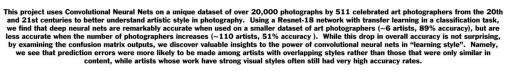
Understanding Photographic Style with Deep Learning

Jeff Sheng, PhD Candidate in Sociology, CS 230 Final Project, jtsheng@stanford.edu





Dataset

One of the strengths of this project is that it uses an original dataset collected from the archives and teaching materials shared by various photography professors for this research project, including those that have taught photography courses at the Massachusetts Institute of Art, the Rochester Institute of Technology, and Harvard University. The training of photography students often involves looking at thousands of "great photographs" taken by hundreds of acclaimed art photographers, with thope that students can absorb the strong sesthetics shown in these images. The final cleaned dataset is comprised of \$11 artists with a total of over 20,000 photographs, which were originally meant to be used as training images for art students. Here, we use neural nets in the place of students.



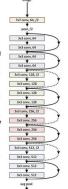




Model

The network architecture is based on a ResNet-18 architecture that starts with The network architecture is based on a ResNet-18 architecture that starts with pre-trained weights from the ImageNet dataset. The final fully-connected layer is replaced with a new layer to calculate a score for each artist in our dataset instead of a score for lmageNet classes. We first held the weights of our pre-trained network constant for 10 epochs, and then allowed all the weights throughout the entire network to update, training for an additional 10 epochs. During the process, we experimented with different regularization techniques including dropout, L2 regularization (weight decay), as well as changing the learning rate in the Adam Optimizer. An accuracy chart was made showing the first 10 epochs with only a reinitialized last layer. A second accuracy chart was made after the additional 10 epochs in fine tuning where all the weights throughout the network were allowed to update. On the right, we see a diagram of the Resnet-18 architecture based on [1] and [2].

Results: Confusion Matrix





An Initial Experiment was done with a small dataset of only 6 prominent black and white photographers, with over 250 images each in this dataset. These included the artists August Sander, Edward Weston, Eugene Atget, Julia Margaret Cameron, Man Ray and Walker Evans. The training set and test set were divided 70%/30%. A decision was made to combine what would normally be a 20% validation set and 10% test set, based on the aims of this project. Here, we present the final accuracy chart showing a final training accuracy of 95.97% and final test accuracy of 89.63%:







This project contributes to the research in art and deep learning in a variety of ways. Previous examinations on the topic have used the WikiArt dataset to examine this topic through painting, a medium that has a much longer history than photography, and one that is also very different [1][4]. Other studies that have used photography with deep learning do so with more popular photography datasets such as Flickr, but use "style" tags that are more commercial in nature and not based in artistic labels from the art world [3]. Using a unique dataset and domain specific knowledge, this project is able to extend the literature of art classification and deep learning to better understand the power of deep neural nets. While our model was able to achieve high accuracy when given a smaller dataset of artists to classify, more interesting results occurred when we expanded the range of artists, to understand how and what our models were learning. The results highlight how art and artists to create the present of the property of the present on contain a high degree of non-linearity and subtleness in its creation, appreciation and knowledge.

Discussion



Future work on this topic will include experimenting with other architectures such as Triplet Loss and other CNN architectures that can be applied to understanding art photography in different ways than just classification. I also plan on expanding and cleaning more data to add to the current dataset to see if there can be other more robust findings with more data. Despite many efforts to increase regularization with parameter tuning to minimize the gap between training and test accuracy in the model variations done with more artists, this would most likely be improved with more data and further experimentation.









60% 40%

A second set of experiments was done with the entire dataset, using variations of the basic model architecture, but varying the number of artists and photographs in each sample. Here, we present the final accuracy chart for the model when tested with 31 artists that had at least 100 photographs in each sample, as well as the corresponding confusion matrix to the right. A set of summary statistics is below showing the best results for each of the 5 model/dataset variations.

ResNet-18 with Transfer Learning (Fine Tuning of All Layers

Results: Summary Statistics

	Each Sample Size	Best Train	Best Tes
Model with 511 artists	> 5 photos	73.54	44.04
Model with 110 artists	> 40 photos	70.06	50.60
Model with 52 artists	> 70 photos	69.02	55.46
Model with 31 artists	> 100 photos	91.88	68,30
Model with 6 artists	> 250 photos	95.97	89.63
*Training (Test auto	F. 700/ /200/	days and day on the	

References: [1] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," arXiv preprint, abs/1512.03385, 2015. [2] https://www.kiagde.com/pytorch/resnet18 [3] S. Karayev, et al. "Recognizing Imag 2013, epinit arXiv:1311.3715 [4] WildArt. https://www.kiagde.com/pytorch/resnet18 [3] S. Karayev, et al. "Recognizing Imag 2013, epinit arXiv:1311.3715 [4] WildArt. https://www.kiagde.com/pytorch/resnet18 [3] S. Karayev, et al. "Recognizing Imag 2013, epinit arXiv:1311.3715 [4] WildArt. https://www.kiagde.com/pytorch/resnet18 [3] S. Karayev, et al. "Recognizing Imag 2013, epinit arXiv:1311.3715 [4] WildArt. https://www.kiagde.com/pytorch/resnet18 [3] S. Karayev, et al. "Recognizing Imag 2013, epinit arXiv:1311.3715 [4] WildArt. https://www.kiagde.com/pytorch/resnet18 [3] S. Karayev, et al. "Recognizing Imag 2013, epinit arXiv:1311.3715 [4] WildArt. https://www.kiagde.com/pytorch/resnet18 [3] S. Karayev, et al. "Recognizing Imag 2013, epinit arXiv:1311.3715 [4] WildArt. https://www.kiagde.com/pytorch/resnet18 [3] S. Karayev, et al. "Recognizing Imag 2013, epinit arXiv:1311.3715 [4] WildArt. https://www.kiagde.com/pytorch/resnet18 [3] S. Karayev, et al. "Recognizing Imag 2013, epinit arXiv:1311.3715 [4] WildArt. https://www.kiagde.com/pytorch/resnet18 [3] S. Karayev, et al. "Recognizing Imag 2013, epinit arXiv:1311.3715 [4] WildArt. https://www.kiagde.com/pytorch/resnet18 [3] S. Karayev, et al. "Recognizing Imag 2013, epinit arXiv:1311.3715 [4] WildArt. https://www.kiagde.com/pytorch/resnet18 [3] S. Karayev, et al. "Recognizing Imag 2013, epinit arXiv:1311.3715 [4] WildArt. https://www.kiagde.com/pytorch/resnet18 [3] S. Karayev, et al. "Recognizing Imag 2013, epinit arXiv:1311.3715 [4] WildArt. https://www.kiagde.com/pytorch/resnet18 [3] S. Karayev, et al. "Recognizing Imag 2013, epinit arXiv:1311.3715 [4] WildArt. https://www.kiagde.com/pytorch/resnet18 [3] S. Karayev, et al. "Recognizing Imag 2013, epinit arXiv:131115 [4] WildArt. https://www.kiagde.com/pytorch/resnet18 [4] S. WildArt. https