\documentclass[10pt,twocolumn,letterpaper]{article}

\usepackage{cvpr}

\usepackage{times}

\usepackage{epsfig}

\usepackage{graphicx}

\usepackage{amsmath}

\usepackage{amssymb}

\usepackage{subcaption}

\usepackage{filecontents}

\usepackage{multirow,tabularx,ragged2e,booktabs,caption}

\usepackage[pagebackref=true,breaklinks=true,letterpaper=true,colorlinks,bookmarks=false]{hyperref}

\setcounter{secnumdepth}{5}

\def\cvprPaperID{\*\*\*\*} % \*\*\* Enter the CVPR Paper ID here

\def\httilde{\mbox{\tt\raisebox{-.5ex}{\symbol{126}}}}

% Pages are numbered in submission mode, and unnumbered in camera-ready

\ifcvprfinal\pagestyle{empty}\fi

\hypersetup{draft}

\begin{document}

%%%%%%%%% TITLE

\title{Transferring and Comparing Elements of Artistic Style}

\author{Shaoyan Pan\\

\\

Institution1 address\\

{\tt\small shaoyank@gmail.com}

% For a paper whose authors are all at the same institution,

% omit the following lines up until the closing ``}''.

% Additional authors and addresses can be added with ``\and'',

% just like the second author.

% To save space, use either the email address or home page, not both

\and

Yahui Liang\\

Institution2\\

First line of institution2 address\\

{\tt\small liangy@msoe.edu}

\and

Josiah Yoder\\

Institution2\\

First line of institution2 address\\

{\tt\small yoder@msoe.edu}

}

\maketitle

%\thispagestyle{empty}

%%%%%%%%% ABSTRACT

% \begin{abstract}

% The paper introduces a new approach to identify the artistic genre and the corresponding style (visual appearance) of a painting. Recent research applied convolutional neural networks (CNN) [cite] to painting style recognition and obtained results

% in terms of not only accuracy [cite], but also a deeper understanding of the association between style and painting image.

% Inspired by these works, we ask: What is the nature of style in painting and can we identify compositions

% of style. Based on the assumption that the properties of style are reflected by the

% brush stroke, color, light, and composition of image [cite? got this from Google running document],

% this paper introduces a new approach, feeding one of the most characteristic properties,

% the enclosed brush stroke,

% of image into a two-pathway convolutional neural network (cite), to recognize different styles of painting. We firstly summarize some traditional algorithms to extract brush stroke of painting. Then we use a two pathway CNN, which separates the predictions of the object content and the texture content of the image and leverage the two types of predictions to obtain the final recognition. We discuss how to adapt the texture pathway to use brush strokes instead of full images for its input. By analyzing the performance of this new two Pathway Convolutional Neural Network (TPCNN), we can evaluate the importance of brush strokes to artistic style.

% \end{abstract}

% % %%%%%%%%% BODY TEXT

% % \section{Introduction}

% % Please follow the steps outlined below when submitting your manuscript to

% % the IEEE Computer Society Press. This style guide now has several

% % important modifications (for example, you are no longer warned against the

% % use of sticky tape to attach your artwork to the paper), so all authors

% % should read this new version.

\section{Introduction}

Artistic style transfer networks have demonstrated the ability to generate compelling images that mimic the genre of existing artwork\footnote{cite gatys, that paper Stephen Schnelle recommended, the other generative paper, etc.}. These two-pathway neural networks (2PNNs) have also demonstrated improved accuracy at style classification tasks [cite]. Given these successes, and the human (if occasionally bizarre) feeling in images generated by neural networks in general \cite{2015arXiv150806576G}, it seems reasonable to hope that artificial neural networks (ANNs) could provide insight into human neural cognition and artistic critique.

A 2PNN includes two pathways: a convolutional neural network (CNN) \cite{2013arXiv1311.3715K} such as VGG-19 [cite VGG-19] in parallel with a texture pathway (Figure \ref{fig:2pnn}). The texture pathway captures the frequency of excitation of the various channels of the feature-maps produced as the output of each layer of the convolutional network [cite].

Style transfer networks take two images as input: A style image and a content image. A work of art produced by one of these networks preserves the major objects and layout of the content image, while producing the artistic effect of the style image. This is done by manipulating the content image such that the output the object pathway is preserved, but the output of the style pathway is changed to match its output when the style image is used as the input. Because the style pathway preserves only the distribution of the feature-maps, and not the pixel positions within the maps, an image with similar color and texture is produced, even at a large scale such as reproducing a star from Van Gogh's starry night. \cite{2015arXiv150806576G} In other words, style transfer networks transfer style by producing an image with the same neural features at every level in the network, but with those features excited in different parts of the image.

\begin{figure}[h]

\begin{center}

\includegraphics[width=0.8\linewidth,height=0.5\linewidth]{CNN\_flow.png}

\caption{Two pathway neural network (2PNN) including both a texture pathway and a CNN pathway.}

\label{fig:2pnn}

\end{center}

\end{figure}

The same networks used for style transfer can also be used to classify the style, with some modification [cite]. Instead of adapting the input image to produce certain outputs from the style and object pathways, fully-connected and softmax layers are added at the end of each pathway and the weights of the network are adapted to predict the style of the input artwork. Including the style pathway in parallel with the standard CNN object pathway improves the accuracy of the prediction by a few percent, perhaps because the object pathway automatically generalizes aspects of style found throughout the image [cite].

% TODO: How does fixing the object pathway weights affect the results?

There are many questions that remain about the internal workings of 2PNNs. Why do the networks produce images that are so visually appealing? Can the performance of classification tasks be improved to the rate of human experts (as suggested by the success of CNNs for object classification [cite])? Do current networks mimic the decision process of human experts or provide a complementary approach? As a step toward answering these questions, we explore the relationship between the formal elements used by art historians and 2PNNs.

Artistic style can be distinguished at many levels; each culture, period, school, and artist has a distinctive style, and even each painting can have a style unique to itself. Every aspect of a work of art contributes to its style -- the medium, the technique, the formal elements, and the principles of design all contribute to the style of a work of art.

In this work, we focus our efforts on the formal elements employed in a painting. To measure the importance of these elements quantitatively, we generate {\em style images} that retain one or more of the formal elements de-emphasizing or removing the others. We then train 2PNNs on these style images, and evaluate the performance of the network at a style classification task.

The main contributions of our paper are:

\begin{itemize}

\item We propose using a 2PNN classification network to quantify the amount of information about a style that specific formal elements convey

\item We evaluate the performance of a neural network on a variety of style images -- such as brush-stroke images [cite] and hue images with or without composition preserved

\item We demonstrate that despite eliminating much of the information in the original image, style images can be used to obtain classification accuracy close to that of the full color image.

\end{itemize}

% TODO: Test on composition images (?) and edge images. Name here.

% TODO: Prior work:

% - color [cite]

% - brush-strokes [cite]

% - composition [cite]

% [e.g., Re using neural nets for segmentation:

% cite: MULTI-SCALE CONTEXT AGGREGATION BY DILATED CONVOLUTIONS

% Fisher Yu]

% - edge [cite]

% \* Work toward stating our contribution on the first page if at all possible.

(Although key objects in an image may associate it with a style, it is quite possible to have two paintings which illustrate the same subject matter, but have markedly different styles. Examples in the Renaissance and Primitivism movements are shown in Figures \ref{fig:mona-r} and \ref{fig:mona-p}).

\section{Related Work}

\subsection{Painting style recognition}

Traditionally, researchers consider the style is the visual appearance represented by a combination of visual traits. Training the machine to detect the visual traits and judge the style of painting by these traits becomes a very common approach used to recognize the painting style. However, the traits selected by researchers and the training algorithms are different.

Feature-based style recognition: This type of algorithm extracts visual features then input these features into a learning algorithm while a single feature may not reliable enough for the recognition. Accordingly, this algorithm needs multiple features to recognize the style of painting. For example, one method develops total fifty different high-level features for color and composition (e.g. saturation and hue are two color features); then this method feed these features into a self-organizing map (SOM) to achieve a binary painting style classification for Impressionism, Expressionism, Impressionism and Post-impressionism.(Lee cite) Another method utilizes the whole HSV space of the image as the color feature; then extract local "texture" by using Gabor Filter and "edge" by using Canny edge detection from gray-scale paintings. These features are fed into several learning algorithms (e.g. K-nearest neighbors (KNN), Support vector machine (SVM), etc.) to classify Realism, Abstract, Impressionism, Cubism, Pop art.(Jana cite) These works represent that every element of a painting can contribute to the style, which inspires us to discover the merits of different elements.

Image-based style recognition: Instead of extracting well-defined visual features, these works train the machine to classify style by a whole image. (cite) Some other papers try to select some parts of the image which contains most of image style information. (cite) This recognition trains machine to decide what feature should be detected then judge the style by these features detected by the machine instead of manually selecting features as the input. For example, CNN can capture different features (e.g. edge and Gabor filter's response) in some intermediate layers and the other layers can contains more information besides the well-defined features, which is an advantage. Currently, CNN is the most leading learning algorithm used by the painting style researchers.

\begin{figure}

\begin{center}

\includegraphics[width=0.5\linewidth]{mona-lisa-real.jpg}

\caption{Mona-Lisa, Leonardo-Da-Vinci, Renaissance}

\label{fig:mona-r}

\includegraphics[width=0.5\linewidth]{mona-lisa.jpg}

\caption{Mona-Lisa, Fernando-Botero, Primitivism}

\label{fig:mona-p}

\end{center}

\end{figure}

Texture-based style recognition: The algorithms deploy Gabor filter to extract texture of the whole painting as one of the feature then input the texture into the machine. The texture is merely an input for learning algorithm. (cite) One recent popular algorithm combining the texture-based and image-based recognition to handle the texture of image inside is called two-pathway CNN, was proposed by Gatys et al to transfer image style originally. This two-pathway CNN adds an texture pathway to a regular CNN to minimize the texture differences between two paintings to transfer the image style to another one without altering the original objects. Since it's capability to isolate and capture the painting style, the two-pathway CNN was extended to classify multiple painting styles. Our work deploy this CNN to capture the texture of brush-stroke, which is more representative than the texture of color and composition, to classify the painting style.

\subsection{Brush stroke detection}

Automated brush-stroke extraction was widely used for artist identification. One common approach is to detect the space surrounding by isolated edges in painting. On the other hand, the K-clustering segmentation to extract brush stroke segments is also popular. Recently, Jia et al. (cite) proposed a new technique, automated brush stroke extraction, which combines edge detection and K-clustering segmentation to extract brush stroke and identify Van Gogh's painting from his contemporaneities. This also showed the brush stroke was unique sufficiently in Van Gogh's paintings by looking at the four marking attributes NBS-NB (Number of Brushstrokes in the Neighborhood), elongatedness, straightness, and BH (Broadness Homogeneity) (cite). This fact inspired us to extend to not only select the brush stroke limited to identify Van Gogh's work but also cluster brush stroke set for each style since different brush stroke may have variances on these aspects.

\section{The Model}

\subsection{Automated Brush stroke Extraction (ABSE) Formulation}

\begin{figure\*}[h]

\begin{center}

\includegraphics[width=0.8\linewidth,height=0.5\linewidth]{brush\_flow.png}

\caption{Two pathway CNN}

\label{fig:brush}

\end{center}

\end{figure\*}

The brush-stroke extraction algorithm

introduced in Jia et al.'s paper was designed to “distinguish Van Gogh's painting from his Contemporaries". The algorithm is based on an observation that

%%

%the space surrounded by edges is the representation of brush stroke. (Yahui Modified)

the space surrounded by edges is the representation of brush stroke.

Accordingly,

%ABSE

the extraction algorithm mainly

uses both edge detection

% This must be based on a different observation right?

to label the edge-surrounding space as potential brush stroke then design a set of tests to filter out the non-brush-stroke components. In the following section we will go through the ABSE process.

\subsubsection{Algorithm}

The brush stroke extraction is shown in Fig. \ref{fig:brush}. By using Canny-edge detection, edge thinning, and edge linking operation, we can extract clearly single pixel edges. In order to extract the spaces surrounding by every single edges, we need to list all the edges which are not enclosed. For the purpose of closing the open edges, the regular morphological operation ( e.g. enclosing) should be avoided since it might ruin the original shapes of edges. Instead, for every open edge, we find out its endpoints and connect them to the nearest edge within the endpoints' neighborhood of size $30\times30$. After we label and extract all the spaces which each of them is fully-compassed by one isolated edge, all these spaces are brush stroke candidates. Only the potential brush strokes satisfy the criterion can be labeled as true brush stroke.

There are four tests for the brush stroke:

\begin{itemize}

\item The size of brush stroke are unlikely too big or too small (In our case the thresholds are 100 to 1300).

\item Besides, the skeleton of the brush stroke cannot be severely branched. We firstly need to identify the skeleton of every brush stroke candidate. For one skeleton, if we detect all the branching points and the edge potion connecting two branching points is shorter than other branches, the branch is considered as severely branch and it should be removed.

\item The ratio of broadness of the skeleton to the length of the skeleton is in a reason range (In our case is 0.05 to 1).

\item $ k = \frac{size\, of\,brush\, stroke}{2\times{skeleton\, length}\times{width}} $ is in a reasonable range (In our case is 0.5 to 2).

\end{itemize}

\subsection{The two pathway convolutional neural network (TPCNN) Formulation}

Convolutional Neural Networks (CNN) were first proposed by LeCun et al. for image classification (cite). Benefiting from large databases for training, CNN provide decent classification since it "stimulate the biological activity in so-called complex cell in primary visual system"(cite) The paper (cite) presented by Gatys et al. re-designed a two-pathway CNN, one pathway for object detection and another for texture detection, for transferring artistic texture. Originally, the two pathway CNN was designed to minimize differences between one original image's texture and another style image's texture without altering the original image's content, which changes the original image's style in human perception.

Therefore, the texture is an important basis for human to recognize styles. Sun et al. (cite) extended the two pathway CNN model specifically for painting style recognition. Our CNN is built upon this network since if it gives reliable performance

%% Something about this phrase feels funny to me (Yahui Modified)

This implementation is benefited from this model to set up the same texture structure for the brush stroke style recognition analysis. In addition, abundance of deep pre-trained CNN models (VGG-19, AlexNet, etc.) and datasets (e.g. WikiPainting, etc.) are accessible.

\subsubsection{Texture pathway}

\begin{itemize}

\item Input: Down-sampled feature maps from one Relu layer

\item Output: image texture representation map

\item Preprocessing: Feed image into a CNN model. Collect sets of the feature map produced by each Relu layer.

\item The first step: Gram matrix. The gram matrix computes the correlation between the filter responses, where the gram matrix $G\_{ij}^l$ is the inner product between the vectorized feature maps $i$ and $j$ in layer $l$:

$$

G\_{ij}^l = \sum\_kF\_{ik}^lF\_{jk}^l \eqno{(7)}

$$

\item The second step: full-connected layer and SoftMax function.

The output of the concatenate layer will be fed into a regular full-connected layer. The detail of this full-connected layer can be seen in the full-connected layer section in the regular CNN. Then we apply the regular SoftMax layer here to give a non-linear properties. Then we can have a texture prediction.

\item The prediction:

The final prediction is calculated by:

$$

P = \alpha{P\_{texture}}+\beta{P\_{object}}

$$

\end{itemize}

The algorithm is shown in fig. \ref{fig:2pnn}

\subsection{TPCNN: Adaptation for brush stroke CNN}

\section{Experiment}

The proposed approach is implemented on TensorFlow by using VGG-19 neural network. The training data set is applied on the styles in WikiPainting dataset. Three experiment sections will be proposed: 1) A prior style transfer experiment. This experiment reveals the reason we make a hypothesis that brush-stroke is the most important factor for painting style. 2) Painting style classifications with different element input. Then the performance of the classification with different elements of painting will be compared. 3) Style classification for brush-stroke image. This section will discuss the influence of brush-stroke to the full-color image and the limitation of brush-stroke.

\subsection{Dataset setup}

\subsubsection{WikiPainting}

The WikiPainting dataset (cite) is one of the largest image style dataset for artistic images. The data set includes 85000 image cover more than 50 different genres of image with their style labels. Most of the images have size from range of 181 $\times$ 251 to 5133 $\times$ 5210. Due to the observation that the brush stroke extraction for image below 1700 $\times$ 1700 lost a great amount of brush stroke information. We applied a super-resolution technique, the Rapid-and-Accurate super-resolution algorithm (cite) to resize the images to at least 1700 $\times$ 1700. Then we can obtain better brush stroke image (Fig below). After the super-resolution, we select total 10 genres of images, includes Impressionism, Baroque, Ink and Wash Painting e.g., which each of them contains more than 1000 resized image. In order to balance the training and testing for the two pathway CNN, we only reserve 500 image in each genre folder. In addition, all the images are resized back to 128$\times$128 pixels to match the VGG-19 CNN. We have Impressionism, Academicism, Baroque, Expressionism, Photorealism, Pointillism and Ink and wash painting for image formal element analysis; Realism, Romanticism, Ukiyo-e as additional styles to pure brush-stroke image classification.

In addition, besides the brush-stroke dataset, we applied the 2PCNN to the other formal elements: color and line.

\begin{itemize}

\item Brush-stroke: The original brush-stroke image is binary. Due to our observation that binary image may not be able to effectively activate the neurons in the neural network. We assign random number to the brush-stroke image. We assigned random integers from 100 to 255 to the brush-stroke region while we assigned random integers from 0 to 30 to the background. We used the same strategy for the line dataset. In the evaluation stage, the images should be re-randomized.

\item Color: With a rich literature suggestions, we convert the original image from RGB to HSV. Then only maintain the pixel values in H channel and set the values in other channels to 255. After that, the image is converted back to RGB model. With that we can provide full lightness and saturation, and only represent hue in the RGB image.

\item Line: We simply apply Canny edge detection to sketch the lines in the image. We also applied the same random number strategy to it.

\end{itemize}

\begin{figure}[ht]

\begin{subfigure}[b]{0.475\linewidth}

\centering

\includegraphics[width=\linewidth]{init.png}

\caption{Full-color image}

\label{fig7:a}

\vspace{2ex}

\end{subfigure}%%

\begin{subfigure}[b]{0.475\linewidth}

\centering

\includegraphics[width=\linewidth]{brush.png}

\caption{Brush-stroke image}

\label{fig7:b}

\vspace{2ex}

\end{subfigure}

\begin{subfigure}[b]{0.475\linewidth}

\centering

\includegraphics[width=\linewidth]{H.png}

\caption{Color image}

\label{fig7:c}

\end{subfigure}%%

\begin{subfigure}[b]{0.475\linewidth}

\centering

\includegraphics[width=\linewidth]{edge.png}

\caption{Line image}

\label{fig7:d}

\end{subfigure}

\caption{Illustration of different cases}

\label{fig7}

\end{figure}

\subsubsection{Dataset enhancement}

In order to enrich and balance our dataset, we crop each image into several pieces to 1) increase the number of the images in the dataset; 2) each dataset has the same number of images. Each image piece is almost as large as its original image. Accordingly, in the training stage, we still have the 2PCNN to predict the styles for each piece; in the validation stage and evaluation stage, we will average the scores from each piece pieces then make a overall prediction for all these pieces. Before the cropping, the training dataset, validation dataset and evaluation dataset have to be separated firstly to prevent there is any image has its piece in different dataset.

In addition, since in the pure brush-stroke image classification, the content of images can be ignored; we did not extract all the brush-stroke from an image. Instead, we only extract brush-stroke from a part of the image then duplicate the brush-stroke throughout the image to save time. We massively use this strategy for Realism, Romanticism, Ukiyo-e.

\begin{figure}[h]

\begin{center}

\includegraphics[width=0.8\linewidth,height=0.5\linewidth]{Fanchart.JPG}

\caption{Dataset separation for experiments}

\label{fig:2pnn}

\end{center}

\end{figure}

%% The number of images for each styles table

%% The image brush stroke for small image and the big image and the super-resolution image

\subsubsection{Two pathway CNN}

The two pathway CNN is built upon a VGG-19 pre-trained CNN. The batch size for the model is set to be 16 by default. If CPU is enabled, this number can be larger. However, GPUs do not usually have as much memory as CPUs, so not many images can be processed at once when the GPU gets enabled.

\begin{itemize}

\item Object pathway: The model is same with the VGG-19 model provided on the Internet, and no changes had been made to the original one. The only parameter adjusted in this path is the output unit of the final fully-connected layer. This number should match with the total number of styles in the dataset. The parameter number of and full-connected layer

\item Texture pathway: On the basis of the object pathway, five layer outputs are gathered to calculate the gram matrix and the style loss, and they are: "relu1-1", “relu2-1”, "relu3-1", "relu4-1", and "relu5-1". Each gram matrix has its own weight. For the single output image of each layer, cols and rows are combined into one dimension, and this forms two dimensional matrix with the color channel. By using the equation (7), gram matrices can be calculated from five activation layers. Since we need these matrices to be fully-connected so that Tensorflow can give us the prediction based on these matrices, five gram matrices are converted into five single dimension array, and they are concatenated with each other to form one single array. Three fully-connected layers are applied to this array to compute the prediction for the image.

\end{itemize}

The whole model gives two sets of predictions at this point, and they are from the object pathway and the texture pathway. In order to calculate the total prediction, we give a weight to each of these two sub-predictions to make the final prediction. In the end of the each train iteration, the loss is calculated by using the

% not sure which function is used for calculating the loss

softmax function, and it gets optimized by using Adam algorithm. Tensorflow handles the work for adjusting weights and bias of the model, so we do not pay much attention to what happens in this process.

\begin{center}

\tiny

\centering

\begin{tabularx}{\columnwidth}{|X|X|X|X|X|}

\hline

\bf Full-connected layer & \bf FC1 \ output size & \bf FC2 \ output size& \bf FC3 \ output size& \bf Prediction Weight\\

\hline

Object FC & 512 & 512 & Number of styles & 0.1\\

\hline

Texture FC & 1024 & 1024 & Number of styles & 0.9\\

\hline

\end{tabularx}

\end{center}

\subsubsection{Evaluation metric}

We evaluate the performance of each formal element case by utilizing accuracy, the number of correct predictions divided by the number of the samples, for multi-label classification. Since the datasets are balanced, the goodness of fit of standard sample accuracy is enough to describe the prediction performance. In brush-stroke style classification section, we also introduce category accuracy; that is, the ratio of correct prediction in one category with the samples in this category.

\subsection{Results}

\subsubsection{style transfer}

To better understanding how the formal element affect the painting style, a simple style transfer experiment was implemented. In this experiment, we transfered the style of brush-stroke image, line image and full-color image to a photo, respectively. The results are below:

\begin{figure}[h]

\begin{center}

\includegraphics[width=\linewidth,height=\linewidth]{44.JPG}

\caption{Dataset separation for experiments}

\label{fig:2pnn}

\end{center}

\end{figure}

The style transfer for each case provides successful result image with a slight difference. The line image transfers solid and sharp edges while the full-color image transfers smoother textures. Both of them has blur in the top half of the result images. The brush-stroke image transfers a certain pattern throughout the whole result image. These style transfers inspire to discover what formal element is the deterministic factor for style generation; and what formal element determines the painting style.

\subsubsection{Formal elements analysis}

Here is the table for showing the performances for the 2PCNN. Object pathway accuracy (OPA) indicates the prediction accuracy from the full-connected layers in the object pathway; Texture pathway accuracy (TPA) indicates the prediction from the Texture full-connect layers. Two pathway accuracy (2PA) is the final prediction from the average value between these two sub-predictions.

\begin{center}

\tiny

\centering

\begin{tabularx}{\columnwidth}{|X|X|X|X|}

\hline

\bf Formal\-elements inputs & \bf OPA (\%) & \bf TPA(\%) & \bf 2PA(\%) \\

\hline

Full-color & 14.3 & 97.1 & 97.5 \\

\hline

Color & 14.3 & 56.1 & 57.8 \\

\hline

Gray-scale & 14.3 & 97.3 & 97.1 \\

\hline

Brush-stroke(Random) & 14.3 & 81.5 & 85.9\\

\hline

Brush-stroke(Gray) & 14.3 & 92.0 & 93.2\\

\hline

Line(Random) & 14.3 & 56.1 & 64.0 \\

\hline

Line(Gray) & 14.3 & 68.2 & 71.6 \\

\hline

\end{tabularx}

\end{center}

To observe the distribution for each formal element to the style classification, a control model is set up for the comparison. With the same network setting and dataset setting, full-color image dataset is the baseline which has completed information for image spatial statistics. The pure color dataset, the brush-stroke (binary) dataset and the edge (binary) dataset are the individual formal element datasets. Their gray-scale version are also set up for revealing the effect of color information.

First of all, the random brush-stroke dataset (brush-stroke with fake gray-scale intensity) provides the highest accuracy, 83.3\%, for the style classification. While a binary brush-stroke dataset maintains a high accuracy, the accuracy for line dataset and color dataset were only 64.0\% and 57.8\%, respectively. This result verifies that brush-stroke image remains most of the necessary style information from the original full-color image. Related to the theory mentioned in the algorithm section, the brush-stroke image describe more sufficient spatial statics of the original image without the original gray-scale intensity and lines. The performances of three different datasets validate another hypothesis: even though brush stroke contributes mostly to the painting style construction, color and line also provide useful spatial statics. The 2PCNN requires all the statics to give the best style recognition.

Secondly, we investigate the prediction accuracy in the object-pathway, texture-pathway. One significant discover is that if the object-pathway and texture pathway are trained jointly, as in the 2PCNN, the prediction from only the object-pathway will merely provide any useful information for classification. Based on our experiment, the object-pathway prediction always eventually constantly recognize every style as one style. However, in all of cases, there is a slight improvement 2PCNN by combining the object logit and the texture logit before the Softmax function that gives the prediction. This result implies that the high-level features extracted by the object pathway, the contents, also contribute to the style. This can be partly explained by each style has its own content preferences. For example, Baroque painters prefer realistic bodies and object shapes; in the meantime Ink and Wash Painting painters usually depict mountains, rivers and natural scenes. However, the predictions from isolated object-pathway are very unreliable so the object-pathway is merely an auxiliary pathway for the style classification.

Besides the isolated element dataset, the full-color image, line image and brush-stroke image in gray-scale version also provide another vision for us. Since the gray-scale intensity value is a linear conversion from RGB, the gray-scale image provide as much color information as the full-color image. The gray-scale image shows as good performance as the full-color image. Then the gray-scale brush-stroke image, which is the combination of brush-stroke region and partial color (it is partial color information since the non-brush-stroke region are black) gave better performance compare to the random-brush-stroke dataset. However, the gray-scale color information's affect cannot be simply concluded by this better performance. A multiple randomized brush-stroke image experiments are required. In this experiment, we did not re-randomized the gray-scale intensity of the brush-stroke image in the evaluation dataset. We simply randomized all the datasets by using the randomize strategy in section 4.1.1, and evaluate the evaluation accuracy without any re-randomizing. This experiment were implemented multiple times to compare with the performance from brush-stroke image with original gray-scale intensity.

\begin{figure}[h]

\begin{center}

\includegraphics[width=\linewidth,height=\linewidth]{33.JPG}

\caption{Dataset separation for experiments}

\label{fig:2pnn}

\end{center}

\end{figure}

The result in this plot may implies a surprising fact that: the original gray-scale intensity, at least the pixel value in the brush-stroke region, does not provide important information for the 2PCNN to classify painting styles. Since the randomized brush-stroke can even provide higher accuracy than the one with the original intensity. Consider to the fact that full-color pixel value is a linear mapping to gray-scale pixel value, the experiment shows that color information are not necessary for 2PCNN to classify painting styles. This discovery is also mentioned by Ikuta et al.'s paper \cite{unknown}, the color should not be a part of style when 2PCNN is utilizing. However, it seems contradict to the standard opinion from art historians; they consider each painting style also has its preference for color, lightness or tone (a term to indicate a composition of lightness and shadow). Further experiments are required for color's investigation.

\subsubsection{Brush-stroke analysis}

\begin{figure}[h]

\begin{center}

\includegraphics[width=\linewidth,height=\linewidth]{22.JPG}

\caption{Dataset separation for experiments}

\label{fig:2pnn}

\end{center}

\end{figure}

Based on observation that the brush-stroke is the most important factor in the chosen formal elements. In this section, we shows the experiment for brush-stroke style classification with three additional styles: Realism, Romanticism and Ukiyo-e. The table shows the performances of the 2PCNN for brush-stroke image in each class.

First of all, for the styles Baroque, Impressionism, Pointillism, Realism and Ukiyo-e, the 2PCNN can provide accurate prediction with both full-color image and brush-stroke image. Obviously, these styles have significant different brush-strokes with each other. For example, Baroque paintings are famous about their loose and clustered brush-stroke to depict motions; while Realism paintings has mode smoother and solider brush-stroke since they try to be closer to realistic objects. Accordingly, the spatial statics created by the brush-stroke, are easier to be identified by the 2PCNN. However, the high accuracy for Impressionism paintings is surprising: their brush-stroke are also characterized as loose, or fragmented, like Baroque's. The 2PCNN can clearly differentiate Baroque and Impressionism can be explained by that, Impressionism paintings have more "personal" brush-stroke. The variation of brush-stroke among different painters in Impressionism is larger so the variation is also captured by the 2PCNN to make the prediction.

For photorealism, the high prediction accuracy with full-color and brush-stroke is also expected. Photorealism paintings are photos, which means they don't have any natural brush-stroke. The brush-stroke algorithm extract the regions which are arbitrary defined as "Brush-stroke", which are very unique in all the non-photo styles . Accordingly, the high accuracy also provides us another vision to the brush-stroke algorithm: the brush-stroke is the collection of areas which maintains most of the spatial statics of the original image. It is possible to extend this definition to be a general definition of brush-stroke in machine vision. Further works will be expected to validate this generalization. (Discussed)

\begin{figure\*}[ht]

\begin{subfigure}[b]{0.2\linewidth}

\centering

\includegraphics[width=\linewidth, height=100pt]{baroque2.jpg}

\caption{Baroque}

\label{fig10:a}

\vspace{2ex}

\hspace{0.3\textwidth}

\end{subfigure}%%

\begin{subfigure}[b]{0.2\linewidth}

\centering

\includegraphics[width=\linewidth, height=100pt]{point2.jpg}

\caption{Pointillism}

\label{fig7:b}

\vspace{2ex}

\hspace{0.3\textwidth}

\end{subfigure}%%

\begin{subfigure}[b]{0.2\linewidth}

\centering

\includegraphics[width=\linewidth, height=100pt]{ro2.jpg}

\caption{Romanticism}

\label{fig7:c}

\vspace{2ex}

\hspace{0.3\textwidth}

\end{subfigure}%%

\begin{subfigure}[b]{0.2\linewidth}

\centering

\includegraphics[width=\linewidth, height=100pt]{photo2.jpg}

\caption{Photorealism}

\label{fig7:d}

\vspace{2ex}

\hspace{0.3\textwidth}

\end{subfigure}%%

\begin{subfigure}[b]{0.2\linewidth}

\centering

\includegraphics[width=\linewidth, height=100pt]{uki2.jpg}

\caption{Ukiyo-e}

\label{fig7:e}

\vspace{2ex}

\hspace{0.3\textwidth}

\end{subfigure}%%

\begin{subfigure}[b]{0.2\linewidth}

\centering

\includegraphics[width=\linewidth, height=100pt]{baroque1.jpg}

\caption{Baroque}

\label{fig7:f}

\end{subfigure}%%

\begin{subfigure}[b]{0.2\linewidth}

\centering

\includegraphics[width=\linewidth, height=100pt]{point1.jpg}

\caption{Pointillism}

\label{fig7:g}

\end{subfigure}%%

\begin{subfigure}[b]{0.2\linewidth}

\centering

\includegraphics[width=\linewidth, height=100pt]{ro1.jpg}

\caption{Romanticism}

\label{fig7:i}

\end{subfigure}%%

\begin{subfigure}[b]{0.2\linewidth}

\centering

\includegraphics[width=\linewidth, height=100pt]{photo1.jpg}

\caption{Photorealism}

\label{fig7:i}

\end{subfigure}%%

\begin{subfigure}[b]{0.2\linewidth}

\centering

\includegraphics[width=\linewidth, height=100pt]{uki1.jpg}

\caption{Ukiyo-e}

\label{fig7:j}

\end{subfigure}%%

\caption{The full-color and brush-stroke pairs. Baroque pair, Pointillism pair and Photorealism pair have completed brush-stroke extraction. Romanticism and Ukiyo-e pairs only extract brush-stroke from a piece of full-color image and the brush-stroke is copied across the image}

\label{fig7}

\end{figure\*}

For Expressionism and Romanticism, the 2PCNN provides quite worse prediction with both full-color image and brush-stroke image In both cases, 2PCNN classifies massive images in Expressionism as Romanticism and verse vice. The results indicates that these two painting genres has similar semantic distribution; so the brush-stroke image, which preserve the semantic information, results in worse prediction as well as full-color images. For Ink and Wash painting, the 2PCNN constantly predicts them as Ukiyo-e with brush-stroke; however, the full-color image offer much high accuracy. Based on this observation, we can conclude that the Ink and Wash painting has similar brush-stroke with Ukiyo-e; the difference between these two genres are contributed by the other formal elements. In addition, for Academicism paintings, even though the prediction accuracy is high with full-color image, the 2PCNN classifies numeric of them as Romanticism with brush-stroke image. In these cases, the brush-strokes are no longer as dominated as it to the other styles for style classification, at least with current 2PCNN model.

Nevertheless, the brush-stroke still shows significant influence on the painting style classification. Due to the time limitation, for a great amount of images, we did not extract all the brush-strokes from each of the images. Instead, we cropped the images into 16 pieces and only extract brush-stroke from a part of the pieces. For the missing pieces, we replace them by a copy of brush-stroke from a random piece of the same image, which indicates that only a part of object shapes and color. This strategy is massively applied in Realism, Romanticism and Ukiyo-e. However, the 2PCNN still provides decent performance for most of the styles. We can therefore validate again that painting styles are the spatial statics, the low-level features regardless of their positions in the image.

\section{Conclusion}

%------------------------------------------------------------------------

{\small

\bibliographystyle{ieee}

\bibliography{research-citations}

}

\end{document}