Classifying digitized art type and time period

Sean Yang
University of Washington
Seattle, WA
tyyang38@uw.edu

Bum Mook Oh
University of Washington
Seattle, WA
bmo5@uw.edu

Daniel Merchant
University of Washington
Seattle, WA
danmerch@uw.edu

Bill Howe
University of Washington
Seattle, WA
billhowe@cs.washington.edu

Jevin West
University of Washington
Seattle, WA
jevinw@uw.edu

ABSTRACT

Millions of art images have been digitized over the last several decades. This has created new opportunities for art scholars and historians. However, searching and navigating these art images is difficult because of the sparsity of metadata and contextual information used to describe these images. Unless one knows the exact title and artist, finding related artwork is a difficult task. The research in this project addresses this challenge by developing unsupervised computer vision methods that generates metadata automatically from artworks. Our dataset includes more than 300,000 art images from three sources: the Metropolitan Museum of Art, WikiArt and Artsy, an online art collection platform. If successful, we plan to build an interactive interface for exploring the extracted features and developing a recommendation system that can be used by art historians, scholars, and art aficionados.

ACM Reference format:
Sean Yang, Bum Mook Oh, Daniel Merchant, Bill Howe, and Jevin West. 2018. Classifying digitized art type and time period. In *Proceedings of International Journal for Digital Art History, London, UK, August 2018 (DSDAH'18),* 5 pages.
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

With the huge success of deep learning techniques like convolutional neural networks (CNNs) on natural images, it seems fitting to extend these methods to other kinds of images, such as artificial images (e.g., digital artwork and scientific figures). Lee. et al [14] utilized state-of-the-art AlexNet[12] and ResNet50 [9] to classify scientific figures into 8 categories and further extract phylogenies from the phylogenetic tree diagrams. Elgammal et al. [5] investigated the style classification of art paintings using several cutting-edge CNNs and compared the learned representations with concepts derived from art history.

However, compared to natural images, artificial images are largely ignored due to comparatively limited access and less organized data. Most recent research [2, 5, 17, 20] makes use of data from, WikiArt[1], which features approximately 250,000 artworks by 3,000 artists. With the recent public release of art images by The Metropolitan Museum of Art (The Met), more digital artworks are now available to the public. Most research in this area focuses on identifying image style [11], such as Impressionism, Expressionism, Cubism, and so on. We seek to classify images on other traits such as time period and art type (e.g., painting or photograph).

We chose to classify images based on time period and art type because both of these characteristics are often used for digital art libraries as a means of data filtering. For example, a user may want paintings from the 1750s, and as such, will search based on these criteria. This search functionality is dependent on the reliable metadata describing the time period, which for many art pieces is not available. Developing a classifier that can accurately predict these features for images missing this information would greatly improve search functionality. We limit the scope of art types to four categories: drawing, painting, prints, and photographs. They comprise the majority of two-dimensional artworks from the dataset. For the time period, we isolated art images from the 15th century to the 20th century.

In this paper, we contributed the following:

- We integrated two open access art data sources, Artsy and the Metropolitan Museum of Art image database that we use for this paper and will make available for other researchers working with these images.
- We reveal that these three different open access platforms feature different types of artwork and therefore different mining opportunities. WikiArt, which is popular for art analysis, mostly includes paintings. However, Artsy and the Metropolitan Museum of Art contain a more diverse collection of artworks, including drawings, prints, photographs, sculpture, and architecture.
- We use state-of-the-art CNN architecture (ResNet50) to classify images based on their time period and type. After removing multiply-labeled data and bucketing images based on the century they were created, we achieved 87.8% accuracy for type classification and 57.4% accuracy for time period classification using a ResNet50 model pre-trained on a corpus of 1.2 million images collected online.

[1]WikiArt dataset: http://www.wikiart.org

DSDAH'18, August 2018, London, UK

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

2 RELATED WORK

Machine learning and deep learning techniques like Convolutional Neural Networks (CNN) have been applied to digitized artwork for multiple applications. Researchers have used CNN to recognize artists [17, 21], classify artistic style [2, 5, 6, 17], transfer art styles to different images [7], and identify uncertain subjects with face recognition [18]. CNN is also utilized for art style transformation. Kong et al. [11] focus on transforming a generic image art style by considering a set of images. Strezoski et al. [20] propose an efficient and accurate method for multi-task learning with a shared representation applied in the artistic domain. Elgammal et al. [5] perform a comprehensive analysis of multiple CNN architectures applied on classification of art images and further investigate the learned representation through correlation analysis with concepts derived from art history. In this paper, we classify two common features of art work (time and art type) in order to improve search of these artworks. Unlike style, genre and artists, time period and art type are barely touched in existing research. However, both are significant attributes for art identification. Therefore, our work could be utilized to improve the current art search engine and we are hoping the tools could bring different perspectives for art historian.

3 DATA

3.1 Data Integration

Unlike most existing works which only utilize digital artworks from WikiArt, our data corpus also collected data from the Metropolitan Museum of Art and Artsy.

3.1.1 Metropolitan Museum of Art. The Metropolitan Museum of Art (The Met) has recently published more than 375,000 art images to use, share, and remix without restriction. For each artwork, information such as title, creator, data, nation of origin, medium, and dimension are provided along with the images. We download approximately 280,000 images with metadata from their open access service[2] and store the data in the viziometrics database of scientific figures and art images [13].

3.1.2 Artsy. Artsy is an online platform for discovering and collecting art. They offer roughly 27,000 artworks for unrestricted use through their public API. We collected 26,966 images and their associated metadata.

3.1.3 WikiArt. WikiArt data has contributed to several recent art-related research projects [2, 5, 17, 20]. For this paper, we were able to compile approximately 140,000 images from WikiArt open access.

We further examine the duplicates by matching title and creator of the artwork and remove 3,480 identical records.

3.2 Data Preprocessing for Training Classifier

Both Artsy and the Met provide labeled type data, but the majority of the data from the Met is labeled with more than one category. To avoid confusion, we only select data with one type label. We are able to acquire 7,713 photographs, 24,303 prints, 13,199 drawings, and 6,542 paintings. It is well known that imbalanced data causes poor performance for learning [8], so we randomly select approximately 7,000 from each category. We further split the data into training, validation and testing sets with 8:1:1 ratio. The distribution of the data across different categories and datasets is shown in Table 1.

[2]https://github.com/metmuseum/openaccess

Table 1: Data distribution with respect to type for CNN training

	Training	Validation	Testing
Paintings	5,223	673	648
Drawings	6,171	748	779
Prints	5,910	751	740
Photographs	6,207	756	752
Total	23,511	2,928	2,919

For the time period model, we remove data without time information and parse the year into century format. We average all values if more than one time value is provided. We obtain 6,965 images for 16th century, 7,737 for 17th century, 7,667 for 18th century, 18,238 for 19th century, and 12,004 for 20th century. We randomly select around 7,000 data for each time period for CNN training. Data is further split into three datasets with the ratio 8:1:1 (train, test, validation). The data distribution for time period is shown in Table 2.

Table 2: Data distribution with respect to time period for CNN training

	Training	Validation	Testing
16th century	4,895	614	605
17th century	6,201	770	769
18th century	6,106	780	783
19th century	5,740	683	743
20th century	5,619	668	710
Total	28,562	3,516	3,611

4 METHODS

In this paper, we present the preliminary results for both unsupervised learning and supervised learning approaches to understand "how machine sees art" and further classify digital artworks.

For visualization of data integration and time period analysis, the method is inspired by [10]. Every digital artwork is embedded into a 2048-dimensional space feature vector using a pre-trained ResNet50[9], which is trained on ImageNet corpus[12] (a corpus of 1.2M natural images collected from the web). Even though the model is pre-trained by natural images, it is believed that the representation should be general, capturing the edges and patterns of the figures. Further, we conduct Principal Component Analysis to reduce the dimension of the embedding vectors and plot the first and second components as a heat map to convey density. We then use the same pre-trained ResNet50 model to train the classifiers.

5 RESULTS

5.1 Data Integration

Figure 1 shows the deviation from the average distribution of figures from different sources. We ran PCA on images from all three

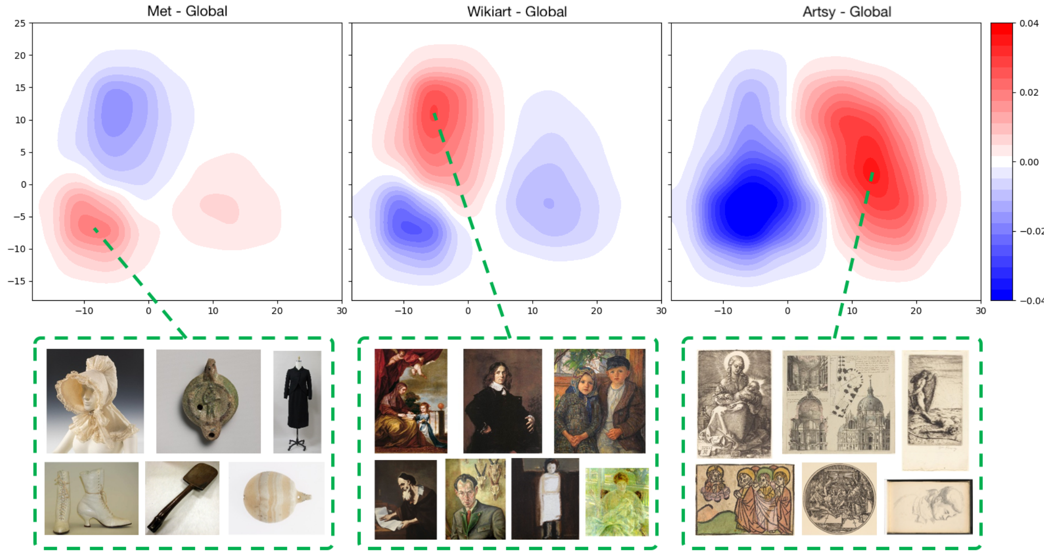

Figure 1: Deviation from the average distribution of figures from the Met, WikiArt, and Artsy. The heatmap for each collection represents the density deviation between the frequency of the collection and global average. The x-axis and y-axis are the first and second principal components, respectively. The red regions represent a positive deviation from the global baseline and blue regions represent a negative deviation.

collections. Then, we separated the collections and looked at how the image collections differed. The red regions represent a positive deviation from the global baseline (high proportion of given image type) and blue regions represent a negative deviation. You can see in Figure 1 that the Met collection contains a higher frequency of objects highlighted in the panel below. For all three highlights, images were randomly selected from the bin which contained the most figures in our analyses. For example, the top-left image in the highlight is a sunbonnet from the Met's costumes collection, and the top-center figure is a Roman terracotta lamp from their terracottas collection. WikiArt contains a higher frequency of paintings when compared to the Met (where this same region is blue). The Artsy dataset only included about 27,000 images, which was far less than the other two collections. This may partially explain the large blue and red regions. In addition, the art samples are quite varied, but the panel indicates a higher proportion of black and white images, as shown by the prints in the highlight, such as Albrecht Durer's *The Virgin with the Swaddled Child* in the upper left corner.

5.2 Time Analysis

Figure 2 shows the deviation from the average distribution of figures from different time periods. We group all the figures by century and demonstrate the density deviation for each century. Also, even though we integrate three different data sources and obtain roughly 350,000 images, selection bias might still exist in this analysis due to copyright issues and limited access to artwork.

In Figure 2, instead of separating by collection, we separate by time period. Image highlights were randomly selected from the bin which contained the most figures in our analyses. As in Figure 1, we highlight image types found in the red region that are in higher proportion when compared to the other time periods. If you compare the 16th and 17th century, you can see that they are similar in their representation - the images highlighted all come from print collections, such as Jacques Bellange's *The Raising of Lazarus* in the left of the 17th century highlight. The prominence of prints in 16th century we observe is probably due to the demand for maps, religious images, and illustrations [3, 19]. The 16th and 17th century image type frequencies are far different than the image type frequencies of the 15th and 20th centuries, which feature African tribal masks (upper left of 15th century) and colorful paintings, like Vincent D. Smith's *Study for Mural at Boys and Girls High School, Bedford-Stuyvesant, Brooklyn*. In the 19th century, we observe a positive deviation on photographs, since photographs were not invented until the early 19th century, and access to photography equipment was limited until the early 20th century [1]. We also show the frequency distribution with respect to time period and art type in Table for reference.

Table 3: Data distribution with respect to time period and art type

	Photographs	Prints	Drawings	Paintings
16th century	0	5550	1317	651
17th century	0	7117	1873	1040
18th century	0	6673	2679	1197
19th century	4854	11996	6169	2845
20th century	6037	7049	3945	2026

5.3 CNN Classifiers

Our classifier's performance is shown in Table 4.

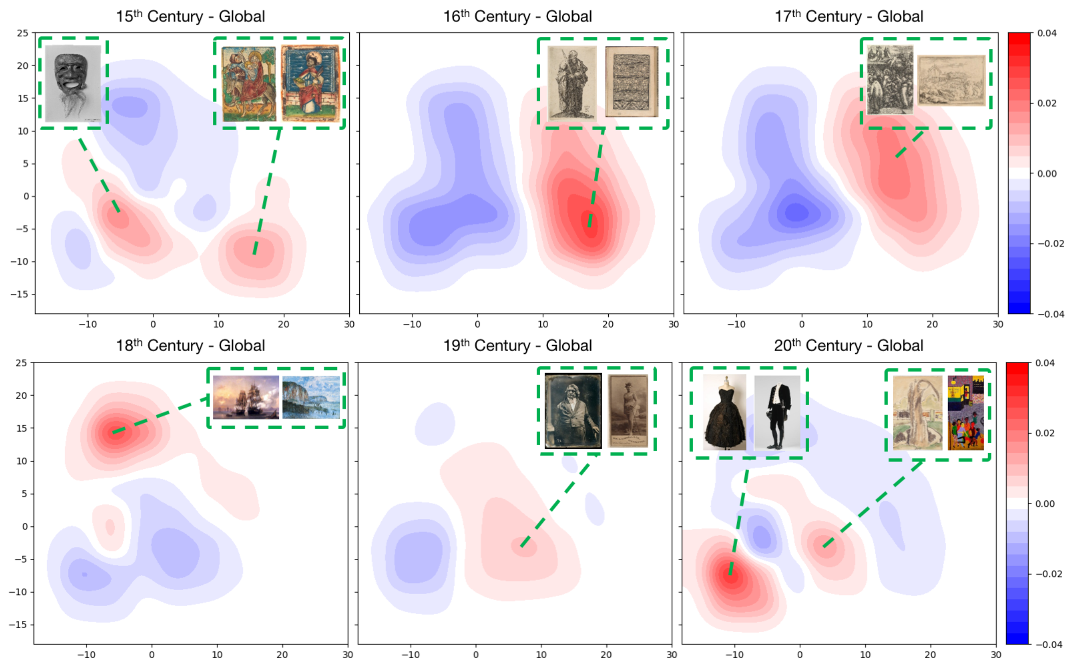

Figure 2: Deviation from the average distribution of figures from 15th century to 20th century. Same as Figure 1, the heatmap for each collection conveys the density differences between frequency of the collection and global baseline.

Table 4: Accuracy on CNN classifiers

	Type Classifier	Time Classifier
Accuracy	0.878	0.574

The classifier was reasonably accurate with respect to type, achieving 87.8% accuracy. The problem of classifying art types is similar to object recognition for natural images, where the model learns to recognize the distinctions among different objects. The challenge of identifying art type among prints, paintings, drawings, and photographs is the high similarity between these types. Even humans have trouble recognizing correct art type without additional information. Therefore, lower accuracy compared to natural image classification (3.57% error rate) [9] seems reasonable but something to look further into.

We achieved limited accuracy of 57.4% from time period classification. This is possibly due to the bucketing size we used when grouping data. Artistic styles and movements change at varying rates, but radical shifts in artistic form can be observed over small periods of time. For example, the cubist movement of Picasso, Braque and Gris existed in the same century as modern photography, but these images are unlikely to share visual traits [4]. In future analysis, we plan to separate these types to see how it improves accuracy.

Another possible factor is the diversity of concurrent art movements, especially with respect to geography. For example, Western European works from the 1600s to 1800s were largely prints and oil-based paintings with Christian imagery. North African art of the same period consists largely of textiles and jewelry, with heavy Islamic influences, such as Arabic calligraphy [15, 16]. Because of these disparate influences and materials used, images from these sets are unlikely to share visual characteristics, despite their shared time period. In the future, limiting the nations of origin to a smaller set could yield improved time-based classification.

6 CONCLUSION

In this study, we show preliminary results on digital artworks analysis based on an integration of data sourced from WikiArt, the Metropolitan Museum of Art and Artsy. Using a CNN classifier, we achieved high accuracy for classifying images based on their type, and limited accuracy classifying images based on their age. Possible explanations for this accuracy with respect to time period are the coarse granularity of the data and the diversity of art movements during a shared time period. The success of the type classifier presents a means of automated metadata enrichment for digital art libraries; by accurately classifying type, digital art databases could use this classifier to fill missing metadata in their dataset, and improve the speed of adding new images to the dataset by allowing computers to tag images instead of requiring a human to tag them.

REFERENCES

[1] Melissa Banta and Robert Burton. 2018. Harvard's History of Photography Timeline. (2018). https://library.harvard.edu/sites/all/themes/HarvardLibraryPortalTheme/timeline/index.html
[2] Yaniv Bar, Noga Levy, and Lior Wolf. 2014. Classification of artistic styles using binarized features derived from a deep neural network. In *Workshop at the European Conference on Computer Vision*. Springer, 71–84.
[3] Encyclopaedia Britannica. 2018. Printmaking - History of Printmaking. (2018). https://www.britannica.com/art/printmaking/History-of-printmaking#ref28335

[4] Michael Cirigliano II, Pac Pobric, Elizabeth A. Weinfeld, and Philomena Mariani. 2018. Heilbrunn Timeline of Art History. (2018). https://www.metmuseum.org/toah/chronology/

[5] Ahmed Elgammal, Marian Mazzone, Bingchen Liu, Diana Kim, and Mohamed Elhoseiny. 2018. The Shape of Art History in the Eyes of the Machine. *arXiv preprint arXiv:1801.07729* (2018).

[6] Zoe Falomir, Lledó Museros, Ismael Sanz, and Luis Gonzalez-Abril. 2018. Categorizing paintings in art styles based on qualitative color descriptors, quantitative global features and machine learning (QArt-Learn). *Expert Systems with Applications* 97 (2018), 83–94.

[7] Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. 2016. Image Style Transfer Using Convolutional Neural Networks. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

[8] Haibo He and Edwardo A Garcia. 2009. Learning from imbalanced data. *IEEE Transactions on knowledge and data engineering* 21, 9 (2009), 1263–1284.

[9] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.

[10] Bill Howe, Po-shen Lee, Maxim Grechkin, Sean T Yang, and Jevin D West. 2017. Deep mapping of the visual literature. In *Proceedings of the 26th International Conference on World Wide Web Companion*. International World Wide Web Conferences Steering Committee, 1273–1277.

[11] Lili Kong, Jiancheng Lv, Mao Li, and Hanwang Zhang. 2017. Extracting Generic Features of Artistic Style via Deep Convolutional Neural Network. In *Proceedings of the International Conference on Video and Image Processing*. ACM, 119–123.

[12] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. 2012. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*. 1097–1105.

[13] Po-shen Lee, Jevin D West, and Bill Howe. 2016. Viziometrix: A platform for analyzing the visual information in big scholarly data. In *Proceedings of the 25th international conference companion on world wide web*. International World Wide Web Conferences Steering Committee, 413–418.

[14] Po-shen Lee, Sean T Yang, Jevin D West, and Bill Howe. 2017. PhyloParser: A Hybrid Algorithm for Extracting Phylogenies from Dendrograms. In *Document Analysis and Recognition (ICDAR), 2017 14th IAPR International Conference on*, Vol. 1. IEEE, 1087–1094.

[15] Metropolitan Museum of Art. 2002. Western North Africa (The Maghrib), 1600âĂŞ1800 A.D. (2002). https://www.metmuseum.org/toah/ht/09/afw.html

[16] Metropolitan Museum of Art. 2003. Central Europe (including Germany) 1600-1800 A.D. (2003). https://www.metmuseum.org/toah/ht/09/euwc.html

[17] Babak Saleh and Ahmed Elgammal. 2015. Large-scale classification of fine-art paintings: Learning the right metric on the right feature. *arXiv preprint arXiv:1505.00855* (2015).

[18] Ramya Srinivasan, Conrad Rudolph, and Amit K Roy-Chowdhury. 2015. Computerized Face Recognition in Renaissance Portrait Art: A quantitative measure for identifying uncertain subjects in ancient portraits. *IEEE Signal Processing Magazine* 32, 4 (2015), 85–94.

[19] Alison Stewart. 2013. The Birth of Mass Media: Printmaking in Early Modern Europe. In *A Companion to Renaissance and Baroque Art*, Babette Bohn and John M. Saslow (Eds.). Wiley-Blackwell, Chichester, 253–273.

[20] Gjorgji Strezoski and Marcel Worring. 2017. OmniArt: Multi-task Deep Learning for Artistic Data Analysis. *arXiv preprint arXiv:1708.00684* (2017).

[21] Nanne Van Noord, Ella Hendriks, and Eric Postma. 2015. Toward Discovery of the Artist's Style: Learning to recognize artists by their artworks. *IEEE Signal Processing Magazine* 32, 4 (2015), 46–54.