Chinese Landscape Painting Classification Using a Convolutional Neural Networks

Olivia Tang
otang@sas.upenn.edu

Rose William
rosewil@sas.upenn.edu

Abstract. In this paper, we discuss our research on image classification as a high-level computer vision task with implications for art history. We utilized roughly 600 photos from publicly available databases to explore if convolutional neural networks could be trained to classify pre-modern artworks into their respective eras of origin. Following are our results of using a multi-layered neural network on a large dataset of digitized images of Song and Ming Dynasty landscape paintings.

Introduction

Historically, the fine arts and their study has been confined to expert eyes who are charged with discerning their origins, era, artist, and influences. With the introduction and democratization of large digitized fine art collections into the public domain through the efforts of museums and universities, art can now be scrutinized not just through connoisseurship but as big data. Inspired by the success of Convolutional Neural Networks (CNN) in image classification projects, we used CNN to categorize our assembled large fine art datasets on Song and Ming Dynasty Chinese landscape paintings into their respective chronological origin. Our reasoning behind doing so was driven by our curiosity at what the neural network would use to discern the two eras, or if it could even discern stylistic differences perhaps imperceptible to untrained human eyes. To examine the capabilities of the deep model in fine-art painting classification, we trained an end-to-end deep convolutional model pulling our dataset from the publicly accessible large-scale "ArtStore" dataset comprising over 2.5 million paintings and eventually building a network with an accuracy of approximately 70%.

Experiment Design/Initial Purpose

The purpose of this project was to investigate the potential role of machine learning in the arts. More specifically, can the nuances of artistic style through the eras be properly analyzed and used by these networks in a way that could benefit the field of art history? Where do the strengths of an algorithm lie as compared to a human? Initially, we toyed with the idea of studying whether or not a CNN could be used to differentiate between Chinese landscape paintings and Japanese kara-e paintings which imitated stylistic elements of Chinese Tang Dynasty paintings[1]. However, after speaking with Professor Nancy S. Steinhardt, a professor of East Asian Art at the University of Pennsylvania, we decided to pivot the project's focus away from kara-e paintings. At the suggestion of Professor Steinhardt, we decided to instead study the paintings from the Chinese Song and Ming Dynasties which offer greater potential for more conclusive results that are not confounded by the influences of painter nationality and culture. According to the Metropolitan Museum of Art, Ming Dynasty paintings emulated the style of ink wash Song Dynasty landscape paintings while advancing the usage of calligraphic technique that has come to characterize this era's artistic fingerprint[2]. We believe that the significant similarities and differences between these two eras of Chinese landscape paintings would provide for a fruitful test of the capabilities and limitations of using CNN in art history applications.

Our network itself is based on PyTorch's CIFAR 10 classifier[3]. We transformed our dataset images to the dimensions of 32x32 and defined batch size as 4 to fit the structure of our CNN which has two convolutional layers, one max-pooling layer, and three fully connected layers. A rough flowchart of this can be seen in **Figure 1**. After transforming the images into normalized tensors, we split the dataset into training and testing images. To train the network, we looped

---

[1]Encyclopædia Britannica, inc. (n.d.). Kara-e. Encyclopædia Britannica. Retrieved May 5, 2022, from https://www.britannica.com/art/Kara-e

[2]Ming Dynasty (1368–1644). Metmuseum.org. (n.d.). Retrieved May 5, 2022, from https://www.metmuseum.org/toah/hd/ming/hd_ming.htm

[3]Training a classifier. Training a Classifier - PyTorch Tutorials 1.11.0+cu102 documentation. (n.d.). Retrieved May 5, 2022, from https://pytorch.org/tutorials/beginner/blitz/cifar10_tutorial.html

over our data iterator to feed our training images into the network to optimize for 50 epochs. We then tested the resulting network on our testing dataset and received a 69% accuracy.

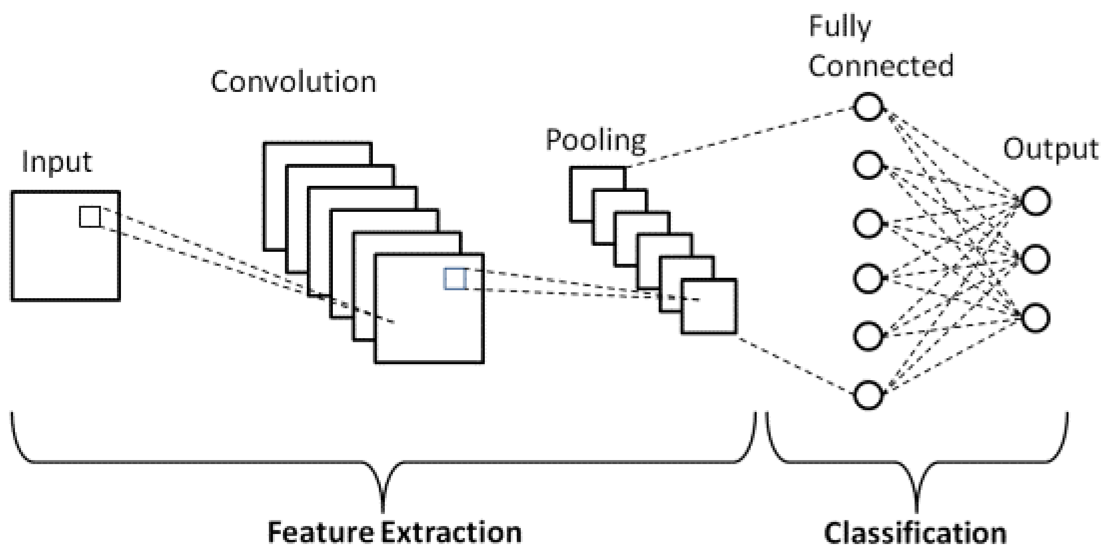

Fig. 1 CNN Architecture from https://shonit2096.medium.com/cnn-on-cifar10-data-set-using-pytorch-34be87e09844

Dataset

50% of our dataset was sourced from ArtStor, a fine arts and humanities database comprised of over 2.5 million pieces of artwork from museums, archives, and educational institutions all around the world. This dataset is considered a curational cornerstone and is constantly being updated by reputable organizations, making it a suitable source of images for our experiment. Pulling from the ArtStor database, we were able to create a dataset of 211 Ming Dynasty and 229 Song Dynasty landscape paintings for a total of 440 images as well as the accompanying metadata relating to the paintings' artist, style, genre, and origin. We then supplemented this dataset with more images pulled from web image search scraping to create a more robust dataset. Our final dataset resulted in 322 Song Dynasty paintings and 311 Ming Dynasty paintings for a total of 633 images. Approximately 80% of those images were used in the training set (506) and

20% were used in the testing set (127). To fine-tune our dataset, we narrowed our ArtStore searches to include only ink and watercolor artworks that depicted primarily landscape scenes on paper. By doing so, we deliberately filtered out other styles of Song and Ming Dynasty art like furniture and calligraphy to eliminate as many confounding variables as possible and thus allowing for the direct comparison of Song and Ming Dynasty landscape paintings. We defined Song Dynasty art as originating from the years 960 until 1279 and Ming Dynasty art from 1368 until 1644. Shown in Figure 1 below are examples of Song and Ming Dynasty landscape paintings pulled from our database.

Fig. 2 Sample images from the dataset. Though they are all landscape paintings, there are clear stylistic and compositional differences.

Hypothesis

The experimental question that our project posed was whether convolutional neural networks can be trained to classify Song and Ming Dynasty landscape paintings into their respective eras of origins. We hypothesize that the model will be unable to classify Song and Ming Dynasty landscape paintings to a high degree of accuracy given our lack of knowledge in training and utilizing neural networks. That being said, this project might provide insights on stylistic differences that exist between Song and Ming Dynasty art that would not be possible if the average layperson were to be tasked with similarly classifying images.

Implications of the Experiment

The results of our project could potentially be seen as a jumping-off point for the utilization of machine learning neural networks in the art history field. Although it is unlikely that this project yields significant results given our limited background in computer science, more rigorous projects would have greater implications in applications of fine art classification. For example, fine art authentication could be impacted especially in the Museum curation space. If neural networks can be trained to reliably and accurately classify paintings into chronological categories or be trained to spot characteristics of forgeries that are imperceptible to the human eye, then the art verification and authentication field could potentially become automated in the future. While this future seems very far off, it is important to think about the consequences of removing humanness from fields that inherently revolve around interpreting an art form that conveys human thoughts and emotions. Perhaps having computers and models exclusively classify paintings would be doing a disservice by denying human interpretation and subjective views on the paintings' origins and their authenticity. These are all important implications to consider as we move forward with our project.

Results or Builds

The results yielded a 68% accuracy for the testing set, with 78.6% for Ming Dynasty paintings and 60% for Song Dynasty Paintings. It also yielded 88% accuracy on the training set. Given that the original dataset was roughly equal in both Song Dynasty and Ming Dynasty inputs (50.8% Song Dynasty paintings and 49.8% Ming Dynasty Paintings), this is significantly better than random chance. The difficulty of this computer-vision task for our chosen neural network and the limited number of digitized Song and Ming Dynasty landscape paintings within our dataset proved to be hurdles to accuracy. Because our dataset was not as large as other datasets used to train neural networks, this may have caused overfitting to occur to some extent. It would be interesting to see how our model would perform on a significantly larger dataset of Song and Ming Dynasty landscape paintings. While the results are not high enough to compete with human experts, there is reasonable evidence to believe that there are discernable differences between Song and Ming Dynasty landscape paintings that can be identified by neural networks.

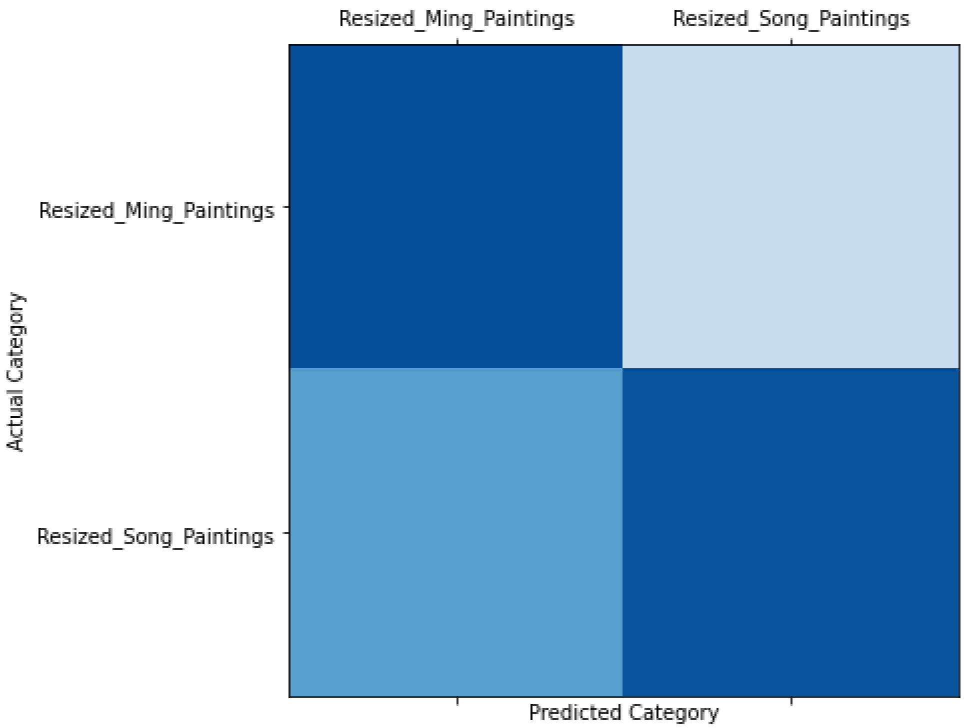

Fig. 3 The confusion matrix for classification. The rows show the real categories and the columns show the predicted categories. Darker colors indicate a stronger correlation while lighter colors indicate a weaker correlation.

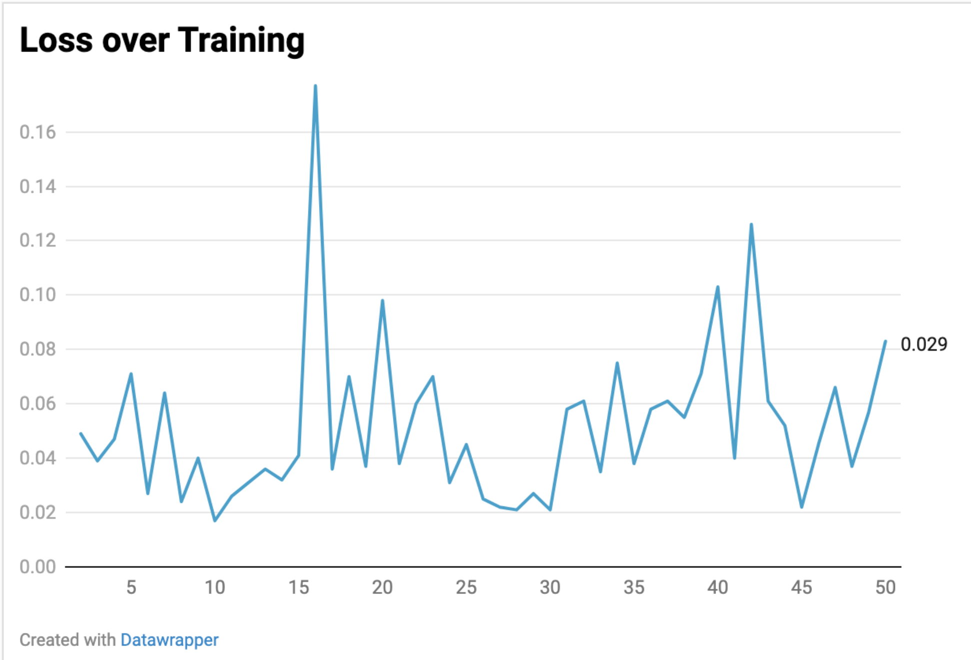

Fig. 4 The loss over training epochs of the CNN.

Conclusion

Fine-art categorization could supplement or possibly replace human labor in categorizing or authenticating artworks as digital artworks. Machine learning can use a larger quantity of data to categorize a broader spectrum of art much faster than a person which is especially relevant with the growth of online art auctions and art ownership that resides entirely in the digital domain like non-fungible tokens. Additionally, the automation of art classification could help art museums build out online exhibitions where paintings are ordered in chronological progression thereby allowing the viewer to see the continuity and change of style and form across eras. That being said, there is still great value in human interpretation that should not be thrown away especially as art historians seek to understand the cultural and social implications of stylistic changes throughout the era. More research is needed to fully harness the potential of computer vision and use it to benefit the current work of art historians and museum curators.

Next Steps

The next steps for the experiment would be to expand our dataset significantly as well as look at different, more advanced neural networks that could be more suited for art classification purposes. Perhaps this is one of the reasons why our algorithm did not reach a high degree of accuracy due to the rudimentary nature of the network we were utilizing as well as the limited dataset. It could also be beneficial to go over images that were misclassified with an art historian and try and see if there is an obvious pattern to these misclassified images. The experiment could also transition to look at Chinese paintings that have a larger chronological gap between them which could potentially make it easier for the algorithm to classify them since the stylistic differences would be more significant.