IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received March 4, 2019, accepted March 17, 2019, date of publication March 28, 2019, date of current version April 12, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2907986

Two-Stage Deep Learning Approach to the Classification of Fine-Art Paintings

CATHERINE SANDOVAL, (Member, IEEE), ELENA PIROGOVA, AND MARGARET LECH, (Member, IEEE)
School of Engineering, RMIT University, Melbourne, VIC 3000, Australia

Corresponding author: Catherine Sandoval (s3465033@student.rmit.edu.au)

This work was supported in part by the Australian Government Research Training Program Scholarship.

ABSTRACT Due to the digitization of fine art collections, pictures of fine art objects stored at museums and art galleries became widely available to the public. It created a demand for efficient software tools that would allow rapid retrieval and semantic categorization of art. This paper introduces a new, two-stage image classification approach aiming to improve the style classification accuracy. At the first stage, the proposed approach divides the input image into five patches and applies a deep convolutional neural network (CNN) to train and classify each patch individually. At the second stage, the outcomes from the individual five patches are fused in the decision-making module, which applies a shallow neural network trained on the probability vectors given by the first-stage classifier. While the first stage categorizes the input image based on the individual patches, the second stage infers the final decision label categorizing the artistic style of the analyzed input image. The key factor in improving the accuracy compared to the baseline techniques is the fact that the second stage is trained independently on the first stage using probability vectors instead of images. This way, the second stage is effectively trained to compensate for the potential mistakes made during the first stage. The proposed method was tested using six different pre-trained CNNs (AlexNet, VGG-16, VGG-19, GoogLeNet, ResNet-50, and Inceptionv3) as the first-stage classifiers, and a shallow neural network as a second-stage classifier. The experiments conducted using three standard art classification datasets indicated that the proposed method presents a significant improvement over the existing baseline techniques.

INDEX TERMS Fine art style recognition, painting classification, machine learning, multi-stage classification, transfer learning, digital humanities.

I. INTRODUCTION

This paper presents a multi-stage machine learning approach to the problem of semantic categorization of images depicting fine art paintings. The proposed approach shows how a machine can efficiently recognize an artistic style. By doing so, the study addresses the semantic gap problem, which is one of the most enduring challenges in machine-based image retrieval and recognition.

Image classification in general can be viewed as a task of assigning a label to an image, which allows it to be placed within a specific category. Depending on the types of labels, images can be classified based on ‘‘what they depict’’. This is known as object recognition (e.g. recognition of landscapes vs portraits, cats vs. dogs, types of written characters, etc.). The second type of labeling is based on ‘‘what is the meaning’’ of the image. This is known as the semantic categorization (e.g. sad images vs happy images, safe vs dangerous road scene, aesthetically pleasing vs unpleasing scenes, etc.). While the current machine learning techniques can very efficiently solve the object recognition task, the semantic categorization, which has a subjective person-dependent character, is still largely an unexplored area.

The ability to recognize an artistic style in fine art paintings is an attribute of highly educated and experienced art scholars who spend years analyzing and learning the specifics and nuances of the fine art objects. For many years, these skills had a very elite character since they had to be acquired through a lengthy process of visual experience. Due to the rapidly expanding availability of the online galleries, as well as various other sources of fine art pictures, fine art has become accessible to the masses, which in turn created a

The associate editor coordinating the review of this manuscript and approving it for publication was Xi Peng.

need to make the art expertise more easily available to the public. One of the solutions to this problem is to transfer the subjective knowledge of human experts into machines. In other words, by training machines with many pictures of fine art paintings labeled by human experts, the machines can learn how to label unknown images and recognize their artistic styles. Machine-based art expertise can be used in automatic image retrieval, in art classes and in labeling of unsigned paintings at auction houses. It can also assist to detect lost masterpieces, or create robots working as human companions with a human-like sense of aesthetics and art appreciation.

Style is one of the most common semantic criteria used in painting classification. In visual arts, style is defined as a set of distinctive elements that can be associated with a specific artistic movement, school or time period [1]. However, the classification of a painting's unique stylistic category is a complex task even for an expert. Some of the challenges include ambiguous interpretability of abstract and aesthetic elements, nuances separating different artistic categories, smooth transitions between art periods, the presence or absence of artistic attributes that belong to multiple styles or do not belong to any style, and variations in the style of the same artist [2], [3].

In the last decade, there has been a growing interest in automatic art classification. Classical approaches focus on determining the optimal set of features to be extracted from the paintings for classification [4]–[9] whereas, more modern, deep learning (DL) approaches address the problem of painting classification through the implementation of the transfer learning (or fine-tuning) of different Convolutional Neural Networks (CNNs). Alternatively, the fine-tuned CNNs can be used to generate features to train various non-network classifiers [10]–[16].

Transfer learning allows the adaptation or reuse of a network model that has been trained for a specific task using a very large dataset to perform a new, related task for which only a small dataset is available. Therefore, a complex network model, pre-trained on a very large dataset of natural images, can be easily modified to perform style painting classification with the advantage of a significant reduction of training time and the use of datasets that are considerably smaller than the database used in the pre-training task. DL approaches have shown to provide promising results in painting style classification, but not as high as results achieved in image object classification [17]–[21]. One of the possible reasons is that standard CNN architectures, used in most studies, require a fixed input image size that is in most cases, significantly smaller than the high-resolution art images offered by fine art datasets. This means that the analyzed images must undergo a downsizing process to fix the dimension to the required standard. This process can lead to geometric distortion, content deformation and loss of relevant information. The associated loss of detailed characteristics related to the texture and composition, such as position, length, width and orientation of the brushstrokes and light, color and shape variations can be detrimental to the accuracy of fine art style categorization [1], [21].

To address this problem, a sub-region approach was introduced in which the original image is divided into smaller regions (or image patches), where the size of each patch matches the standard image input size required by the CNN model [22]. Despite the apparent advantages of the sub-region painting analysis, this approach has not yet been comprehensively explored. Therefore, it is difficult to evaluate its efficacy in comparison with the standard method of the full-image analysis and classification.

This study provides an extensive investigation into the patch-based approach to fine art classification. The investigation leads to the proposal of a new, synergistic two-stage fine art categorization method. First, the proposed approach divides the input image into patches and applies a deep neural network to train and classify each patch. At the second stage, the final style label is generated by a shallow neural network trained on the probability vectors given by the first-stage classifier. The proposed method is evaluated and compared with a baseline approach using three standard fine-art classification databases. The effects of different complexity of the pre-trained CNN are investigated by testing six popular CNN models, and a comparison with the current state-of-the-art techniques is presented.

The remainder of this paper is organized into five sections: a brief review of previous studies about style painting categorization is presented in Section II, Section III describes the proposed method, Section IV provides the detail of the experiments and datasets, Section V presents the experimental results and discussion and finally, the paper concludes in Section VI.

II. RELATED WORKS

The fine art classification task has been addressed through a wide range of methods. Most can be grouped into two major categories: classical and deep learning techniques. Classical approaches address the task of style classification using image descriptors, in which a set of low-level parameters is first extracted from the input image and then categorized using one of the standard classification algorithms [4]–[9].

Early studies tested the feasibility of style classification on very small datasets of images and using only a few style categories. For example in [4], three styles were classified using a dataset of 513 images. The features were extracted using several transforms, such as Fourier, Chebyshev and Wavelet, whereas the Weighted Nearest Neighbor (WNN) method was implemented as a classifier. Other examples of relevant works based on the extraction of low-level features include methods proposed in [5] and [6]. The former applied techniques such as the color scale-invariant feature transform (CSIFT) and the opponent-SIFT (O-SIFT) algorithm to extract features and classify seven different art styles using a training dataset of 490 paintings. The latter implemented features extracted from the Local Binary Patterns (LBP), color LBP, the Generalized Image Search Tree

(GIST), the Pyramid of Histograms of Orientation Gradients (PHOG), the Histogram of Oriented Gradients (HOG) parameters and the Scale-Invariant Feature Transform (SIFT) techniques using a dataset (Painting-91) with 13 styles and a total of 4266 images. Both studies implemented the Support Vector Machine (SVM) algorithm to classify the features. It was concluded that low-level features yield relatively poor multi-class classification results [5], [6]. Different combinations of features and classifiers were explored in [7]. In a recent work [8], the classification of three artistic styles based on the qualitative color descriptors and color similarity was explored using the SVM and the k-Nearest Neighbors (k-NN) as classifiers. In [9] unsupervised feature extraction was investigated, for the classification of 6776 paintings into eight stylistic groups. However, in all cases, no significant improvements were achieved. The first breakthrough was due to the application of DL techniques and the development of image classification CNNs that were pre-trained on very large datasets of images.

Recent studies are dominated by various applications of fine-tuned, pre-trained, CNN models that can both learn the features and infer the style label. The concept of fine-tuning of a pre-trained network on a small dataset is known as transfer learning. Several DL approaches have adapted a pre-trained CNN as a feature extractor [10]–[16]. Since the features in these cases were given as the network parameters, there was no longer a need to determine knowledge-based features. The classification process of the network parameters was done by a linear classifier, such as the SVM. One of the first large-scale studies of the fine art classification was reported in [10].

Twenty-five styles represented by a large dataset of digitized paintings were categorized using features generated by a pre-trained CNN model. The SVM was implemented as a classifier. There was a consistency in the outcomes of these studies. They confirmed that style classification models, based on DL features, outperform classical models which are based on engineered knowledge-based features [10]–[16]. In [16], for example, the accuracy obtained using a big set of hand-designed visual descriptors was compared with the accuracy achieved by the implementation of feature extraction using pre-trained CNN models. It demonstrated that the CNN-feature extraction technique performed significantly better. Further improvement of style classification was achieved through the application of transfer learning with the CNN model used to do both, feature learning and label inference.

One of the first systematic studies using this method was reported by Tan et al. [17]. A pre-trained object classification CNN, known as AlexNet, was used to classify a large number of paintings representing over 27 stylistic categories. It was demonstrated that the CNN fine-tuned models can outperform standard, non-network classifiers trained on features extracted from CNNs as well as CNN model trained ''from scratch'' (i.e. without pre-training). This was confirmed by a number of other similar studies [17]–[21]. In [20], it was suggested that transfer learning based on a CNN model, pre-trained on scene recognition or image sentiment analysis, rather than object classification data, coul sification results. Whereas, in [21] style classification results obtained by fine-tuning three different CNN models were compared.

Learned CNN representations of styles were analyzed using the principal component analysis (PCA). The findings revealed a strong correlation between the CNN features and the chronology of paintings.

One of the distinct lines of research closely related to the current study stems from [22], where painting authorship recognition based on image sub-sections, called patches, was explored. The study implemented the SVM model trained on features extracted from the CNN model to perform a binary identification of Vincent van Gogh's paintings. The CNN was generated using a relatively small dataset of 333 images of paintings. Each of the analyzed images was divided into several non-overlapping patches, and each patch was individually classified. The final decision was given by the patch with the highest classification score. Image analysis based on datasets containing images of varied sizes was explored in [23]. Artist recognition based images of paper prints of artworks were performed using the average score obtained by independent CNNs trained on different image scales. A similar artist recognition approach was reported in [24] where a multiscale pyramid framework comprised of three layers was applied. Fixed-size input images were analyzed by the first CNN layer, while the second layer analyzed four patches, and the third layer analyzed sixteen patches extracted from the input images enlarged by two and four times, respectively. The final result was determined by the category that scored the highest average class entropy.

Another interesting patch-based approach was proposed in [25]. A complex three-branch CNN structure was trained using a dataset of 2338 images with 13 categories to classify paintings by style. Three random patches were used as inputs to the CNN. Two of these patches were taken from the original input image, and the third one was extracted from a downsized version of the same image. A boosted ensemble of SVMs was proposed in [26] to classify artistic styles over a large dataset of images. Color histograms and topographic descriptors of images were implemented as features. The final decision was determined by the majority voting of the classification outcomes resulting from the analysis of the whole painting and several random sub-regions.

As shown in our previous work [27], a significant improvement of the classification results can be achieved using a weighted sum of outcomes for individual patches. We have introduced an optimized approach, where individual patches were independently classified by a CNN, and then a weighted average of the classification outcomes for individual patches was estimated to determine the final style label. The weight values were determined via numerical optimization with an objective to maximize the overall style classification accuracy. In this study, we propose a new two-stage classification algorithm which offers further improvement of the patch-based style classification results.

III. PROPOSED METHOD

Inspired by the patch-based analysis and the multi-stage classification method introduced in [28], a two-stage classification system with four data processing steps is proposed to achieve the style classification task. The functional block-diagram of the classification framework is shown in Fig. 1.

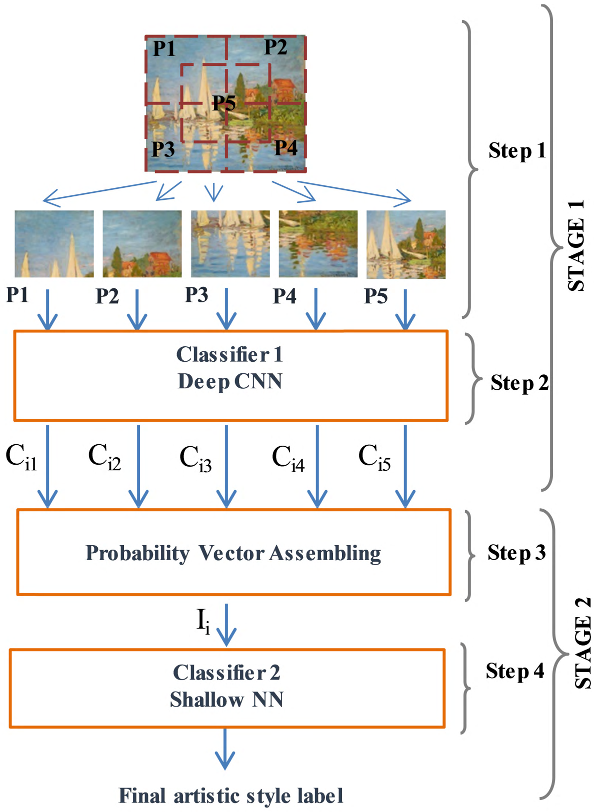


FIGURE 1. The proposed two-stage classification framework using five image patches; i is the analyzed image index.

The style classification is achieved in two stages. In the first stage, the analyzed images are divided into five patches (sub-images P1-P5). A deep CNN model is used to classify the artistic style for each patch. In the second stage, the intermediate CNN classification outcomes (probability vectors C1-C5) for the individual patches are assembled into a single input vector to a shallow neural network trained to provide the final style label. While the first stage deep CNN classifier is trained on images, the second stage shallow NN is trained on the class-probability vectors resulting from the first stage classification. Using as analogy, the first stage classifier effectively works as an assembly of assessors, each making judgments based on a different part of the original image, while the second stage classifier learns how to assess the classification skills of the first stage assessors, and how to use this knowledge to make the final decision. The following sub-sections explain the details of the proposed methodology.

A. STEP 1 - PATCH EXTRACTION

The analyzed image is divided into five patches. Before this division takes place, the image needs to be scaled up by an appropriate factor to achieve patch sizes that comply with the input requirements of a given CNN model. As shown in Fig. 1, the image upper right, upper left, lower right and lower left sections are represented by the first four patches: P1, P2, P3 and P4, respectively; whereas the center of the painting is represented by the fifth patch (P5). The fifth patch (P5) overlaps with 25% of each of the other four patches.

B. STEP 2 - DEEP CNN CLASSIFIER

Each of the five patches (P1-P5) generated in Step 1 are then passed to a CNN model (Classifier 1 in Fig. 1) to infer the intermediate style classification outcome (C1-C5) for each patch. Depending on the available computational and data resources, the CNN model can be either trained from scratch or transfer learning can be applied to fine-tune a pre-trained model. While the first option requires lengthy training and large data, the latter option can be efficiently used when the resources are scarce. When using the transfer learning option, the last three layers, including the last fully connected layer, the softmax layer and the classification output layer of the pre-trained CNN structure, must be customized for the given style classification task to match the required number of individual-patch outcomes and the number of possible classes. The number of different artistic styles determines the size of the last fully-connected layer. The softmax layer delivers a vector describing what are the probabilities for each patch to belong to each of the possible artistic style categories. The classification output layer assigns one of the mutually exclusive stylistic categories to each patch. The $C_{i,j}$ output vectors generated in Step 2 (see Fig. 1) contain style probabilities $p_{i,j,k}$ estimated for each patch j of a given image i as in (1).

$$C_{i,j} = \{p_{i,j,k}\}_{k=1,\ldots,L} \tag{1}$$

where i is the index of the analyzed input image ($i = 1, \ldots, M$), j is the patch number ($j = 1, \ldots, N$), and k is the style index ($k = 1, \ldots, L$).

C. STEP 3 - PROBABILITY VECTOR ASSEMBLING

The probability vectors $C_{i,j}$ generated in Step 2 that belong to the same image i ($i = 1, \ldots, M$) are concatenated into a single vector of probabilities I_i as given in (2).

$$\mathrm{I}_i = [p_{i,1,1}, p_{i,2,1}, \ldots, p_{i,\mathrm{N},1}, \ldots\ldots, p_{i,1,\mathrm{L}}, p_{i,2,\mathrm{L}}, \ldots, p_{i,N,L}] \tag{2}$$

The vector of probabilities I_i is passed as the input features to the second stage classifier (Classifier 2 in Fig. 1).

D. STEP 4 - SHALLOW NN CLASSIFIER

The concatenated probability vectors I_i from Step 3 are presented as features to the second classifier that delivers the final style classification label. A standard classifier, such as the multiclass SVM, Gaussian Mixture Model (GMM), or a shallow neural network, can be used to classify the features. The second-stage classifier is trained alongside with the first stage CNN in a single two-stage process, where the CNN training scores (style probabilities) obtained at the first stage are used as the training features for the second classifier. Once both classifiers are trained, the same two-stage procedure can be used to infer a style label for an unlabeled input image.

IV. VALIDATION EXPERIMENTS

A. DATASETS

Three datasets of digital images of paintings collated from publicly available fine art collections, with the addition of an Australian Aboriginal art dataset created by the authors, were used to empirically validate the efficacy of the proposed method.

1) DATASET 1

Dataset 1 included 30870 images representing six artistic styles. The stylistic classes were balanced, and each class was represented by 5145 images (16.66% of the total number of pictures). The styles included: Australian Aboriginal Art, Expressionism, Impressionism, Post Impressionism, Realism and Romanticism. The last five styles were selected from the otherwise larger WikiArt dataset [29], as the only styles with a sufficiently large number of images to match the number of Aboriginal-style images. Since the WikiArt images were not labeled by art experts, but by general-public volunteers, the pictures had to be manually verified to ensure the correctness of labels, and to eliminate images having poor quality or not depicting fine-art paintings.

2) DATASET 2

Dataset 2 was created to cover a larger number of artistic styles than Dataset 1. As for Dataset 1, Dataset 2 was largely generated from the WikiArt collection with the addition of Australian Aboriginal paintings. The original WikiArt collection included more than 85000 paintings categorized into 27 styles. However, the numbers of pictures representing different styles were highly imbalanced varying from 12000 to 98 pictures per style [29]. To achieve more balanced representation, only 23 out of 27 stylistic categories were selected from the WikiArt dataset, with three classes related to cubism merged in one single class, giving a total of 21 WikiArt styles. With the addition of the Australian Aboriginal style, the Dataset 2 included 26400 images and 22 styles. The style representation was balanced, with each style represented by 1200 images (5% of the total number of pictures). Apart from giving a good style representation, these numbers kept the computational effort needed to process the pictures within the limits of the available hardware. Fig. 2 shows the styles and their distribution for Dataset 2.

3) DATASET 3

Dataset 3 contained images of fine art paintings from the Paintings Dataset for Recognizing the Art Movement

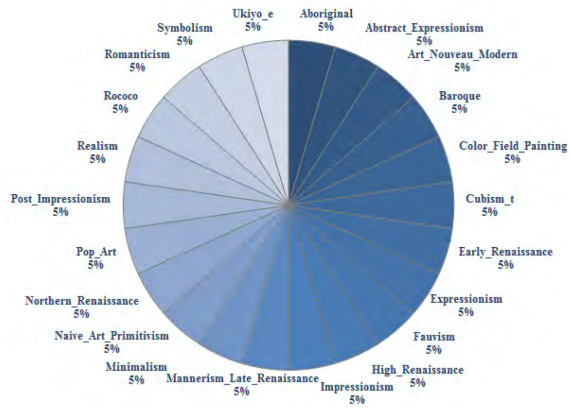


FIGURE 2. Dataset 2 - the number of pictures per style (in percentage).

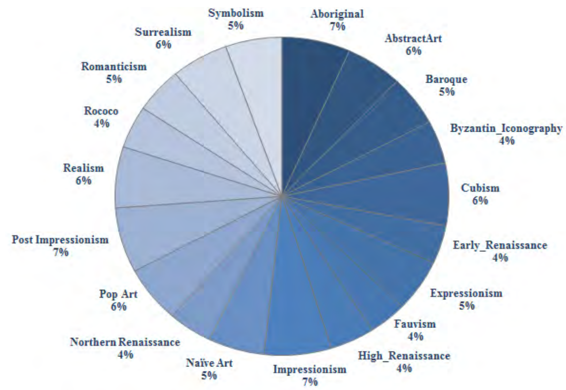


FIGURE 3. Dataset 3 - the number of pictures per style (in percentage).

(Pandora 18K) collection [30], [31] plus the Australian Aboriginal style. In total, Dataset 3 comprised of 19320 images and 19 styles. As shown in Fig. 3, the numbers of images were quite evenly distributed across styles. The most important advantage of the Pandora 18K dataset over the WikiArt collection was the high validity of the labeling process. Unlike the WikiArt labels that were made by public volunteers, the Pandora labels were strictly assigned by art experts.

B. EXPERIMENTAL SETUP

The proposed method was evaluated using six standard CNN architectures, AlexNet [32], VGG-16 [33], VGG-19 [33], GoogLeNet [34], ResNet-50 [35], and Inceptionv3 [36]. These networks have been pre-trained on the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) dataset that contains 1.2 million natural images of objects with 1000 categories [37].

The last three fully connected layers of each CNN were modified to comply with the number of classes for a given dataset. The modified networks were then fine-tuned to the specific task of art style recognition. Table 1 shows the most relevant characteristics of the CNN models. AlexNet, VGG-16, VGG-19 have a linear stack or sequential architecture while GoogLeNet, ResNet-50 and Inceptionv3 have

TABLE 1. CNN models characteristics.

Model	No of Layers	Architecture	Input Image Size	Size (MB)	Parameters (Millions)
AlexNet	8	linear	227 x 227	227	61
VGG-16	16	linear	224 x 224	515	138
VGG-19	19	linear	224 x 224	535	144
GoogLeNet	22	Inception module	224 x 224	27	7
ResNet-50	50	Residual blocks	224 x 224	96	25.6
InceptionV3	48	Inception module	299 x 299	89	23.9

more complex, deeper, non-linear architectures comprised of inception modules or residual blocks that act as sub-networks.

To train the CNN models, 80% of the data was used and the system performance was tested with the remaining 20%. A three-fold cross-validation scheme was adopted, and the reported results are given as the average classification accuracy. The classification accuracy was defined as [38]:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

where TP, TN, FP and FN denote the number of true positives, true negatives, false positives and false negative predictions, respectively. For each dataset, and for each one of the six pre-trained CNN models, seven different classification scenarios were evaluated and compared. Scenarios 1 and 2 represent simple single-stage classification methods, whereas Scenarios 3-7 implement various versions of the proposed two-stage method, where the outcomes of the first stage classification are used at the second (decision making) stage to derive the final label. Two versions of the proposed method are tested in Scenarios 6 and 7.

1) SCENARIO 1 - BASELINE

Only a single stage classification was applied. There was no division of the CNN input images into patches. Each image had to be re-sized to fit into the dimensions required by a given CNN architecture. The accuracy of each CNN model was calculated as the average over all tested images.

2) SCENARIO 2 INDIVIDUAL PATCHES

As in Scenario 1, only single stage classification was applied. However, the input images were divided into five patches, and each patch was classified individually (i.e. a separate style label was provided for each patch). No information about connections between patches (belonging to the same image) was provided to the system. To comply with the input size requirement of CNNs, each of the original images had to be up-sampled to achieve twice of the required CNN input size and then divided into five patches. Due to this sub-division, the training dataset was five times larger than the one used in the baseline test. The final classification accuracy was calculated as the average over all testing patches.

3) SCENARIO 3 – MAJORITY VOTING

In this testing scenario, the same image sub-division procedure as in Scenario 2 was used, however, for each of the original input images the final style label was determined by a simple majority voting system over the five patches classification outcomes. The label that was voted most often was selected as the category of the painting. In the case of a tie between two or more labels, the final decision was made by choosing the label with the highest probability value. The accuracy of each CNN model was calculated as the average over all tested images.

4) SCENARIO 4 - AVERAGE PROBABILITY

This scenario was identical to Scenario 3, except that instead of using the majority voting system, an average probability for each class was calculated over five patches of a given image. The final label was assigned to the class with the largest average probability. The accuracy of each CNN model was calculated as the average over all tested images.

5) SCENARIO 5 - WEIGHTED AVERAGE

The same patch classification procedure as in Scenarios 3 and 4 was applied. However, instead of calculating the majority vote or the average probability value, the style of a given painting was determined as a weighted average of the classification outcomes (probability vectors) obtained for each of the five patches. The optimal weight values were derived by a numerical optimization algorithm that aimed to maximize the overall system classification accuracy. Details of this approach can be found in [27].

6) SCENARIO 6 – PROPOSED TWO-STAGE CLASSIFICATION USING PATCHES ONLY (CNN/NN-P)

This scenario applied the proposed two-stage method described in Section III. The probability vectors obtained during the first-stage classification of individual patches of a given input image were assembled into feature vectors, as in (2) and passed to the second stage classifier to determine the final label for the analyzed image. To determine the most suitable second-stage classifier, preliminary tests were conducted with the probability vectors being generated by the AlexNet and passed to either a Subspace Discriminant classifier, coarse K-NN, multiclass SVM, or a shallow neural network (NN). The SVM and the NN showed the best results with Dataset 1, and the NN presented the best results with Datasets 2 and 3. Given its good performance, high training flexibility and relative simplicity of implementation, a shallow NN was chosen as the second classifier. Table 2 shows the essential parameters of the shallow NN that were used for each dataset. A five-fold cross-validation scheme was adopted to train the shallow NN with 65% of the data for training, 15% for validation and 20% for testing.

TABLE 2. Characteristics of the shallow NN classifiers.

Dataset	No Layers	No Nodes	Probability Vector Size	No Categories
Dataset1	3	3600	30	6
Dataset2	5	2850/570	110	22
Dataset3	5	2375/475	95	19

7) SCENARIO 7 - PROPOSED TWO-STAGE CLASSIFICATION USING PATCHES AND THE WHOLE IMAGE (CNN/NN-P&I)

In this scenario, the same approach as in Scenario 6 was implemented, however, in addition to the classification of individual patches, the first-stage classification also included classification of the whole image (resized to fit the CNN input). Thus, the feature vectors passed to the second-stage shallow NN were longer, as they were assembled by concatenating the patch probabilities with the image probabilities of different artistic styles. In other words, the whole image was added to the system as a "sixth patch".

V. RESULTS AND DISCUSSION

The overall system performance is presented in Tables 3-5. Both versions of the proposed two-stage classification approach (Scenario 6 and Scenario 7) clearly outperformed all other methods. This applies to all three datasets and to all CNN models. It confirms the benefits of having a second-stage classifier trained not on the images but on the class-probability vectors. The second-stage classification appears to compensate for the mistakes made during the first-stage classification. It is also important to note that because both stages are trained independently, the second-stage training does not alter results of the first stage training. This means that in general, if the first stage of classification performs well, the second stage may not contribute a lot, but it will not disturb the results of the first stage. However, if the first stage is not performing well (possibly due to a poor choice of features or insufficient training), the second stage may provide significant benefits while eliminating the need for a longer more data consuming training of the first-stage classifier. The nonlinear NN modeling applied in the second stage clearly outperformed simple linear decision-making approaches such as the majority vote, average probability and the weighted average. Again, this observation applies to all investigated datasets and CNN models.

The pre-trained CNN models are listed in Tables 3-5 in order of increasing structural complexity with the AlexNet being the simplest and the Inception V3 the most complex. It can be noticed that the higher the CNN complexity, the better the classification outcomes disregarding the method and the database. This outcome confirms general expectations that the more complex structures tend to perform better in various classification tasks. However, it is also known that the more complex is the network structure, the more training data is needed to achieve the full benefits. Since the data size used in our experiments was the same for all CNN models, it is likely that the results achieved for the most complex models, such as the GoogLeNet or the ResNet-50, could be further improved with a greater amount of labeled data. Less conveniently, this also implies the availability of much higher computational resources. While some of the recent relevant studies [18], [21] suggested that the ResNet-50 is the best performing CNN model for painting style classification, the results obtained in this study show that the InceptionV3 model outperforms the ResNet-50 by 0.4% - 1.10% for Dataset 1, 0.30% -0.8% for Dataset 2 and 0.5 - 1.63% for Dataset 3 (Table 3-5).

Another important factor affecting the classification outcomes was the quality of the training dataset. Clearly, the best results across all methods and CNN models were achieved with Dataset 3. Dataset 2 and Dataset 3 gave similar performance, however, in both cases it was about 10% lower than for Dataset 3. These results are consistent with the high quality of Dataset 3. It was created using professional expertise and expert labeling, whereas Datasets 1 and 2 were both labeled by public volunteers.

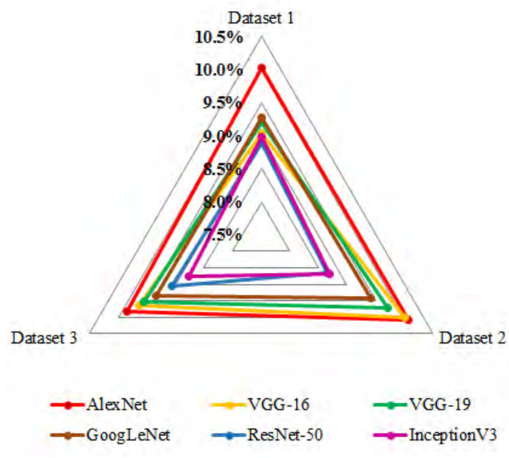


FIGURE 4. Differences between the accuracy of the proposed method (Scenario 6), and the baseline (Scenario 1) when using different CNN models.

Similar conclusions can be derived from Fig. 4. It shows the differences between the percentage accuracy of the proposed method (Scenario 6), and the baseline (Scenario 1) when using different CNN models and different datasets. Generally, the higher was the CNN model complexity, and the better was the quality of the training data, the smaller was the difference between the baseline and the proposed method.

A comparison between Scenario 6 and Scenario 2 directly determines the effect of adding the second-stage classifier on the overall classification results. Thus, for example, when using the simplest CNN, AlexNet, the classification accuracy in Scenario 6 (two stages) is 13% higher than in Scenario 2 (one stage) for Dataset 1, 15% higher for Dataset 2, and 15% higher for Dataset 3. When using the most complex CNN, InceptionV3, the classification

TABLE 3. Dataset 1 - Average accuracy (%) of style classification across different classification scenarios and CNN models.

CNN Model	Scenario 1: Baseline	Scenario 2: Individual Patches	Scenario 3: Majority vote	Scenario 4; Average Probability	Scenario 5 Weighted Average	Scenario 6: **CNN/NN – P (patches only)**	Scenario 7: **CNN/NN -P&I (patches and image)**
AlexNet	51.50%	48.50%	52.02%	53.00%	54.40%	**61.53%**	**62.46%**
VGG-16	52.80%	51.40%	53.10%	54.50%	56.04%	**61.84%**	**62.69%**
VGG-19	52.90%	51.60%	53.40%	54.30%	55.90%	**62.10%**	**62.81%**
GoogLeNet	54.50%	52.00%	55.50%	56.10%	57.60%	**63.78%**	**64.42%**
ResNet-50	56.80%	53.40%	56.90%	57.40%	58.50%	**65.70%**	**66.64%**
InceptionV3	57.20%	54.50%	57.80%	58.30%	59.40%	**66.18%**	**67.16%**

TABLE 4. Dataset 2 - Average accuracy (%) of style classification across different classification scenarios and CNN models.

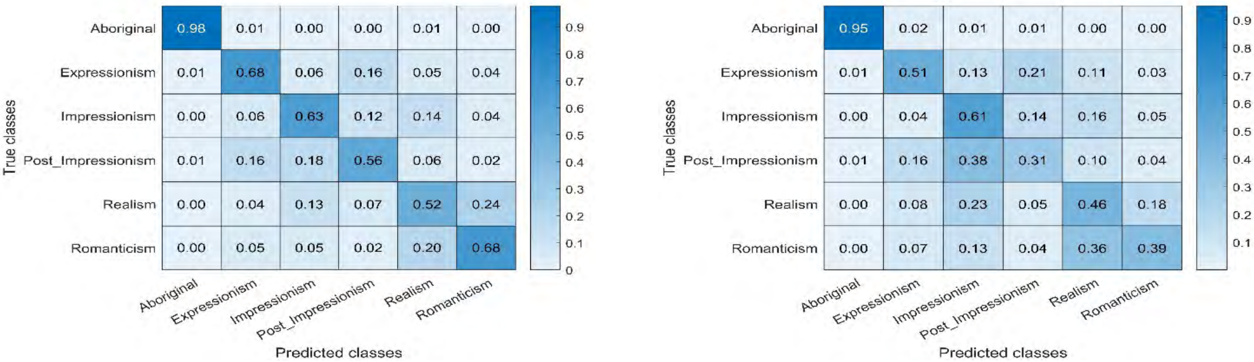
CNN Model	Scenario 1: Baseline	Scenario 2: Individual Patches	Scenario 3: Majority vote	Scenario 4; Average Probability	Scenario 5 Weighted Average	Scenario 6: **CNN/NN – P (patches only)**	Scenario 7: **CNN/NN -P&I (patches and image)**
AlexNet	49.30%	44.80%	51.15%	52.10%	54.50%	**59.37%**	**60.27%**
VGG-16	51.15%	47.60%	53.14%	54.60%	56.70%	**61.17%**	**62.11%**
VGG-19	51.85%	48.10%	52.65%	55.10%	57.10%	**61.57%**	**62.49%**
GoogLeNet	53.95%	50.00%	54.78%	57.20%	58.90%	**63.37%**	**64.27%**
ResNet-50	56.55%	51.80%	57.43%	58.90%	60.90%	**65.13%**	**66.02%**
InceptionV3	57.10%	52.10%	57.97%	59.70%	61.70%	**65.83%**	**66.71%**

TABLE 5. Dataset 3 - Average accuracy (%) of style classification across different classification scenarios and CNN models.

CNN Model	Scenario 1: Baseline	Scenario 2: Individual Patches	Scenario 3: Majority vote	Scenario 4; Average Probability	Scenario 5 Weighted Average	Scenario 6: **CNN/NN – P (patches only)**	Scenario 7: **CNN/NN -P&I (patches and image)**
AlexNet	62.20%	57.30%	62.62%	63.50%	64.80%	**72.04%**	**73.11%**
VGG-16	63.00%	58.20%	63.15%	64.20%	65.30%	**72.64%**	**73.61%**
VGG-19	62.87%	58.40%	63.01%	63.97%	65.07%	**72.41%**	**73.36%**
GoogLeNet	64.10%	60.10%	64.43%	65.10%	66.00%	**73.44%**	**74.37%**
ResNet-50	65.97%	63.90%	66.57%	67.77%	68.57%	**75.23%**	**76.14%**
InceptionV3	67.60%	63.40%	67.98%	69.40%	70.20%	**76.57%**	**77.53%**

accuracy in Scenario 6 is 12% higher than in Scenario 2 for Dataset 1, 14% higher for Dataset 2, and 13% higher for Dataset 3. It shows that the better the first-stage classification, the smaller the improvement given by the second stage. In other words, the addition of the second stage is most effective in cases where the first stage performs poorly due to either insufficient training or poor quality of features.

Several further observations can be derived by a more detailed comparison between different classification approaches. The patch-based classification in Scenario 2 (with each patch being individually classified) showed the lowest accuracies. There was a decrease in the accuracy by 2.1% to 4.9% with Dataset 3, by 3.5% to 5% with Dataset 2, and by 1.4% to 3.4% with Dataset 1 when compared to the baseline (Scenario 1). This can be explained by the fact that in Scenario 2, no information about the links between patches belonging to the same image was used during the classification process, and the labeling was based only on partial information about the image composition. Whereas, the baseline recognition used the whole image, thus the overall composition could be analyzed. This could also explain the fact that the proposed approach based only on the classification of image patches (Scenario 6) led to about 1% lower accuracy compared with Scenario 7, where the classification was based on both patches and the whole image. Similarly, when the results of the individual patch classification were combined to generate the final decision by using the majority voting method (Scenario 3) or by calculating the non-weighted (Scenario 4) or the weighted average (Scenario 5), the information about the links between patches that belong to the same image was re-introduced to the system. This was possibly the main factor that increased the accuracy obtained in Scenarios 3-5 compared to the baseline (Scenario 1). Therefore, it can be concluded that the local patch-based information, as well the global, whole image information, both play an important role during stylistic art analysis.



a) b)

FIGURE 5. Confusion arrays using the Inceptionv3 CNN model tuned on Dataset 1. a) Proposed approach (Scenario 6 - patches only). b) Patch-based baseline (Scenario 2). The values represent percentage accuracy (divided by 100), and the shade intensity of the table cells increases with the percentage accuracy value.

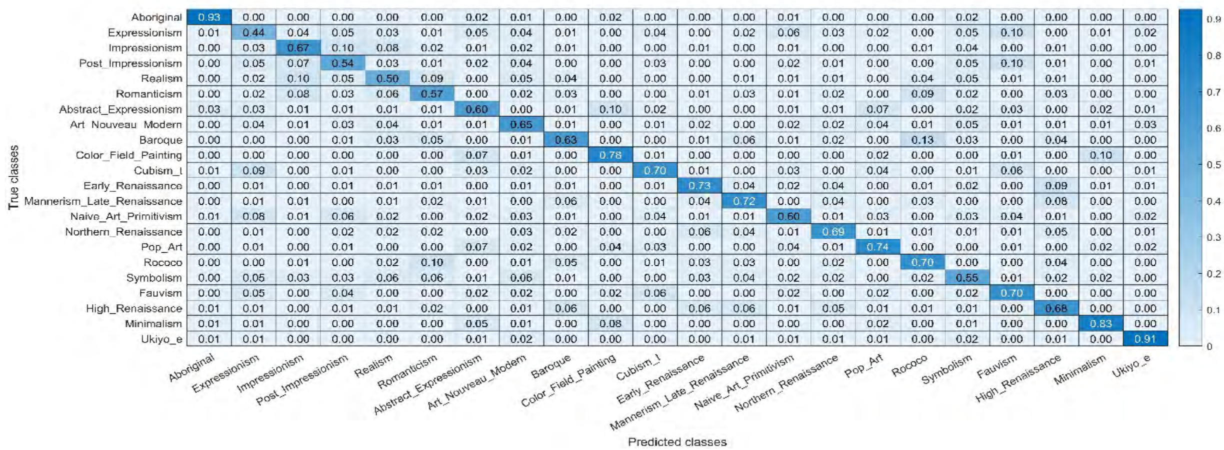


FIGURE 6. Confusion array for the proposed method with the Inceptionv3 CNN model tuned on Dataset 2. The values represent percentage accuracy (divided by 100), and the shade intensity of the table cells increases with the percentage accuracy value.

In comparison with other similar studies, it is worthwhile to note that a recent study based on the Pandora18K dataset [26] reported an average accuracy of 63.5% for the method using a large set of visual descriptors, sub-region analysis and boosted SVMs and an accuracy of 62.1% when using a fine-tuned ResNet-50 model. To obtain a direct comparison of the proposed two-stage classification method with the results reported in [26], an additional experiment was conducted using the Dataset3 with the exclusion of the Australian Aboriginal style. While an average classification accuracy of 63.98% consistent with [26] was achieved for the baseline Scenario 1 using the ResNet-50 model, a significant improvement was observed in Scenarios 6 and 7 when applying the proposed method. Namely, in Scenario 6 the average accuracies of 73.63% and 74.87% were achieved when using the ResNet-50 and the InceptionV3 models respectively. Whereas Scenario 7 provided accuracies of 74.84% and 75.98% for the ResNet-50 and the InceptionV3 models respectively.

While the technical details of image analysis and the classification system design are essential from the engineering perspective, a closer look into the classification of individual styles can reveal some interesting insights for art analysis. This can be done by the analysis of confusion arrays shown in Fig. 5, Fig. 6 and Fig. 7 for Dataset 1, 2 and 3, respectively. For simplicity, only the best performing case of Scenario 7 is illustrated in these figures.

Fig. 5 compares two confusion arrays; one obtained in Scenario 6 (the best performing case for patches only), and the other in Scenario 2 (the baseline patch-based approach). In both cases Dataset 1 was used with only six artistic categories. Ideally, one would like to see large values within the diagonal cells (top left to bottom right) of the confusion array (indicating high numbers of TP and TN classifications) and small numbers elsewhere (indicating small numbers of FP and FN classifications). The patterns in Fig. 5 show that the proposed method increased the diagonal values and improved their distribution across styles in comparison with

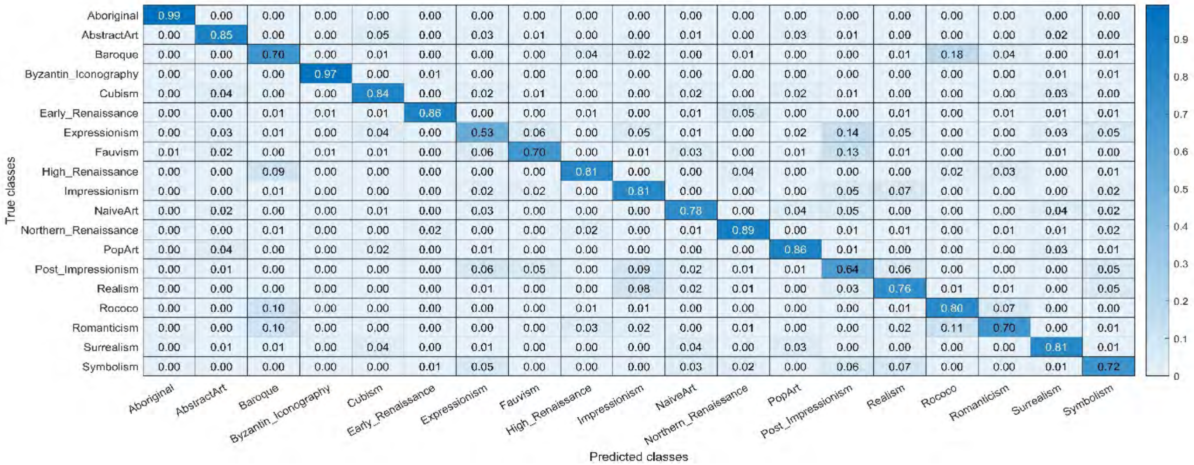


FIGURE 7. **Confusion array for the proposed method with the Inceptionv3 CNN model tuned on Dataset 3. The values represent percentage accuracy (divided by 100), and the shade intensity of the table cells increases with the percentage accuracy value.**

the baseline technique. For example, the proposed approach reduced misclassification of Romanticism as Realism from 36% to 20% and Post-Impressionism as Impressionism from 38% to 18%. This in turn contributed to the increase of the recognition accuracy from 39% to 68% for Romanticism, and from 31% to 56% for Post-Impressionism. Interestingly, both approaches achieved the highest recognition accuracy (above 90%) for the Australian Aboriginal art paintings.

Fig. 6 and Fig. 7 illustrate confusion arrays based on the proposed best performing approach (Scenario 7), the only difference was the database used to train the CNN InceptionV3 model. In the case illustrated in Fig. 6, Database 2 with labels generated by public volunteers was used, whereas in the case illustrated in Fig. 7, Database 3 with high-quality expert labeling was applied. Consistent with our previous observations, the lower labeling quality of the Dataset 2 is reflected in the lower classification accuracy achieved across individual styles. Although a direct comparison is difficult due to style differences between these two datasets, one can observe that the diagonal values (percentage accuracy per style) in Fig. 7 show generally higher and more evenly distributed values compared to the diagonal values in Fig. 6.

Amongst the styles represented in Dataset 2 (Fig. 6), the Ukiyo-e and the Australian Aboriginal styles were classified with the highest accuracies of 91% and 93%, respectively. This can be attributed to the fact that these two styles were the only non-western styles included in the Dataset 2, and as such their features were likely to stand out from the rest and make identification easier. Expressionism yielded the lowest accuracy of 44% and exhibited relatively high confusion rates with Fauvism, Naïve Art and Post-Impressionism. Baroque showed 13% of confusion with Rococo. There were also high confusion levels between Rococo and Romanticism, Early Renaissance and High Renaissance, and between Early Renaissance and Mannerism Late Renaissance. These results are consistent with difficulties in distinguishing between stylistically similar groups that are related or that exhibit smooth transitions between artistic movements [39].

Finally, looking at the style classification based on Dataset 3 (Fig. 7), the Australian Aboriginal and the Byzantine Iconography styles were recognized with the highest accuracies of 99% and 97%, respectively. The lowest accuracy was achieved for Expressionism (53%) which was mostly confused with Post-Impressionism. Relatively high confusion rates were also observed between Baroque and Rococo, and between Fauvism and Post-Impressionism. These results are expected as these styles are known to be closely related [39].

VI. CONCLUSION

A new machine learning method for automatic fine art style classification was presented and evaluated. The proposed approach applied two independently trained stages of classification. While the first stage applied a deep CNN trained directly on image data, the second stage used a shallow NN trained on the class probability vectors generated by the first-stage classifier. The experimental validation tests compared the proposed method with a baseline image classification technique and four other related methods using six different CNN models and three different datasets of images depicting fine art paintings. The findings show clear benefits of the proposed approach in comparison with the other current techniques. A strong dependency of the classification outcomes on the type of CNN and the quality of the training data was observed. In addition, our findings indicate that for stylistic art analysis the best results are achieved when local patch-based analysis is combined with the holistic analysis of the entire image. Confusion between artistic styles was found to be consistent with the historical similarity between styles. Future research directions will aim to reduce the confusion between specific styles. Hierarchical structures of information-sharing, inter-dependent deep and shallow networks trained to differentiate between styles showing high level of similarity will be investigated.

REFERENCES

[1] L. Fichner-Rathus, *Understanding Art*, 9th ed. Belmont, CA, USA: Wadsworth, 2010, p. 560.

[2] T. E. Lombardi, "The classification of style in fine-art painting," Ph.D. dissertation, Dept. Doctor Prof. Stud. Comput., School Comput. Sci. Inf. Syst. Pace Univ., New York, NY, USA, 2005.

[3] (May 20, 2018). *The Art Story: Modern Art Movement Timeline*. [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 Trans. Appl. Perception*, vol. 7, no. 2, pp. 1–17, Feb. 2010. doi: 10.1145/1670671.1670672.

[5] R. S. Arora and A. Elgammal, "Towards automated classification of fine-art painting style: A comparative study," in *Proc. 21st Int. Conf. Pattern Recognit. (ICPR)*, Tsukuba, Japan, 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," *Mach. Vis. Appl.*, vol. 25, no. 6, pp. 1385–1397, Aug. 2014. doi: 10.1007/s00138-014-0621-6.

[7] S. Agarwal, H. Karnick, N. Pant, and U. Patel, "Genre and style based painting classification," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, Jun. 2015, pp. 588–594. doi: 10.1109/WACV.2015.84.

[8] Z. Falomir, L. Museros, I. Sanz, and L. Gonzalez-Abril, "Categorizing paintings in art styles based on qualitative color descriptors, quantitative global features and machine learning (QArt-Learn)," *Expert Syst. Appl.*, vol. 97, pp. 83–94, May 2018. doi: 10.1016/j.eswa.2017.11.056.

[9] E. Gultepe, T. E. Conturo, and M. Makrehchi, "Predicting and grouping digitized paintings by style using unsupervised feature learning," *J. Cultural Heritage*, vol. 31, pp. 13–23, May/Jun. 2018. doi: 10.1016/j.culher.2017.11.008.

[10] S. Karayev *et al.*, "Recognizing image style," in *Proc. Brit. Mach. Vis. Conf. (BMVC)*, Sep. 2014, pp. 1–20.

[11] Y. Bar, N. Levy, and L. Wolf, "Classification of artistic styles using binarized features derived from a deep neural network," in *Proc. Workshop at Eur. Conf. Comput. Vis.*, Cham, Switzerland: Springer, Mar. 2015, pp. 71–84. doi: 10.1007/978-3-319-16178-5_5.

[12] N. van Noord, E. Hendriks, and E. Postma, "Toward Discovery of the Artist's Style: Learning to recognize artists by their artworks," *IEEE Signal Process. Mag.*, vol. 32, no. 4, pp. 46–54, Jul. 2015. doi: 10.1109/MSP.2015.2406955.

[13] C. Hentschel, T. P. Wiradarma, and H. Sack, "Fine tuning CNNS with scarce training data—Adapting ImageNet to art epoch classification," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2016, pp. 3693–3697. doi: 10.1109/ICIP.2016.7533049.

[14] B. Saleh and A. Elgammal, "Large-scale classification of fine-art paintings: Learning the right metric on the right feature," *CoRR*, vol. abs/1505.00855, pp. 1–21, May 2015. [Online]. Available: https://arxiv.org/abs/1505.00855

[15] W.-T. Chu and Y.-L. Wu, "Image style classification based on learnt deep correlation features," *IEEE Trans. Multimedia*, vol. 20, no. 9, pp. 2491–2502, Sep. 2018. doi: 10.1109/TMM.2018.2801718W.

[16] F. Bianconi and R. Bello-Cerezo, "Evaluation of visual descriptors for painting categorisation," *IOP Conf. Ser., Mater. Sci. Eng.*, 2018, vol. 364, no. 1, Art. no. 012037. doi: 10.1088/1757-899X/364/1/012037.

[17] 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 *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Phoenix, AZ, USA, Sep. 2016, pp. 3703–3707. doi: 10.1109/ICIP.2016.7533051.

[18] A. Lecoutre, B. Negrevergne, and F. Yger, "Recognizing art style automatically in painting with deep learning," in *Proc. Asian Conf. Mach. Learn.* Gujarat, India: ACML, Nov. 2017, pp. 327–342.

[19] T. Sun, Y. Wang, J. Yang, and X. Hu, "Convolution neural networks with two pathways for image style recognition," *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4102–4113, Sep. 2017. doi: 10.1109/TIP.2017.2710631.

[20] E. Cetinic, T. Lipic, and S. Grgic, "Fine-tuning convolutional neural networks for fine art classification," *Expert Syst. Appl.*, vol. 114, pp. 107–118, Dec. 2018. doi: 10.1016/j.eswa.2018.07.026.

[21] A. Elgammal, B. Liu, D. Kim, M. Elhoseiny, and M. Mazzone, "The shape of art history in the eyes of the machine," in *Proc. 32nd AAAI Conf. Artif. Intell.*, New Orleans, LA, USA, Feb. 2018, pp. 1–9.

[22] G. Folego, O. Gomes, and A. Rocha, "From Impressionism to expressionism: Automatically identifying van Gogh's paintings," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Phoenix, AZ, USA, Sep. 2016, pp. 141–145. doi: 10.1109/ICIP.2016.7532335.

[23] N. van Noord and E. Postma, "Learning scale-variant and scale-invariant features for deep image classification," *Pattern Recognit.*, vol. 61, pp. 583–592, Jan. 2017. doi: 10.1016/j.patcog.2016.06.005.

[24] K. A. Jangtjik, T.-T. Ho, M.-C. Yeh, and K.-L. Hua, "A CNN-LSTM framework for authorship classification of paintings," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 2866–2870. doi: 10.1109/ICIP.2017.829680.

[25] S. Bianco, D. Mazzini, and R. Schettini, "Deep multibranch neural network for painting categorization," in *Proc. Int. Conf. Image Anal. Process. (ICIAP)* (Lecture Notes in Computer Science), vol. 10484, S. Battiato, G. Gallo, R. Schettini, and F. Stanco, Eds. Cham, Switzerland: Springer, Sep. 2017, pp. 414–423.

[26] C. Florea and F. Gieseke, "Artistic movement recognition by consensus of boosted SVM based experts," *J. Vis. Commun. Image Represent.*, vol. 56, pp. 220–233, Oct. 2018. doi: 10.1016/j.jvcir.2018.09.015.

[27] C. S. Rodriguez, M. Lech, and E. Pirogova, "Classification of style in fine-art paintings using transfer learning and weighted image patches," in *Proc. 12th Int. Conf. Signal Process. Commun. Syst. (ICSPCS)*, Dec. 2018, pp. 1–7. doi: 10.1109/ICSPCS.2018.8631731.

[28] M. N. Stolar, M. Lech, R. S. Bolia, and M. Skinner, "Towards autonomous machine reasoning: Multi-stage classification system with intermediate learning," in *Proc. 11th Int. Conf. Signal Process. Commun. Syst. (ICSPCS)*, Dec. 2017, pp. 1–6. doi: 10.1109/ICSPCS.2017.8270486.

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

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

[31] C. Florea, C. Toca, and F. Gieseke, "Artistic movement recognition by boosted fusion of color structure and topographic description," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2017, pp. 569–577. doi: 10.1109/WACV.2017.69.

[32] I. S. A. Krizhevsky and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, vol. 25, 2012, pp. 1097–1105.

[33] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *CoRR*, vol. abs/1409.1556, pp. 1–14, Sep. 2014. [online]. Available: https://arxiv.org/abs/1409.1556

[34] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. pattern Recognit.*, Boston, MA, USA, Jun. 2015, pp. 1–9. doi: 10.1109/CVPR.2015.7298594.

[35] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778. doi: 10.1109/CVPR.2016.9.

[36] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Las Vegas, NV, USA, Jun. 2016, pp. 2818–2826. doi: 10.1109/CVPR.2016.308.

[37] O. Russakovsky *et al.*, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015. doi: 10.1007/s11263-015-0816-y.

[38] W. Zhu, N. Zeng, and N. Wang, "Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations," in *NESUG: Proceedings: Health Care Life Sciences*, Baltimore, MD, USA: 2010, pp. 19–67.

[39] D. J. DeWitt, R. M. Larmann, and M. K. Shields, *Gateways to Art: Understanding the Visual Arts*, 2nd ed. New York, NY, USA: Thames & Hudson, 2015, pp. 465–547.

CATHERINE SANDOVAL received the B.S. degree in electronic engineering from Pontificia Javeriana University, Colombia, in 2001, and the M.S. degree in electronic engineering from RMIT University, Australia, in 2015, where she is currently pursuing the Ph.D. in electrical and electronic engineering. Her research interests include machine learning and image processing.

ELENA PIROGOVA received the B.Eng. degree (Hons.) in chemical engineering from National Technical University, Ukraine, in 1991, and the Ph.D. degree in biomedical engineering from Monash University, Australia, in 2002. She is currently an Associate Professor in biomedical engineering with the School of Engineering, RMIT University, Australia. Her research interests include biomedical electronics and instrumentation, bioelectromagnetics, and protein modeling.

MARGARET LECH received the M.S. degree in physics from Maria Curie-Skłodowska University, Poland, the M.S. degree in biomedical engineering from the Warsaw University of Technology, Poland, and the Ph.D. degree in electrical engineering from the University of Melbourne, Australia. She is currently an Associate Professor in signal processing and artificial intelligence with the School of Engineering, RMIT University, Australia. Her research interests include psychoacoustic, speech and image processing, system modeling, and optimization.