AlexNeterov: An Automated Art Historian

Debnil Sur
Department of Computer Science, Stanford University

Background

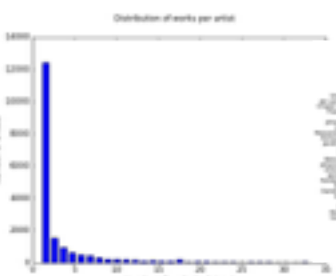
Identifying a painting's artist is necessary to establish a painting's cultural, historical, and economic value. It also represents the intersection of two vision challenges: scene understanding and cross-depiction. The number and variety of objects makes scene understanding difficult, and algorithms that work well on photographs are under-studied in different materials. Convolutional neural networks are primed to effectively answer both questions.

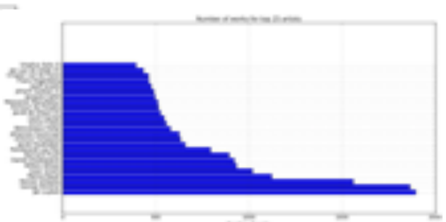
Problem Statement

Using convolutional neural networks, can we accurately identify the artist of a painting?

Data

The Rijksmuseum in Amsterdam released a digitized version of its collection (with an API) and created the Rijksmuseum Challenge in 2014 [1]. This dataset consists of 112,039 high-quality artwork images. The dataset is significantly right-skewed, so we stored the 13,564 images created by the 25 most-occurring artists in the dataset. A train/validation/test split of 70/10/20 was used. Images were down sampled and center cropped to 224 x 224 pixels for compatibility with PyTorch. The graphs below display the right-skew of the distribution of works per artist (left) and the most frequent artists' numbers of paintings.





Methods

We used four main approaches: (1) HOG with SVM, (2) transfer learning, (3) CNN with HOG, and (4) style and content embeddings. In all approaches, paintings were processed in batches of 50 to optimize time and memory.

Approach #1: Classical Classifier
To generate a conventional baseline, we generated HOG features for each batch of paintings and piped them into a 25-class logistic regression. This worked decently but suffered due to overfitting, even with L2-reg. The best performance was **69% train and 39% test** accuracy.

Approach #2: Transfer Learning
Transfer learning has worked well for using pre-trained image classification nets on object detection. Paintings present two challenges: (1) scenes, not objects, and (2) cross-depictional material. We trained a new classifier layer and then retrained the convolutional layers. Resnet18 performed excellently, with **99.6% train and 82.5% test** accuracy.

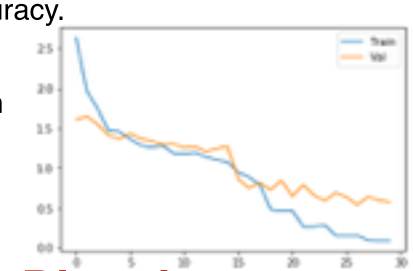
Approach #3: CNN with HOG
Can we combine conventional features and deep learning? Here, we concatenate a HOG feature vector with a painting's convolutional layer (from transfer learning) and re-train new fully connected layers. This model is **still training** but expected to be no better than Approach #2.

Approach #4: Content and Style Embeddings
Neural style transfer demonstrated the ability to separate a painting's content and style [2]. This can create style and content embeddings for paintings, which we can use to classify them by artist. The confusion matrices can hopefully help identify significant relationships between artists from their works. This model is **still under development** but expected to perform best, from the success of style transfer.

Results

We won the Rijksmuseum Challenge via transfer learning using ResNet18! These results have several insights for transfer learning using ResNet, which is rarely used in the literature.
1. Generalizability of ResNet: This significantly outperformed AlexNet (> 40% train accuracy). We will compare it with VGG-16, the most common in transfer learning.
2. Fewer layers are better: It also had much better test accuracy than ResNet50. More convolutional layers are harder to retrain.
3. Retraining helps with cross-depiction: Slightly retraining the convolutional weights significantly helped both train and test accuracy.

The train and validation losses (right) for pre-trained ResNet18. It's learning!



Future Directions

We will to test the effects of (1) concatenating a conventional HOG vector and (2) converting style and content transfer into embeddings for classification. Hopefully, the latter defeats a transfer learning approach. Future cross-depictional work can also focus on object detection, localization, and similar image tasks that have successfully used transfer learning.

References

[1] Mensink, Thomas, and Jan Van Gemert. "The rijksmuseum challenge: Museum-centered visual recognition." Proceedings of International Conference on Multimedia Retrieval. ACM, 2014.
[2] Girshick, Ross, et al. "Rich feature hierarchies for accurate object detection and semantic segmentation." Proceedings of the IEEE conference on computer vision and pattern recognition. 2014.
[3] Gatys, Leon A., Alexander S. Ecker, and Matthias Bethge. "Image style transfer using convolutional neural networks." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.