CONTENT-BASED IMAGE RETRIEVAL FOR PAINTING STYLE WITH CONVOLUTIONAL NEURAL NETWORK

(Date received: 19.01.2022/Date accepted: 07.04.2022)

W S Tan[1], W Y Chin[2*], K Y Lim[3]

*[1,2] Department of Mathematical and Data Science, Tunku Abdul Rahman University College,
Kampus Utama, Jalan Genting Kelang, 53300 Kuala Lumpur,
Wilayah Persekutuan Kuala Lumpur, Malaysia*

*[3] Department of Computing and Information Technology, Tunku Abdul Rahman University College,
Kampus Utama, Jalan Genting Kelang, 53300 Kuala Lumpur,
Wilayah Persekutuan Kuala Lumpur, Malaysia*

**Corresponding author: chinwy@tarc.edu.my*

ABSTRACT

With the advancement of digital paintings in online collection platform, new image processing algorithms are required to manage digital paintings saved on database. Image retrieval has been one of the most difficult disciplines in digital image processing because it requires scanning a large database for images that are comparable to the query image. It is commonly known that retrieval performance is largely influenced by feature representations and similarity measures. Deep Learning has recently advanced significantly, and deep features based on deep learning have been widely used because it has been demonstrated that the features have great generalisation. In this paper, a convolutional neural network (CNN) is utilised to extract deep and high-level features from the paintings. Next, the features were used for similarity measure between the query image and database images; subsequently, similar images are ranked by the distance between both pair features. Our experiments show that this strategy significantly improves the performance of content-based image retrieval for the style retrieval task of painting. Besides, the extracted feature to retrieve the right classes from the query image has achieved over 61% accuracy which beat the current-state-of-art results. However, the result can be further improved in future research by leveraging CNN representations visualisation approaches for a better understanding of how CNN extract features from paintings.

Keywords: Content-based Image Retrieval, Deep Learning, Convolutional Neural Network

1.0 INTRODUCTION

With the continuous expanding due to advancement in digital imaging and internet usage, online artwork collection such as WikiArt, Artsper and Mutual Art have been one of the fastest growing databases. As a result, existing algorithms are incapable of managing these large databases, necessitating the use of robust and quick approaches. Among the several domains of image processing, image retrieval has been always one of the popular approaches in recent years. Image retrieval, which involves scanning a large database for photos that are similar to the query image, was first developed in 1970 by text-based image retrieval (TBIR), in which the system accepts a query word from the user and searches for images that include the text (Rui *et al.,* 1999). However, the concept of an image is much more complex than a few words, and it often turns out not to be so effective. This is due to the subjectivity of the task compared to the meaning of its semantic content. Therefore, content-based image retrieval (CBIR) was invented in 1990. The CBIR has been applied in numerous disciplines, including medical imaging (Campbell, 1994), video processing (Karimi and Bashiri, 2011), crime prevention and other areas that need image recognition (Hwang and Lee, 2012 and Jabalemali *et al.,* 2012).

Feature extraction is a critical operation in signal processing, image, video, and speech processing (Zade *et al.,* 2014 and Pasandideh *et al.,* 2016). It is also one of the critical components of any image retrieval system. The features of an image can be described in two different categories: At the digital level, low-level features mainly are colour-based, texture, and shape features. At the semantic level, the image can be interpreted as having at least one meaning. Unfortunately, paintings are defined digitally in today's information system, while users are more interested in their semantic concept, rather than visually similar. The semantic gap between low-level features and human concept is huge, and it is currently difficult to identify correspondences between the digital painting and semantic levels. Although it may be able to extract increasingly complicated low-level features from images, the size of the feature vector will grow, and the retrieval speed will slow as the calculation time increases. As a result, it is necessary to extract appropriate abstracted features in order to maximize retrieval precision while minimizing retrieval time. Thus, deep learning is one of the ways that has been shown to reduce the semantic gap between low-level features and human perception (Zade *et al.,* 2016) and achieve a good efficiency of image retrieval.

It is commonly known that CBIR is a system that retrieves images from an image database using visual contents. Because it can successfully address the challenges written above, this system has now become vital for image retrieval. In CBIR, visual contents are the features extracted from digital images, and its performance is strongly influenced by the features extracted and similarity measures. Due to these reasons, CNN as a successful subfield of deep learning was used to extract deep and appropriate features for CBIR to process for image retrieval in order to improve the performance of CBIR. In addition, we should not overlook the reality where the research for CBIR has been thriving and particularly strong over the past decades such as CBIR with handcrafted features (Hiremath and Pujari, 2007; Alhassan and Alfaki, 2017 and He *at al.,* 2018). However, the amount of attention obtained in the search for paintings in CNN image retrieval is minimal because there is no specific mechanism for visual art interpretation. One of the reasons could be that the visual likeness of paintings can be highly variable, with broad criteria in judging the similarity ranging from a little object, texture, brushstroke, to the entire configuration of the painting itself (Seguin, 2009). To be more explicit, developing a general content-based image retrieval system is easier than developing a domain-specific application, which necessitates domain knowledge. In short, developing a specific domain image retrieval application is difficult yet rewarding research.

In this paper, the work was motivated by the advancement and the efficiency of the features extraction in CNN. Handcrafted features methods such as SIFT (Lowe,1999), SURF (bay *et al.,* 2006) and GIST (Oliva and Torralba, 2016) were popular in CBIR, however, we wish to understand if we can profit more fully and flawlessly from deep CNN to increase the efficiency for features extraction in CBIR process. Moreover, creative artwork, such as fine art painting, has attracted much attention from various researchers to seek potential applications. Undoubtedly, several researchers have published numerous publications regarding paintings' characteristic recognition and retrieval task. For instance, Cetinic *et al.,* (2018) introduced an approach that are similar to Tan *et al*. (2016) for addressing the fine art classification with fine-tuning CNN, where the model can classify painting's characteristic and also explored on the applicability of the model for retrieving similar paintings based on the query image in either style or content. In the following two years, Cetinic *et al*. (2020) presented another work which used CNN for learning features that are relevant for understanding properties of artistic styles described by Heinrich Wolfflin. Their evaluations suggested that the models learn to discriminate meaningful features that correspond to the visual characteristic of the artistic concepts. Two of these papers indicate that CNN could perform very well and able to measure the artistic style or content in paintings with proper settings. Gontheir *et al*. (2021) recently did a similar experiment. The authors employed approaches to show network internal representations in order to offer information about what a network learns from aesthetic imagery. They also shown that a twofold fine-tuning using a medium-sized artistic dataset may improve the classification on smaller datasets, even when the classification task changes. Besides, Chen *et al*. (2019) expanded on previous research on the use of CNNs for style categorization by observing that various layers in existing deep learning models exhibit varied feature responses for the same input picture. To fully use the input from various levels, the authors presented an adaptive cross-layer model that incorporates responses from both lower and upper layers to capture style. Sandoval *et al*. (2019) contributed by proposing a two-stage picture classification strategy to enhance style categorization. The approach divides the input image into patches and utilises a CNN model to categorise the artistic style for each patch in the first step. The CNN's probability values are then combined into a single feature vector, which is sent as an individual input to a shallow neural network model, which conducts the final classification. The suggested technique is based on the idea that separate patches act as independent evaluators for different parts of the same image, with the final model combining those evaluations to make the ultimate judgement. As is typical in this research, there was some misunderstanding between historically related styles. In short, we conclude that differentiating visual styles remains a difficult task. However, it can be observed that with the rapid development of deep learning framework, without using conventional CBIR methods such as handcrafted features, but instead build a domain-specific CBIR could get a better classification of painting's style and also better accuracy in similar image retrieval-based models fine-tuned for style recognition. It closes various gaps in prior methodologies and pave a new direction for domain specific CBIR which provides valuable additional information into classifier decision-making processes.

The rest of this paper is organized as follows. Section 2 presents the proposed approach and then evaluate the solution through experiments and the application of using proposed approach for similarity measure in Section 3. Finally, Section 4 presents the conclusion and future direction of this work.

2.0 PROPOSED METHODOLOGY

2.1 Convolutional Neural Network Model Configuration

The overall structure of CNN was a modified version of VGG16 (Simonyan and Zisserman, 2014) where it has five convolutional layers, three max-pooling layers, and followed by Global Average Pooling layer (GAP) (Lin *et al.,* 2013). GAP is a procedure that computes the average output of each feature map, decreasing the total number of parameters in the model, and preparing the model for the final classification layer. The intention of replacing fully connected layers from VGG16 with GAP was to reduce the parameters and lower the risk of overfitting to the training data set. Each convolutional layer yields 64, 128, 256, 512, and 512 feature maps, respectively. Filter size of 3 x 3 was use throughout the whole net, which are convolved with the input with only stride 1. Then, the pooling layers with a size of 2 x 2 and step 2 for down sampling. The activation function will be rectification linear unit (RELU) in all weight layers except the last output layer, which will utilize the softmax function as activation and operate as a multi-class classifier to predict the painting categorization as shown in Figure 1. However, in order to measure the similarity of the paintings, the output layer is removed after training and the features will be extracted from the GAP layer.

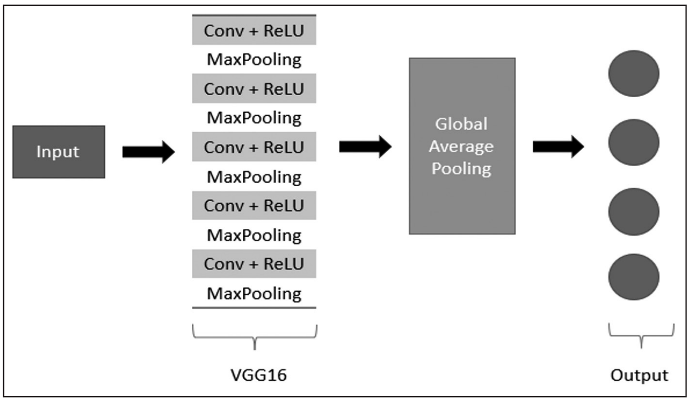


Figure 1: Architecture of CNN

2.2 Dataset

The first data source is Imagenet Dataset which used in Imagenet Large-Scale Visual Recognition Challenge (ILSVRC), in the fine-tuning process, we pre-trained the network using this dataset. It consists of 1.2 million object images with roughly 1000 images in each of the 1000 categories. In all, there are about 1.2million training images, 50,000 validation images and 150,000 testing images. The second data source WikiArt (Saleh and Elgammal, 2015), which is now the largest online available collection of digital paintings. The WikiArt paintings dataset contains over 80,000 fine-art paintings by more than 1,000 artists, it includes artworks from a wide period of time, ranging from the fifteenth century to modern times, and its particular focus on the 19th and 20th centuries, as well as contemporary art. The collection contains 27 different art styles and 45 different genres. WikiArt is also a well-organized collection that incorporates a diverse variety of metadata such as artist name, style, genre, nationality, and so on. Meanwhile, with a total of around 83,000 of samples in the dataset was split into training, validation, and testing with a ratio of 70%, 15% and 15% respectively.

2.3 Experimental Set-up

2.3.1 Input Layer and Preprocessing

The input data with a dimension of 224×224×3 where 224×224 is the width and height of the image and 3 is the number of channels which is RGB colour image. The preprocessing of input image will be subtracting the mean value of RGB over the Wikiart dataset for each pixel. No data augmentation was applied.

2.3.2 Training Details

The model is trained using stochastic gradient descent (SGD) with a batch size of 64 samples. The rest of the parameters are set as momentum of 0.9, decay rate of 0.00001 and the initial learning rate of 0.0001. The weight initialization was taken from the pre-trained VGG16 model, where it was trained with over 1.2 million images for object recognition. Since object recognition and painting's style classification have the same data consistency and share the same data type. The learnt features from object recognition can be easily transfer to the new domain images. This could help in reducing the computational cost for retraining from scratch.

2.3.3 Method for Similarity Measure

After the training process, the trained model with the painting dataset will be used to extract the features from each image. The softmax last layer is removed, and the GAP output feature will be stored to measure the similarity between images (refer to Table 1). In particular, features extracted from GAP is 512 feature vectors. Based on retrieved feature vectors, the distance between feature vectors was calculated using the k-NN brute-force approach, and Euclidean distance measure is utilised as a distance metric to calculate the painting similarity. The general formulation for points given by Cartesian coordinates in n - dimensional Euclidean space is as follows:

$$d(p,q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_n - q_n)^2} \quad (1)$$

Table 1: Related parameters of Convolutional Neural Network

Type	Size/Stride	Output Size
Conv1	3×3/1	64 × 224 × 224
MaxPool1	2×2/2	64 × 112 × 112
Conv2	3×3/1	128 × 112 ×112
MaxPool2	2×2/2	128 × 56 × 56
Conv3	3×3/1	256 × 56 × 56
MaxPool3	2×2/2	256 × 28 × 28
Conv4	3×3/1	512 × 28 × 28
MaxPool4	2×2/2	512 × 14 × 14
Conv5	3×3/1	512 × 14 × 14
MaxPool5	2×2/2	512 × 7 × 7
GlobalAvrPool	-	512
Softmax	-	27

2.4 CNN Architecture Details

2.4.1 Global Average Pooling

Global average pooling (GAP) (Lin *et al.,* 2013) is a pooling operation designed to replace the conventional architecture of CNN that uses fully connected layers as the standard configuration. Fully connected layers usually consist of too many parameters, and this has led to slow training speed of the network. By replacing a fully connected layer with GAP, it not only can eliminate the use of parameter, but also is able to avoid overfitting. Instead of adding fully connected layers on top of the feature maps as in conventional CNN model in GAP, the average of the entire pixels of each feature map is taken, and the resulting vector is fed directly into the softmax layer for classification.

2.4.2 VGG16 Pre-Training Model

As shown in Figure 1, the first five convolutional layers which is the VGG16 (Simonyan and Zisserman, 2014) pretrained model that acts as the base model in the proposed cascade CNN architecture. The unique characteristic of VGG16 is that instead of having a large number ofconvolution filters, the authors focused on having convolution layer of 3 x 3 filter with stride 1 and followed by max pooling layer of 2 x 2 filter with stride 2. This base model is responsible on learning the low-level features that can better adapt to various problems and high-level features for domain specific problems. Also, pretrained networks for VGG are available freely on the internet, the weights can be downloaded and used for transfer learningwhere it can shift the learnt features from one domain (object recognition) to the new domain (paintings classification).

3.0 EXPERIMENTAL RESULT AND DISCUSSION

The result in Table 2 shows the proposed fine-tuning model has a competitive result as compared to the current state-of-art methods (Tan *et al.,* 2016 & Cetinic *et al*. 2018) without any additional mechanisms. By comparing the current state-of-art result with the proposed model with further fine-tuning, the proposed model implementation with retraining achieved 55.6% of accuracy for classifying painting's style. In addition, proposed model has lesser parameter with approximate of 44-millions which resulted in reducing the computational cost compared to Tan *et al*. (2016) about 61-millions parameter by just replacing the fully connected layer with global average pooling layer. As a result, it was conjecturing that the proposed model can be further improved with different classifying approach and applying different data pre-processing can lead to a huge boost of performance as shown in the work of (Lecoutre *et al.,* 2017). In section 3.2, different approach of classifying painting's style was further explored, and it was able to beat the current-state-of art.

Table 2: Comparison of CNN fine-tuned results on Wikiarts dataset with model pretrained with ImageNet Dataset

Reference	Methods	Number of classes	Accuracy (%)
Our model	Proposed Model (VGGNet)	27	55.6
Tan *et al.* (2016)	CNN fine tuning (AlexNet)	27	54.5
Cetinic *et al.* (2018)	CNN fine-tuning (CaffeNet) – Hybrid model (with best fine-tune scenario)	27	57

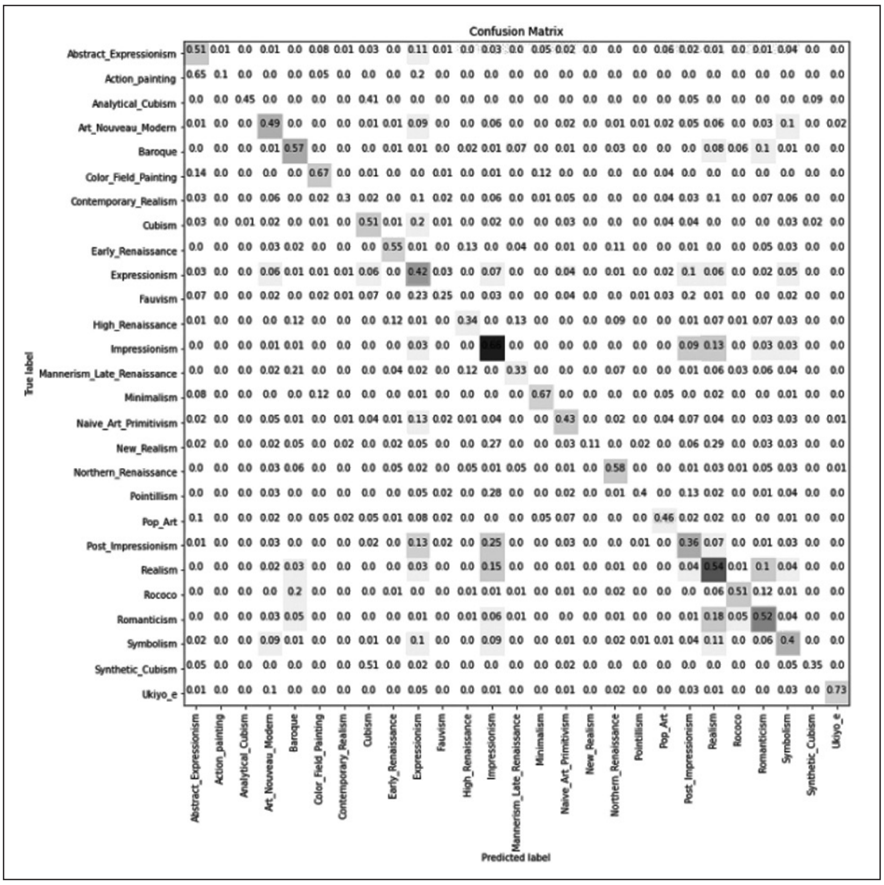


Figure 2: Confusion Matrix for style classification

Following the style classification result and comparison with the existing state-of-the-art, a further examination of the style classification was carried out by looking into the per-class classification performance that merited attention. Figure 2 depicts the WikiArt dataset's confusion matrix for each classification class. It can be observed that there are several classes perform relatively better as compared to other classes due to their distinct visual appearance, such as Ukiyo-e (73%), Minimalism (67%) and Colour Field Painting (67%). Ukiyo-e shows in general the best result as it is a type of art that flourished in Japan that has a very special characteristic. Secondly, the proposed CNN can distinctly classify the Impressionism with 66% accuracy from the other styles. The high accuracy might be due to the high number of training data with approximately 13-thousands images in the Impressionism category. This is consistent with the finding of Goodfellow *et al*. (2016) that a neural network requires around 5,000 labels per class to achieve human-level classification performance. Action paintings (10%) was wrongly classified as Abstract Expressioniosm (51%), this was because action painting was evolved in the 1940s and 1950s during a time of unrest following World War II which can be seen as both of the styles are belong to the same groups. For the misclassification on the dataset, normally two groups of styles share the common conceptual ground. For example, 25% of Post Impressionism was wrongly classified as its elder brother, Impressionism. 51% of Synthetic Cubism was classified as Cubism which Synthetic Cubism was known as later phase of Cubism started from 1908-1912. New realism was wrongly classified nearly 30% to Realism which come from the same root as it was new form of realism developed at the beginning of 20th century. This is similar to Rococo (51%) and Baroque (57%) as these two styles are historically related. The misclassification of style also explained that it have a poor performance from the classification task as cascade model generally struggle on differentiating historically related styles. It can be concluded that artwork style is not only associated with common visual properties but contextually dependant concept.

3.1 Measure Similarity with Proposed Model

CNN models fine-tuned for style identification were used to retrieve images with similar style or content. As shown in Figure 3, each query image with four of the most similar images were retrieved. We can see from these examples that the suggested CNN fine-tuned model for style recognition focuses more on style attributes like brushwork or amount of detail. Despite some incorrectly obtained class image, it can nevertheless retrieve similar painting in terms of content by including certain items and similar compositions. In addition to the result above, we conjecture those further improvements in style-specific classification performance will result in greater distinguishability between style-similar images. Therefore, in order to validate this hypothesis, further investigation of the model features is performed to study the effect of before and after fine-tuning by transfer learning in Section 3.2. It is well known that the ImageNet dataset was used to train various pre-trained models (VGG also pre-trained with the ImageNet dataset). As a result, in most cases, they provide an excellent starting point for similarity computations. However, if these models were adjusted to suit the specific problem, they would find similar images even more accurately.

3.2 Comparison of CBIR Performance Before and After Fine-Tuning with Specific Domain Knowledge

Similar process mentioned from Section 2.3.3 which extract the features and retrieve the similar images, but from this experiment, we identify the worst-performing categories, fine-tune, and then see how the accuracy change. For every image in the WikiArt dataset, it uses the brute-force approach to determine the closest neighbours for each image in the

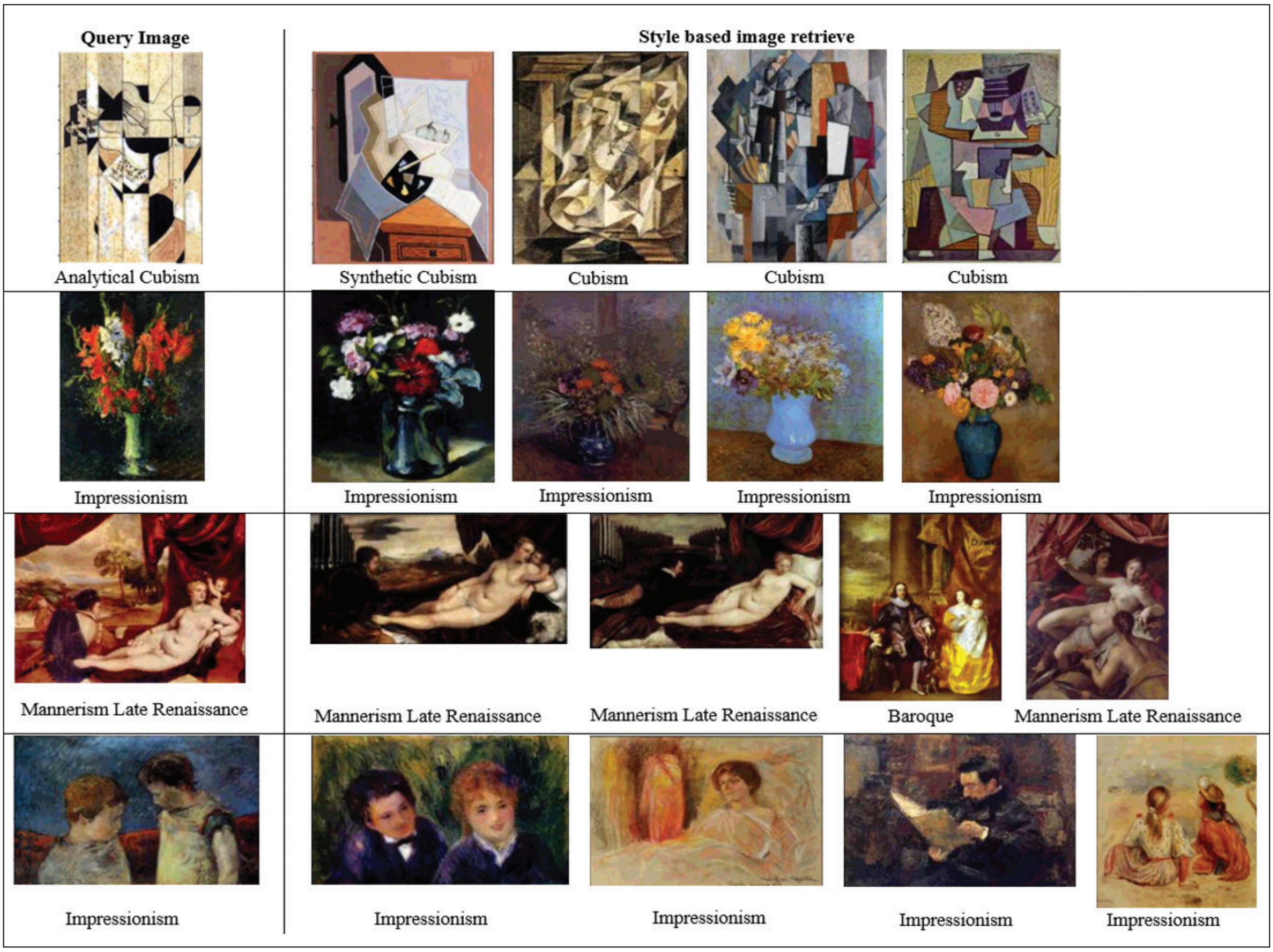


Figure 3: Examples of paintings with style label retrieved as most similar to the query image when using the fine-tuned proposed CNN model as feature extractors

dataset and then returns the top-10 classes with the lowest accuracy. The analysis would provide an overview on how fine adjustment affects the results.

Table 3: Top 10 Lowest-Accuracy Classes

No.	Class (Before fine-tuning)	Retrieval accuracy (%)	Class (After fine-tuning)	Retrieval accuracy (%)
1	New Realism	11.58	Fauvism	39.92
2	Fauvism	21.6	New Realism	40.56
3	Mannerism Late Renaissance	23.8	High Renaissance	48.12
4	High Renaissance	24.89	Mannerism Late Renaissance	50.4
5	Pointillism	25.51	Action Painting	50.82
6	Rococo	29.07	Post Impressionism	52.57
7	Post Impressionism	30.66	Expressionism	52.81
8	Early Renaissance	32.45	Synthetic Cubism	54.1
9	Action Painting	32.84	Contemporary Realism	55.48
10	Baroque	34.16	Symbolism	55.84
Average Correct Prediction Accuracy (%)		**39.2**		**61.33**

With the extracted feature vectors before fine-tuning model, it can be observed from Table 3 that the retrieval accuracy is quite poor as the lowest accuracy was only 11.58% while the highest accuracy in the Top-10 least accuracy classes was at 34.16%. This result shows that the model suffered from discriminating the correct classes when retrieving similar images. Using these feature vectors in applications such as image retrieval systems may be a bad idea because obtaining a clean plane of separation between classes may be difficult. It is hardly surprising that the retrieval accuracy performed so poorly in this nearest-neighbour-based categorization task due to the learned features being based on the natural images. In contrast, after retraining with domain dataset, the outcome is intriguing; the Top-10 least accurate classes have some changes, and retrieval accuracy has skyrocketed. Previously, the feature vectors from the model before fine-tuning achieved an overall correct prediction accuracy of only 39.2%. The new feature vectors after fine-tuning deliver a whopping 61.33% accuracy.

From Table 4, the prior works classification accuracy was again act as a benchmark. As we compared the result with prior work, it shows that with our approach could outperforms the current state-of-art reported for the WikiArt dataset. In Tan *et al*. (2016) paper, the authors achieved the best result with 54.5% by fine-tuning the Alexnet network which also pre-trained with ImageNet dataset. On the other hand, Cetinic *et al*. (2018) achieved an even better result with 57% by implementing different domain-specific weight initialization and different training settings. However, with our approach where basically extract the feature vectors from fine-tuned model and further classified with nearest-neighbour approach led to a better performance in overall. To summarise the discussion, the hypothesis expressed in the previous section was valid in which additional increases in style-specific classification performance will result in higher distinguishability across style-similar images. As a result, we may conclude that domain-specific initialization and task-specific fine-tuning can have a considerable impact on obtaining CBIR performance.

Table 4: Comparison of results to prior works on the style classification task with new feature extraction method

References	Methods	Accuracy (%)
Proposed model	Proposed Model (VGGNet)	61.33
Tan *et al.* (2016)	CNN fine-tuning (AlexNet)	54.5
Cetinic *et al.* (2018)	CNN fine-tuning (CaffeNet) – Hybrid model (with best fine-tune scenario)	57

4.0 CONCLUSIONS

In this work, we presented a study using CNN as a feature extractor for measuring similarity between painting's styles. We successfully applied the extracted feature to retrieve the right classes from the query image that achieve over 61% accuracy. This improvement is mainly due to the idea of transfer learning and the importance of retraining. As suggested by our experiments, CNN retraining is required to build a specific domain CBIR that can outperform general CBIR and is suitable for measuring the similarity of painting and feasible for use in online art galleries. However, the inclusion of a larger painting dataset should allow the model to learn more from scratch rather than via transfer learning. As a result, we intend to expand the dataset so that we may fully retrain the deep learning models. We also plan to deepen our multidisciplinary collaboration in the future by doing research on the importance of the findings to specific art history study areas. Investigate how a deep neural network may be used to extract high-level and semantically significant components that can be utilised to discover new knowledge patterns and meaningful connections between individual artworks. Increase the knowledge and interpretability of deep learning models, on the other hand, by leveraging CNN representations visualisation approaches such as activation maximisation, saliency maps, and class activation maps, as well as other visualisation techniques for a better understanding of how CNN extract features from paintings

5.0 ACKNOWLEDGMENTS

This project is funded by CMG Holdings Sdn Bhd, we would also like to express our gratitude towards our collaborator Cashierbook for their useful suggestions and special thanks to the reviewers for their feedback in this paper. ■

REFERENCES

[1] Alhassan, A.K. and Alfaki, A.A., 2017, January. Color and texture fusion-based method for content-based image retrieval. In 2017 International Conference on Communication, Control, Computing, and Electronics Engineering (ICCCCEE) (pp. 1-6). IEEE.

[2] Bay, H., Tuytelaars, T., and Van Gool, L., 2006, May. Surf: Speeded up robust features. In European conference on computer vision (pp. 404-417). Springer, Berlin, Heidelberg.

[3] Campbell, J.F., 1994. Integer programming formulations of discrete hub location problems. European Journal of Operational Research, 72(2), pp.387-405.

[4] Cetinic, E., Lipic, T. and Grgic, S., 2018. Fine-tuning convolutional neural networks for fine art classification. Expert Systems with Applications, 114, pp.107-118.

[5] Cetinic, E., Lipic, T. and Grgic, S., 2020. Learning the Principles of Art History with convolutional neural networks. Pattern Recognition Letters, 129, pp.56-62.

[6] Chen, L. and Yang, J., 2019, October. Recognizing the style of visual arts via adaptive cross-layer correlation. In Proceedings of the 27th ACM international conference on multimedia (pp. 2459-2467).

[7] Gonthier, N., Gousseau, Y. and Ladjal, S., 2021, January. An analysis of the transfer learning of convolutional neural networks for artistic images. In International Conference on Pattern Recognition (pp. 546-561). Springer, Cham.

[8] Goodfellow, I., Bengio, Y. and Courville, A., 2016. Deep learning. MIT press.

[9] He, T., Wei, Y., Liu, Z., Qing, G., and Zhang, D., 2018, January. Content-based image retrieval method based on SIFT feature. In 2018 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS) (pp. 649-652). IEEE.

[10] Hiremath, P.S. and Pujari, J., 2007. Content Based Image Retrieval based on Color, Texture, and Shape features using Image and its complement. International Journal of Computer Science and Security, 1(4), pp.25-35.

[11] Hwang, YH, and Lee, YH, 2012. Uncapacitated single allocation p-hub maximal covering problem. Computers & Industrial Engineering, 63(2), pp.382-389.

[12] Jabalameli, M.S., Barzinpour, F., Saboury, A., and Ghaffari-Nasab, N., 2012. A simulated annealing-based heuristic for the single allocation maximal covering hub location problem. International Journal of Metaheuristics, 2(1), pp.15-37.

[13] Karimi, H. and Bashiri, M., 2011. The hub covers location problems with different types of coverage. Scientia Iranica, 18(6), pp.1571-1578.

[14] Lecoutre, A., Negrevergne, B. and Yger, F., 2017, November. Recognizing art style automatically in painting with deep learning. In Asian conference on machine learning (pp. 327-342). PMLR.

[15] Lin, M., Chen, Q. and Yan, S., 2013. Network in network. arXiv preprint arXiv:1312.4400.

[16] Lowe, D.G., 1999, September. Object recognition from local scale-invariant features. In Proceedings of the seventh IEEE international conference on computer vision (Vol. 2, pp. 1150-1157). IEEE.

[17] Oliva, A., and Torralba, A., 2001. Modeling the shape of the scene: A holistic representation of the spatial envelope. International Journal of Computer Vision, 42(3), pp.145-175.

[18] Pasandideh, S.H.R., Niaki, S.T.A. and Sheikhi, M., 2016. A bi-objective hub maximal covering location problem considering time-dependent reliability and the second type of coverage. International Journal of Management Science and Engineering Management, 11(4), pp.195-202.

[19] Rui, Y., Huang, T.S. and Chang, S.F., 1999. Image retrieval: Past, present, and future. Journal of Visual Communication and Image Representation, 10(1), pp.1-23.

[20] Saleh, B. and Elgammal, A., 2015. Large-scale classification of fine art paintings: Learning the right metric on the right feature. arXiv preprint arXiv:1505.00855.

[21] Sandoval, C., Pirogova, E. and Lech, M., 2019. Two-stage deep learning approach to the classification of fine-art paintings. IEEE Access, 7, pp.41770-41781.

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

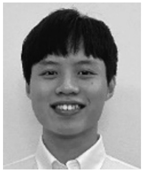
[23] Tan, W.R., Chan, C.S., Aguirre, H.E. and Tanaka, K., 2016, September. Ceci n'est pas une pipe: A deep convolutional network for fine-art paintings classification. In 2016 IEEE international conference on image processing (ICIP) (pp. 3703-3707). IEEE.

[24] Van der Maaten, L. and Hinton, G., 2008. Visualizing the data using t-SNE. Journal of Machine Learning Research, 9(11).

[25] Zade, A.E., Sadegheih, A., and Lotfi, M.M., 2014. A modified NSGA-II solution for a new multi-objective hub maximal covering problem under uncertain shipments. Journal of Industrial Engineering International, 10(4), pp.185-197.

[26] Zade, A.E., Sadegheih, A., and Lotfi, M.M., 2016. Fuzzy multi-objective linear programming for a stochastic hub maximal covering problem with uncertain shipments. International Journal of Industrial and Systems Engineering, 23(4), pp.482-499.

PROFILES



TAN WEI SHENG received the B. Sc degree in Management Mathematics with Computing from Tunku Abdul Rahman University College, Malaysia in 2019. Presently pursuing his master's degree in Mathematical Sciences at Tunku Abdul Rahman University College, Malaysia. His research interest includes Statistics, Machine Learning, Deep Learning, and other areas.
Email address: tanws-wa15@student.tarc.edu.my

DR CHIN WAN YOKE received B.Sc. and Ph.D. degrees from Universiti Teknologi Malaysia in the year 2009 and 2016, respectively. She is currently a postgraduate programme leader at Tunku Abdul Rahman University College and is actively involved in the activities organized by the Faculty of Computing and Information Technology, Center for Data Science and Analytics, and Centre for Postgraduate Studies and Research. Her main research focus is mathematics and statistics. She is a multidisciplinary researcher who is involved in several projects including data analytics, predictive modeling, optimization, statistical hydrology, painting styles classification, and social media engagement.
Email address: chinwy@tarc.edu.my

DR LIM KHAI YIN received her Master of Science (Artificial Intelligence) and PhD (Visual Analytics) in 2009 and 2017, respectively. Currently, she is a senior lecturer in TAR UC. Her research interests include computer vision, medical image processing, machine learning, and data analytics. She has also published papers both in journal and conferences relating to artificial intelligence (AI) and computer vision, specifically in fuzzy theory, neural networks, and medical imaging during her postgraduate study. She has experiences in guiding final year project (FYP) students on the topics that involves AI and computer vision. Besides that, she has been involving in some industry projects, mainly in the topics of classifying painting styles and engagement prediction in social media.
Email address: limky@tarc.edu.my