Determining Style of Paintings Using Deep Learning

Sean Chang, Jeffrey Chen, Patrick Mogan | {schang18, jchen623, pjmogan}@stanford.edu

Background

- Accurately identifying the style of an art piece requires a keen eye and a great deal of knowledge and experience. Furthermore, many art pieces have somewhat ambiguous styles. We decided to take the first steps in implementing a tool that can help inexperienced art enthusiasts classify the style of any given art piece.

Data

- We are using Kaggle's dataset "Painters By Numbers" to train, validate, and test our model [1]. This dataset consists of ~103,000 unique images of differently -sized colored paintings primarily from WikiArt.org, and their corresponding title, style, genre, artist, and date of creation. There are originally 125 styles accounted for throughout the dataset of which we extracted the 10 most common styles.
- The data was split into an 80/10/10 train, dev, and test set split. We resized images to 256 x 256 x 3 pixels and applied data augmentation to ensure that each class has roughly 8,000 training images.

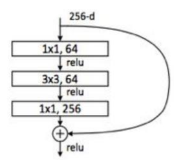
Models

- For our baseline, we used a simple three layer CNN to quickly get a model training and interpret results [2].
- For our subsequent models, we used a ResNet-50 model, which contains 16 three-layer residual blocks that allow for shortcut connections between layers and make it safer to train deeper models [3].

ResNet-50 Architecture

7×7, 64, stride 2	
3×3 max pool, stride 2	
$3 \times$	$\begin{Bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{Bmatrix}$
$4 \times$	$\begin{Bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{Bmatrix}$
$6 \times$	$\begin{Bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{Bmatrix}$
$3 \times$	$\begin{Bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{Bmatrix}$
Average pool	
1000-dimensional FC	
Softmax	

ResNet 50 residual block



- Additional testing focused heavily on hyperparameter turning, such as the number of residual blocks left unfrozen during training and dropout rate for fully connected layers.

Results

Validation Results

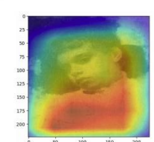
Model	Precision	Recall	F1 Score
Baseline*	0.075	0.024	0.023
Resnet-50: Transfer Learning*	0.475	0.472	0.466
Resnet-50: Transfer Learning	0.516	0.514	0.509
Resnet-50: 1 Residual Block Retrained	0.592	0.588	0.587
Resnet-50: 2 Residual Block Retrained	0.587	0.581	0.582
Resnet-50: 3 Residual Block Retrained	0.608	0.592	0.593
Resnet-50: 2 Residual Blocks Retrained and Dropout	0.595	0.59	0.587

*model trained without data augmentation

- Our best performing model was a ResNet-50 with three retrained residual blocks trained on augmented data. Results on the test data can be seen below:

Precision	Recall	F1 Score
0.604	0.593	0.595

- Class activation map shows whitespace most pertinent to classification, illustrating a potential model weakness against black and white images.



Discussion

- Our model did not perform as well as a panel of art experts would, but it performed comparably to models built for similar tasks [4].
- Most mistakes involved classifying a painting's style as one that is represented more frequently in the original dataset, suggesting the need to train on more paintings from underrepresented classes.
- Class activation maps give better insight into model performance and specific feature detection that cause incorrect classification of painting style (see Results) [5].
- Style is a large umbrella for classification, as style is often quite broad and paintings classified as the same style can differ drastically.

Future Work

- Reduce overfitting and improve performance with early stopping, weight decay, and increasing amount of training data for worse-performing classes.
- Test the model on and tailor it to other art classification tasks, such as year of creation or artist.

References

[1] Kaggle. *Painter by Numbers*. 2017. Web. 16 January 2019. https://www.kaggle.com/c/painter-by-numbers.
[2] perseus784. BvS. 2018. Web. 12 February 2019. https://github.com/perseus784/BvS.
[3] K. He, X. Zhang, S. Ren, and J. Sun. "Deep Residual Learning for Image Recognition." *Computer Vision and Pattern Recognition*. Dec, 2015. [Online]. Available: Arvix, arvix.org. Accessed March 16, 2019.
[4] A. Lecoutre, B. Nevgrevergne, and F. Yger, "Recognizing Art Style Automatically in painting with deep learning." *JMLR: Workshop and Conference Proceedings*. 2017. [Online]. Available: Univeriste Paris Dauphine. Accessed March 16, 2019.
[5] http://cs231n.stanford.edu/reports/2017/pdfs/410.pdf