step 3, instead of a simple unweighted average, a weighted average was used (purple bars in Figs. 5 and 6).

TABLE I. CNN MODEL SPECIFICATIONS

CNN Model	CNN model characteristics		
	Input image size	**Number of Layers**	**Replaced layers (fine tuning)**
AlexNet	227x227	8	fc8, prob, output
VGG-16	224x224	16	fc8 ,prob, output
VGG-19	224x224	19	fc8 ,prob ,output
GoogLeNet	224x224	22	loss3-classifier, prob, output
ResNet-50	224x224	50	fc1000,fc1000_softmax ,ClassificationLayer_fc1000'
Inceptionv3	299x299	48	predictions, predictions_softmax, ClassificationLayer_predictions

The results indicate that the proposed approach (Step 4) showed the highest accuracy, outperforming all other cases (Steps 1-3). The average accuracies obtained with the patch weight optimization method (Step 4) were, depending on the CNN model, 1% - 1.3% (dataset 1) and 1.1% - 1.6% (dataset 2) higher than the simple average patch aggregation method (Step 3), and between 1.9% - 3.2% better than the baseline (Step 1).

The lowest accuracy was in all cases obtained when the patches were classified individually (Step 2) with either dataset, regardless of what CNN model was used. Accuracies obtained with the patch weight optimization method (Step 4) were, depending on the CNN model, 4.3% - 5.9% (dataset 1) and 4.7% - 7.5% (dataset 2) higher than for the individual patch accuracy (Step 2). When comparing Step 2 with the baseline test (Step 1), the accuracy decreased by 2.1% - 4.9% with dataset 2 and by 1.4% and 3.4% with dataset 1.

In summary, the patches provided higher resolution but only partial content of an image. Therefore, style classification based on individual patches (Step 2) led to lower results than the baseline classification of lower-resolution but full-content images (Step 1). When the patches were considered in relation to the whole image (Step 3) the classification results were better than in Step 1, indicating that the higher resolution of the patches added important stylistic information. The results were further improved in Step 4 by assigning weights to each patch, which implied or indicated their importance to the final classification decision.

The results presented in Figs. 5 and 6 show that the classification accuracy improves with the increasing complexity and training data size of the CNN network. Thus, the highest accuracies were obtained with the pre-trained CNN InceptionV3 model and the lowest with the AlexNet model. This was true for both datasets. The InceptionV3 network outperformed the AlexNet by 5.3% for the dataset 1, and by 5.7% for the dataset 2. The accuracies obtained with both VGG models were very similar, and only represented around 1% of improvement compared to the AlexNet model. The tests with the GoogLeNet model improved by 2% - 3%, and the tests for Resnet-50 by 3.8% - 5.4% compared to AlexNet. Therefore, it can be observed that by increasing the complexity, and depth of the CNN models from AlexNet, with a linear architecture of 8 layers, to InceptionV3, with 48 layers, the average classification accuracy was improved by about 5.5%.

978-1-5386-5602-0/18/$31.00 ©2018 IEEE

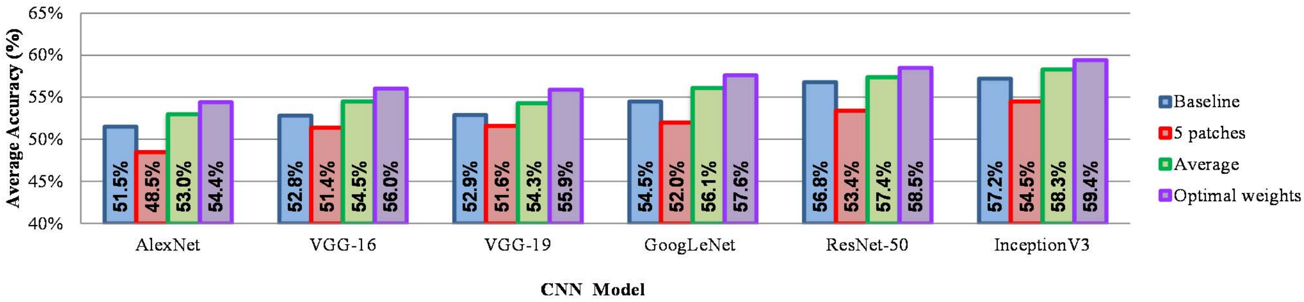

Fig. 5. Style classification results based on dataset 1.

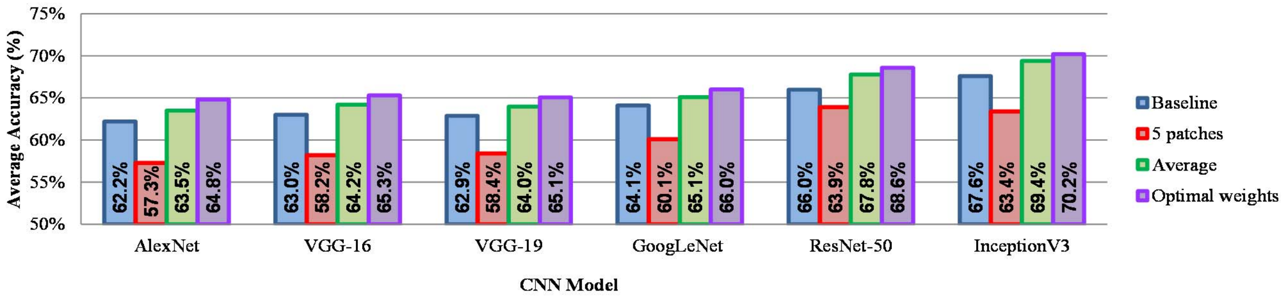

Fig. 6. Style classification results based on dataset 2.

However, the increase of classification accuracy was associated with increasing computational cost. It was estimated that the processing time was increased five times when changing from AlexNet to InceptionV3. One of the most recent research studies [15] suggested that the ResNet-50 is the best performing CNN model for painting style classification. However, the test results presented in this study show that InceptionV3 can provide a further improvement of classification accuracy. Namely, the current study shows that the classification accuracy given by InceptionV3 outperforms ResNet-50 by about 0.4% - 0.9% (dataset 1) and 0.5 - 1.7% (dataset 2).

Studies based on the Pandora 18K dataset [14] have reported an average accuracy of 50.1% using a large set of visual descriptors and the SVM as a classifier, which is a significantly lower result compared to the results given by the proposed method, as well as the presented CNN baseline techniques.

Fig. 7 shows an example of the confusion matrix for style classification based on the InceptionV3 CNN model, optimized weights and dataset 1. The Australian Aboriginal style provided the highest accuracy, while the Post Impressionism achieved the lowest value. The most significant style confusions were observed between the Post Impressionism and Impressionism artistic movements. The Expressionism style also had a noticeably high error rate.

Fig. 8 shows an example of the confusion matrix for the style classification based on the InceptionV3 CNN model and individual classification of the patches (Step 2). Comparing Fig. 7 and Fig. 8, the classification accuracy per style increased due to the application of the optimized weights. The weights act as correction factors, improving system performance. It can be also observed that the accuracy for the Aboriginal style showed only a small increase of 1% with the application of weights, while the accuracy for the Impressionism was significantly increased by 8%.

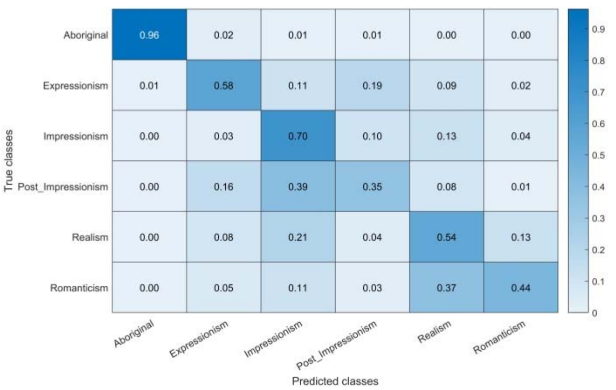

Fig. 7. Confusion matrix of style classification results based on dataset 1 using Inceptionv3 model and weighted patches.

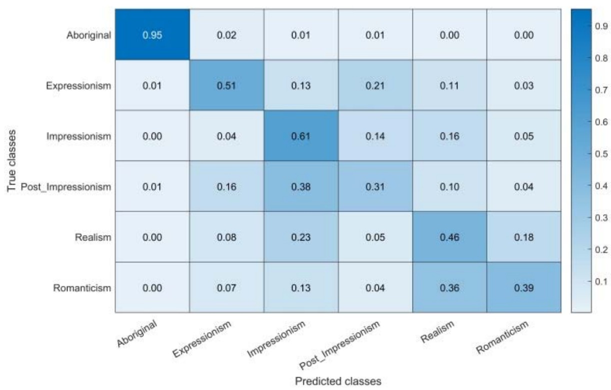

Fig. 8. Confusion matrix of style classification results based on dataset 1 using Inceptionv3 model and patches individual classification.

978-1-5386-5602-0/18/$31.00 ©2018 IEEE

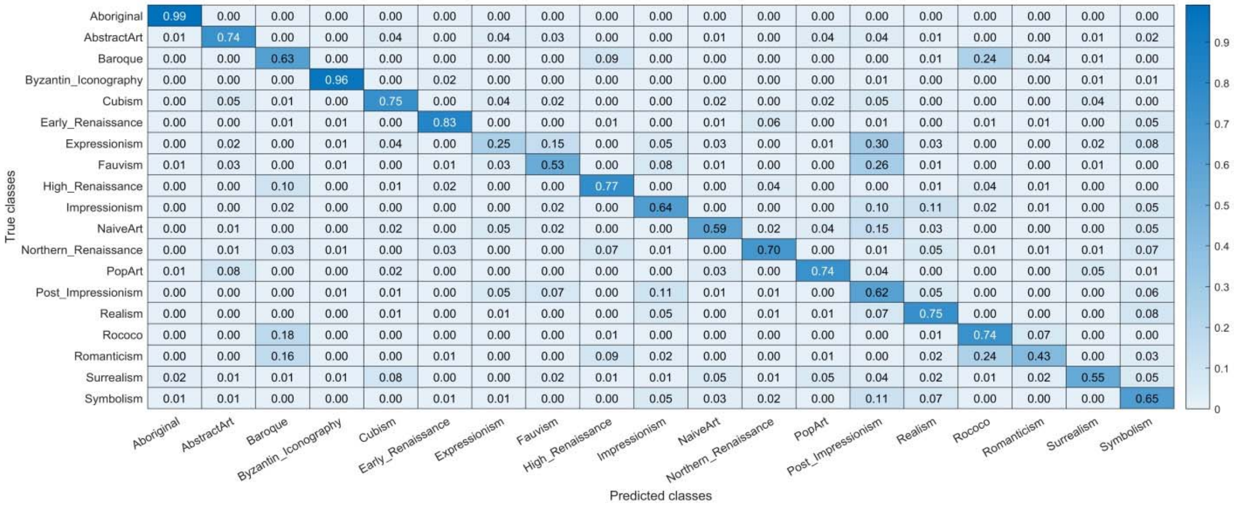

Fig. 9. Confusion matrix of style classification based on dataset 2 using Inceptionv3 model and weighted patches.

The confusion matrix for the classification test based on the dataset 2 using InceptionV3 model and optimized weights is shown in Fig. 9. The Australian Aboriginal style achieved the highest accuracy (above 95%) in all tests. The Byzantine Iconography style was also recognized with a very high accuracy of around 90%. These styles are noticeably different from other artistic movements. Expressionism style, on the other hand, provided the lowest accuracy. Both the Expressionism and Post Impressionism styles presented the highest error rates. Baroque and Rococo indicated a significant confusion, which could be attributed to the fact that these styles are historically related.

V. CONCLUSIONS

The present study investigated the application of a new fine art classification technique using transfer learning and weighted image patches. The classification results indicated that the proposed method provides a computationally efficient way of improving the fine-art style classification accuracy without the need to define and fully train new CNN model structures, or to increase the sizes of existing databases.

Empirical tests demonstrated that regardless of the type of the CNN model, the proposed method outperforms the baseline approach. Classification of image patches performs better than the classification of the complete images when the relation to the original source-image is kept by either calculating the unweighted or weighted average accuracy across all patches. In contrast, when image patches are classified individually as independent from the original source-image, the classification accuracy decreases.

The pre-trained CNN network Inceptionv3 model was found to provide the best results within the proposed classification paradigm; however, it is also the most demanding model in terms of computational costs.

Classification of images using weighted image patches could be used in applications where the label of the image is kept over all sub-regions. Therefore, the categorization of emotions, quality, aesthetic in artworks and photographs could be explored as future work.

ACKNOWLEDGMENT

This research is supported by an Australian Government Research Training Program Scholarship.

REFERENCES

[1] OxfordWords, style | Definition of style in English by Oxford Dictionaries, on: May 20, 2018. [Online]. Available: https://en.oxforddictionaries.com/definition/style.

[2] T. E. Lombardi, "The Classification of Style in Fine-Art Painting " Doctor of Professional Studies in Computing Ph.D. dissertation, School of Computer Science and Information Systems Pace University New York, 2005.

[3] The Art Story: Modern Art Movement Timeline, on: May 20, 2018. [Online].Available: http://www.theartstory.org/section_movements_timeline.htm.

[4] L. Shamir, T. Macura, N. Orlov, D. M. Eckley, and I. G. Goldberg, "Impressionism, expressionism, surrealism, Automated recognition of painters and schools of art," ACM Transactions on Applied Perception, vol. 7, no. 2, pp. 1-17, Feb. 2010. doi:10.1109/ICIP.2016.7532335.

[5] R. S. Arora and A. Elgammal, "Towards automated classification of fine-art painting style: A comparative study," in *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, Tsukuba, Nov. 2012, pp. 3541-3544.

[6] F. S. Khan, S. Beigpour, J. van de Weijer, and M. Felsberg, "Painting-91: a large scale database for computational painting categorization," Machine Vision and Applications, vol. 25, no. 6, pp. 1385-1397, Aug. 2014. doi:10.1007/s00138-014-0621-6.

[7] J. D. Olga Russakovsky, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, Li Fei-Fei, "ImageNet Large Scale Visual Recognition Challenge," International Journal of Computer Vision (IJCV), vol. 115, no. 3, pp. 211–252, Dec. 2015. doi:10.1007/s11263-015-0816-y.

[8] S. Karayev et al., "Recognizing image style," in *Proceedings of British Machine Vision Conference*, arXiv preprint arXiv:1311.3715, 2014.

[9] B. Saleh and A. Elgammal, "Large-scale Classification of Fine-Art Paintings: Learning The Right Metric on The Right Feature," arXiv preprint arXiv:1505.00855, 2015.

[10] C. Hentschel, T. P. Wiradarma, and H. Sack, "Fine tuning CNNS with scarce training data – Adapting imagenet to art epoch classification," in *2016 IEEE International Conference on Image Processing (ICIP)*, Phoenix, AZ, USA, Sep. 2016, pp. 3693-3697. doi: 10.1109/ICIP.2016.7533049

[11] W. T. Chu and Y. L. Wu, "Image Style Classification Based on Learnt Deep Correlation Features," IEEE Transactions on Multimedia, vol. 20, no. 9, pp. 2491-2502, Sep. 2018. doi:10.1109/TMM.2018.2801718.

[12] W. R. Tan, C. S. Chan, H. E. Aguirre, and K. Tanaka, "Ceci n'est pas une pipe: A deep convolutional network for fine-art paintings classification," in *2016 IEEE International Conference on Image Processing (ICIP)*, Phoenix, AZ, USA, Sep. 2016, pp. 3703-370. doi:10.1109/ICIP.2016.7533051.

[13] T. Sun, Y. Wang, J. Yang, and X. Hu, "Convolution Neural Networks With Two Pathways for Image Style Recognition," IEEE Trans on Image Process, vol. 26, no. 9, pp. 4102-4113, Sep. 2017. doi:10.1109/TIP.2017.2710631.

[14] C. Florea, C. Toca, and F. Gieseke, "Artistic Movement Recognition by Boosted Fusion of Color Structure and Topographic Description," in *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Santa Rosa, CA, USA, Mar. 2017, pp. 569-577. doi:10.1109/WACV.2017.69.

[15] A. Elgammal, M. Mazzone, B. Liu, D. Kim, and M. Elhoseiny, "The Shape of Art History in the Eyes of the Machine," in *32nd AAAI Conference on Artificial Intelligence*, New Orleans, USA, Feb. 2–7, vol. abs/1801.07729, 2018.

[16] S. Bianco, D. Mazzini, and R. Schettini, "Deep Multibranch Neural Network for Painting Categorization," in *Image Analysis and Processing - ICIAP 2017*, Lecture Notes in Computer Science, S. Battiato, G. Gallo, R. Schettini, and F. Stanco, Eds. Cham: Springer International Publishing, 2017, vol. 10484, pp. 414-423. doi:10.1007/978-3-319-68560-1_37.

[17] G. Folego, O. Gomes, and A. Rocha, "From Impressionism to Expressionism: Automatically identifying van Gogh's paintings," in *2016 IEEE International Conference on Image Processing (ICIP)*, Phoenix, AZ, USA, Sep. 2016, pp. 141-145. doi:10.1109/ICIP.2016.7532335.

[18] V. A. Encyclopedia.Wikiart, WikiArt.org - Visual Art Encyclopedia, 2018. [Online]. Available: https://www.wikiart.org.

[19] C. Florea, R. Condorovici, C. Vertan, Pandora18K, 2018. [Online]. Available: http://imag.pub.ro/pandora/pandora_download.html.

[20] C. Florea, R. Condorovici, C. Vertan, R. Butnaru, L. Florea, and R. Vrânceanu, "Pandora: Description of a painting database for art movement recognition with baselines and perspectives," in *2016 24th European Signal Processing Conference (EUSIPCO)*, Budapest, Hungary, Aug. 2016, pp. 918-922. doi:10.1109/EUSIPCO.2016.7760382.

[21] I. S. A. Krizhevsky, and G. E. Hinton "Imagenet classification with deep convolutional neural networks," Advances in Neural Information Processing Systems 25 (NIPS 2012), pp. 1097–1105, 2012.

[22] K. S. a. A. Zisserman, " Very deep convolutional networks for large-scale image recognition," CoRR, vol. 1409.1556, 2014.

[23] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, et al., "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, Jun. 2015, pp. 1-9. doi:10.1109/CVPR.2015.7298594.

[24] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770-778. doi:10.1109/CVPR.2016.90.

[25] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 2818-2826. doi:10.1109/CVPR.2016.308.