Broad Strokes: Classifying an Artist's Impression Using Machine Learning

Ariyana Miri
(*Waste Management)*[2]
Computer Science Department,
Florida Southern College
Land O' Lakes, FL, USA
ariyanahmiri@gmail.com

Savanna Moss
(*Waste Management)*[2]
Computer Science Department,
Florida Southern College
Clermont, FL, USA
savannamoss9@gmail.com

Hannah Wilberding
(*Waste Management)*[2]
Computer Science Department,
Florida Southern College
Richmond, VA, USA
hannah.wilberdingbw@gmail.com

Brock Wilson
(*Waste Management)*[2]
Computer Science Department,
Florida Southern College
Longview, TX, USA
brockwilson199@gmail.com

Abstract

In this paper, we present three machine learning algorithms that can measurably predict the artist of a painting given an image of the work. We accomplished this by running a dataset of 2,730 images from 10 different artists using three supervised learning algorithms: a Convolutional Neural Network, K-Nearest Neighbors classifier, and Support Vector Machine. The data was partitioned using an 80:20 ratio, resulting in 2,184 training images and 546 testing images. All images were resized to 224px by 224px, and transformed to greyscale. When this data was run through the algorithms listed above, the respective accuracies were: 38%, 38.64%, and 50%.

Keywords: Artificial Intelligence, Machine Learning, Supervised Learning, Image Recognition, Computer Vision, Convolutional Neural Network, K-Nearest Neighbors, Support Vector Machine

I. Introduction

The motivation for this research paper came about with the question: Is art quantifiable? Art is something typically regarded as purely human, not something a computer can understand or create, yet we consistently place categorical and numerical values on art in the form of prices, grades, artistic elements, composition… Machine learning algorithms have been used to recognize vehicles, faces, thumb prints… but can they recognize something as intrinsic as an artist's impression? If so, it will be another testament to the power of artificial intelligence, and an indication of future possibilities in the field.

II. Background

The term machine learning (ML) refers to the study of computer algorithms that improve automatically through experience. For the purposes of this paper, we are focusing on the branch of ML concerned with computer vision. Below, we outline key ideas related to this field.

A. Image Recognition in Machine Learning

The first ML programs were written in the early 1950s. Image recognition was introduced to the field in the late 1960s, with many of those algorithms still being used today. This branch of the field, also referred to as computer vision, focuses on how computers can process and gain high-level information from digital images or videos.

There are many popular algorithms today in the ML field capable of processing visual data. For our research, we chose to utilize three of these models: a Convolutional Neural Network, a K-Nearest Neighbors classifier, and a Support Vector Machine. While these models are all able to classify data, specifically visual data, they accomplish this using different methods. This is why we chose to use three different programs, so that we could not only test the ability of AI to classify art, but also find which model preforms best.

B. Challenges Presented with using Art

There are both technical and conceptual challenges with using art as our data. On the technical side, we're dealing with images, which are harder to collect, store, and process than text data. They also make programs run much slower, since they have to load in such a large amount of data. This ended up making the data collection and processing phase of our research take much longer than initially planned, as well as the code testing phase.

Conceptually, a painting differs from an object. Objects tend to have more universal and distinct features, whereas paintings and works of art in general have fewer rules and less measurable consistency, even within works from the same artist. This will make it difficult for our models to identify specific attributes to use when classifying our data. Also, we converted all the images to greyscale, so our models will be missing the key feature of color when processing the data, having to rely solely on patterns present within the greyscale to classify.

C. Related Work

The world of art and AI has existed for years, but recently, there have been significant advances made demonstrating how far AI can go to understand art. There are many existing research papers using AI to classify various features of a painting, from style all the way to identifying objects depicted within the work. There have also been programs created to turn flat paintings into 3D VR experiences. Recently, a lot of attention has also been put into AI artists, that is, machine learning algorithms that can create their own "unique" works based on learned data from initial training images. It is overall an exciting new field that presents major potential for future applications of computer vision and machine learning.

III. Data Processing

The data needed for this project consists of a large set of paintings with accompanying artist labels for each artwork.

A. *Collection*

The images used in this work were mined from the Painter by Numbers [1] competition dataset. We used 2,730 images comprised of work from 10 different painters [1], with a varying number of images for each artist. The artists were chosen by us based on popularity, but the data for each was mined randomly.

B. *Transformation*

After the data was collected, it needed to be transformed in order to make it usable. Images need to be the same size in order to be classified, but due to the nature of the data, we didn't want to utilize cropping. Instead, we compressed the images and set them all to 224px by 224px. This caused varying degrees of warping within the images depending on the initial dimensions, but allowed us to keep all the data.

After the images were resized, we converted them from RGB values to a greyscale. This turns the three-part RGB value of each pixel into one greyscale value. Doing this makes the data easier to work with, and also creates a shorter run time for our programs. Now when the models read in the image data, they will flatten the data from each painting into a one-dimensional array of 50,176 greyscale pixel values.

IV. Application of Algorithms

Once the dataset was processed, we divided it into training and testing data using an 80:20 ratio, resulting in 2,184 training images and 546 testing images. After the training data had been run through the models, the models classified the testing data based on information learned in the training process. The details and results of this are presented below. All code described in the following sections is publicly available on GitHub [2].

A. *Convolutional Neural Network*

Our CNN starts by itterating over the individual pixels of an image, then slowly moves to larger and larger sections. Our final CNN consisted of 8 layers: 2 Convolutional, 2 Max Pooling, 3 Dense, and 1 Output. This resulted in 12,113,174 trainable weights, and a classification accuracy of 38%. A model of the network can be seen in **Figure 1.1** below.

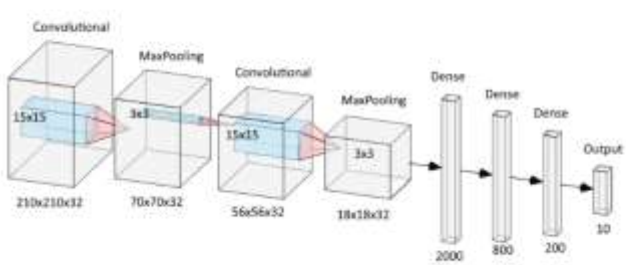

Fig. 1.1 CNN Model

1 Claude Monet, Edvard Munch, Edward Hopper, Frida Kahlo, Jackson Pollock, Johannes Vermeer, Leonardo da Vinci, Pablo Picasso, Salvador Dali, Vincent van Gogh

2 https://github.com/SavannaMoss/BroadStrokes

B. *K-Nearest Neighbors*

The KNN focuses on comparing the pixel values of the images. We implemented Principal Component Analysis (PCA) in order to perform dimension reductions on all our images, increasing the efficiency of the model. After testing several values in the program, we chose to use 28 neighbors and 80 PCA components, resulting in a final accuracy of 38.64%. In **Figure 2.1**, we've graphed the values of each testing sample's first three components (the components with the highest variance) generated by PCA.

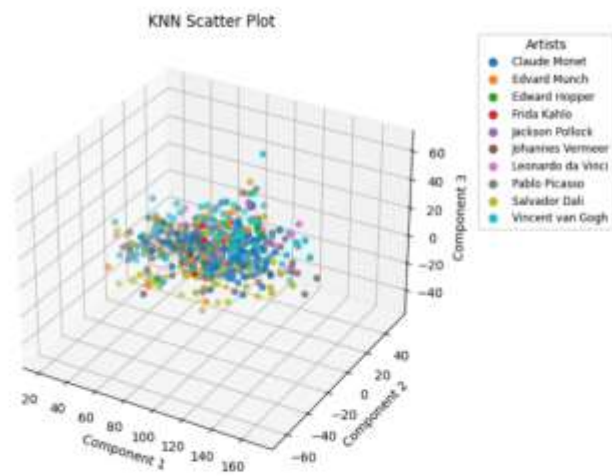

Fig. 2.1 KNN PCA Values

C. *Support Vector Machine*

Our SVM cell size was set to 16px-by-16px, with 14-by-14 cells per block. The model loops over the blocks while implementing HOG features, and uses a total of 45 different image orientations. With our random seed being set to 0, our final testing accuracy was 50%. We were initially worried that, due to the nature of the data, transforming images would not have the same affect it does with object recognition. Since paintings are made to be looked at from a certain orientation, and this could be an influential feature is helping recognize an artist's work, we were worried it might actually lower the accuracy of the program. However, implementing HOG features ultimately increased our accuracy by a substantial margin, making our SVM the most accurate model of the three.

V. Results

A. *Preformance Metrics*

The final testing accuracies of our CNN, KNN, and SVM were 38%, 38.64%, and 50% respectively. Since the accuracy that would result from random classification of the data would be 10%, these accuracies show that our models were all able to have measurable success in increasing the rate of successful classification. The second performance metric we used for evaluation was a confusion matrix. The matrices for our KNN and SVM can be seen in **Figures 3.1** and **3.2.** We also have a chart depicting the accuracy over time of our CNN in **Figure 3.3**.

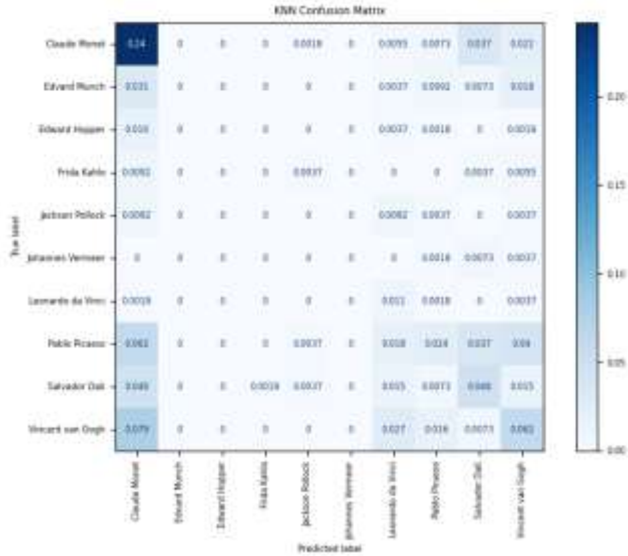

Fig. 3.1. KNN Confusion Matrix

Fig. 3.2. SVM Confusion Matrix

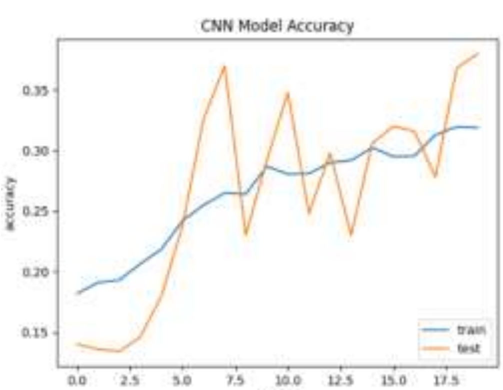

Fig. 3.3. CNN Accuracy over Time

B. *Conclusion*

Based on the results of this paper, machine learning algorithms can in fact recognize and classify an artist's impression, though some are better than others. Our results show that an SVM performs best with this type of data out of the three. This is likely due to the way in which each model processes the data. Since an SVM iterates over batches instead of pixels, it's able to better recognize objects and patterns within the images than a KNN, which primarily iterates over individual pixels. However, we were expecting a higher accuracy on our CNN. The underperformance may be due to the way in which we built out the model, and is something to review in later versions of this research.

The results also showed us the specific accuracy breakdown of each individual artist. Upon reviewing the data provided by the confusion matrices, it was clear to us that the amount of data each artist had in the dataset significantly affected the classification accuracy of that artist. In order to better show this, we created a graph at **Figure 3.4** to compare the percent of data made up by the artist, and the percent of correct classifications of that artist given by the confusion matrices.

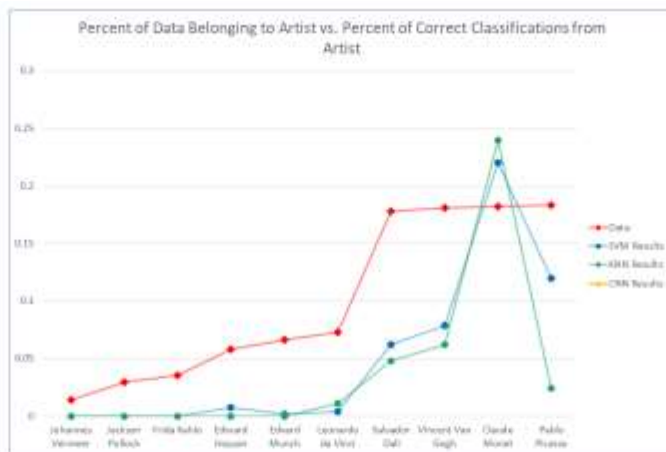

Fig. 3.4. Class Data vs. Result

There is a clear correlation between the amount of data and the resulting accuracy for each artist, with Picasso being the outlier. We're not sure why Picasso's accuracy was so low given that he had the most data in the dataset. We would have to do more analysis of where in the models the data is being misclassified in order to better understand the problem. This anomaly was particularly interesting to us since Picasso has such a unique painting style, one that even people without much art knowledge are able to distinguish. This shows us that there are still a *few* things humans can do better than AI.

VI. Future Application

A. *Next Steps for Improvement*

There are several ways in which we can improve the models presented in this paper. The first our group noted was keeping the RGB values of the images instead of converting them to greyscale. While converting them was good for the processing capabilities and time constraints of this project, we believe having color as another attribute the models can use would increase our accuracies. However, given that the models didn't have color and were still able to make accurate predictions based on greyscale values shows both the impressive capabilities of machine learning algorithms, and that painters have a unique impression left in their work outside of color that even machines can detect. There would also likely need to be extra measures taken to avoid overfitting once color is introduced to the model.

Another way to increase our accuracies, as shown in our conclusion section, would be to have more data for the artists that have fewer artworks represented in the dataset. It would also be beneficial to have more data in general from all classes. This would allow the algorithms to learn more information from the training data, and therefore have higher prediction accuracies for our classes.

B. *Implications for the Field*

Computer vision is one of the most popular branches of machine learning today, and is continuing to grow rapidly. We already use and rely on technologies created by the field every day in the form of self-driving cars, thumb and facial recognition, language translation, waste management… Nearly every major business and field of research relies on some form of AI, with more and more of them integrating computer vision into their systems. Fields like medicine, engineering, and agriculture all rely heavily on image recognition to model, simulate, and predict a wide range of things using varying sets of data. This paper serves as yet another testament to the capabilities of machine learning, and indicates an impressive and ever improving future for image recognition as it becomes more and more integrated into our lives.

REFERENCES

[1] Kiri Nichol KN. Nichol ***Painter by Numbers***, Kaggle, October. 22, 2020. Accessed on: April. 29, 2016. Website. Available: https://www.kaggle.com/c/painter-by-numbers/data