Neural Recognition of Media of Art Pieces

Jane Boettcher [1] Shridhar Athinarayanan [1]

[1]Department of Computer Science, Stanford

Introduction and Problem

In the realm of art classification, most research has classified art pieces by artist and style, but there is little work on classification by medium.

We work to solve the problem of classifying artwork into its respective media. The input to our algorithm is any 224x224 image of artwork. We use a Deep Residual Neural Network, or ResNet model, paired with a binary classification system to output predicted classes from 10 medium description classes.

Dataset

Our final model is trained on a dataset we constructed of images from the Tate collection, one of the United Kingdom's largest national collections of British art. Since the images are of varying sizes, they were then center cropped to 224x224 pixels and converted to tensors with RGB channels and normalized them.

Methods

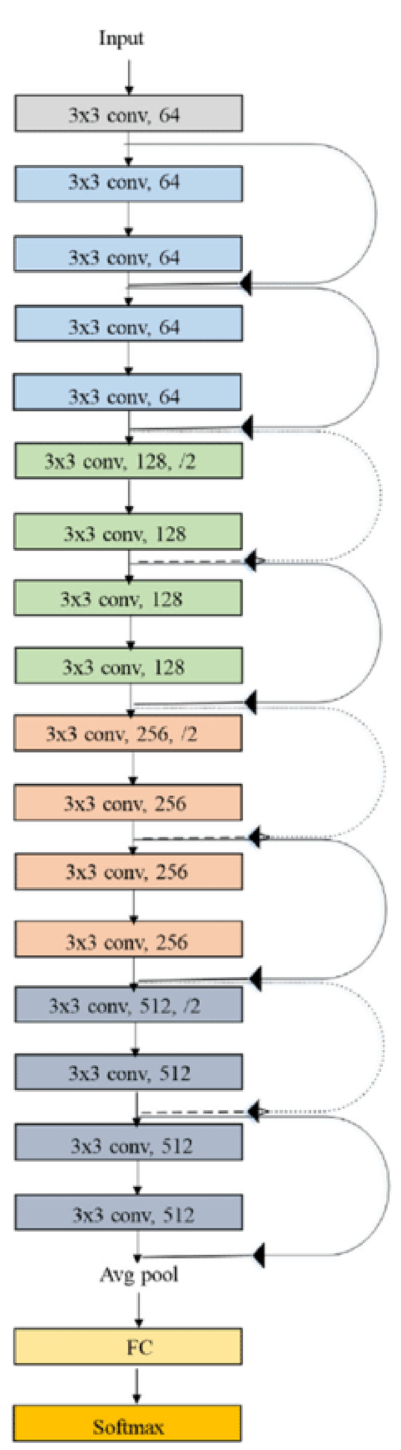

Figure 1. ResNet18 [1]

Simple CNN

Our baseline model BiDAF combines convolutional, relu, and maxpooling layers to create a simple CNN network shown below:

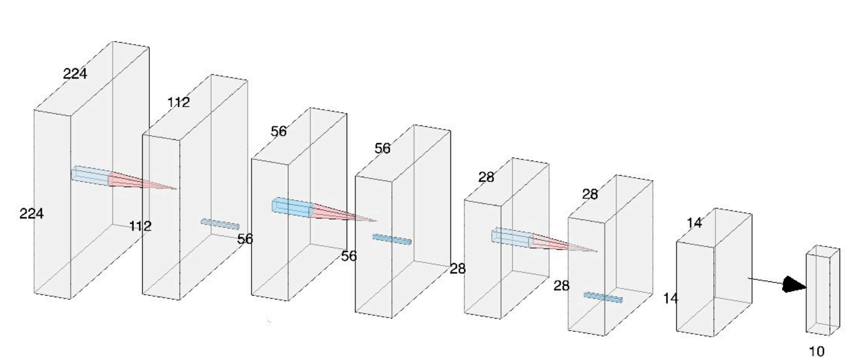

Figure 2. Simple CNN architecture

ResNet18

- ResNet18 = deep residual neural network with 18 deep layers (72 individual layers total).
- Multiple deep layers arises vanishing gradients issue
- ResNet uses residual blocks with skip connections, connecting past certain training layers, solves vanishing gradients [2].

Experiments

1. Baseline simple Convolutional Neural Network model (shown in Figure 2)
2. ResNet18 (Figure 1)
3. Augmented ResNet18 with data augmentation and frozen gradients
4. Multi-binary classifier system using ResNet18 (Medium description classes converted to lists of media tags (10), each tag with separate binary classifiers.)

We used the SGD Optimizer with a learning rates of 2e-3, 4e-3, 7e-3, and 2e-3 respectively.

	Baseline	ResNet18	ResNet18 (Aug)	Binary
Train	.7454	1	1	1
Validation	.5700	.7660	.7760	.4760
Test	.5650	.7430	.7490	.4770

Performance of various models on train, validation, and test sets

Analysis

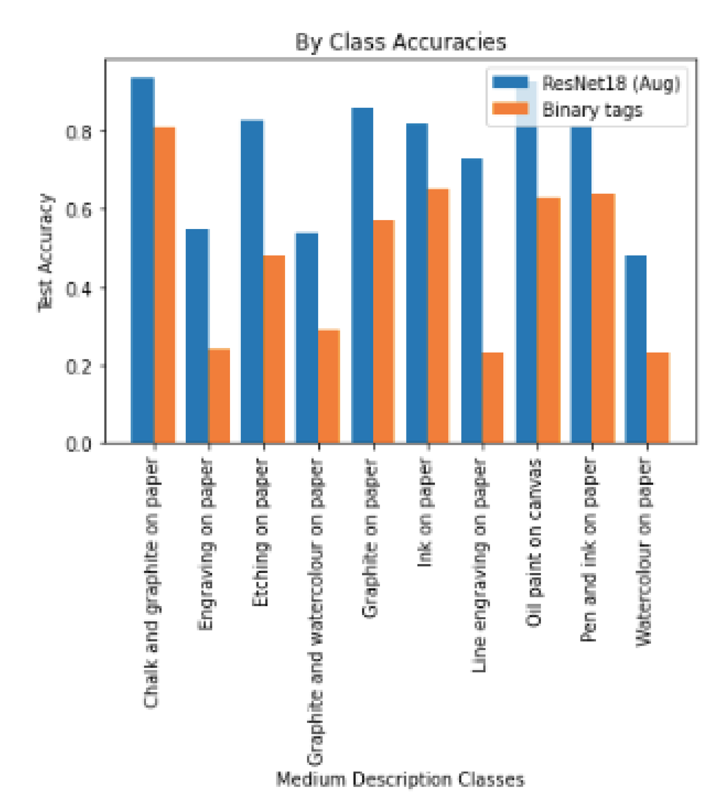

Figure 3. By medium description class test accuracies

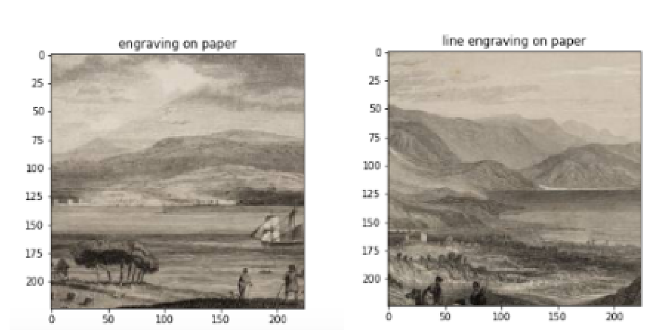

Figure 4. Engraving and line engraving images that don't seem discernibly different to the non-domain expert human eye. Both were tagged by the binary classifiers as both 'line engraving' and 'engraving (non-line)' media

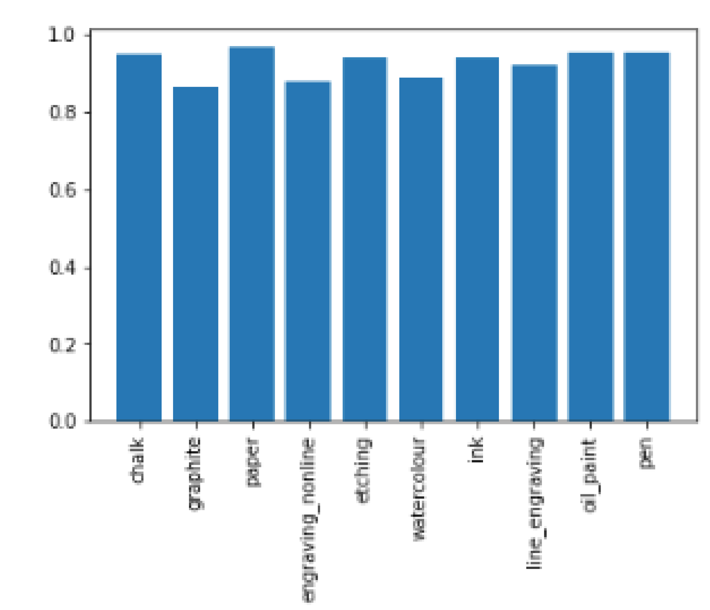

Figure 5. By tag binary classifier test accuracies

- Individual binary tag classification accuracy high (Figure 5), but overall low due to 2^{10} possibilities (Figure 3)
- Confusion matrix (Figure 6) reflects the problem that categorizing works into mixed media categories can pose
- i.e. 'Graphite and watercolour on paper' works most often get confused with 'graphite on paper' and 'watercolour on paper' works

Analysis cont.

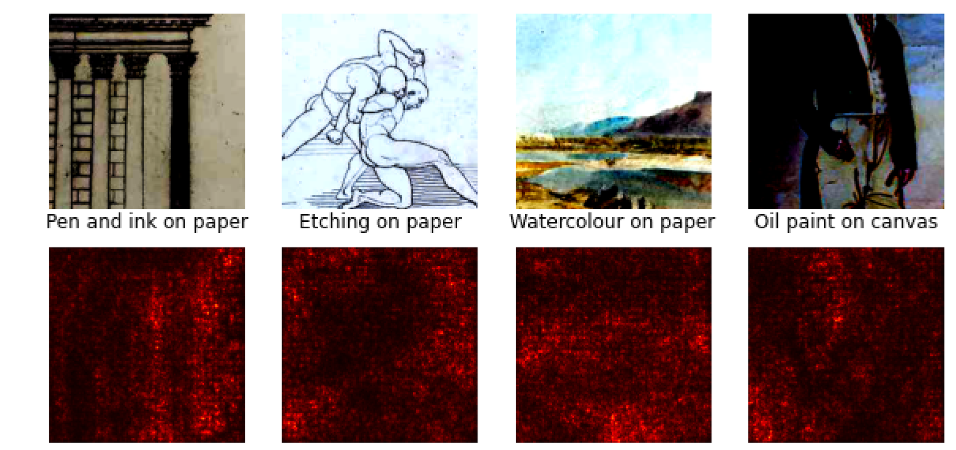

Figure 6. Saliency Maps for 4 different media

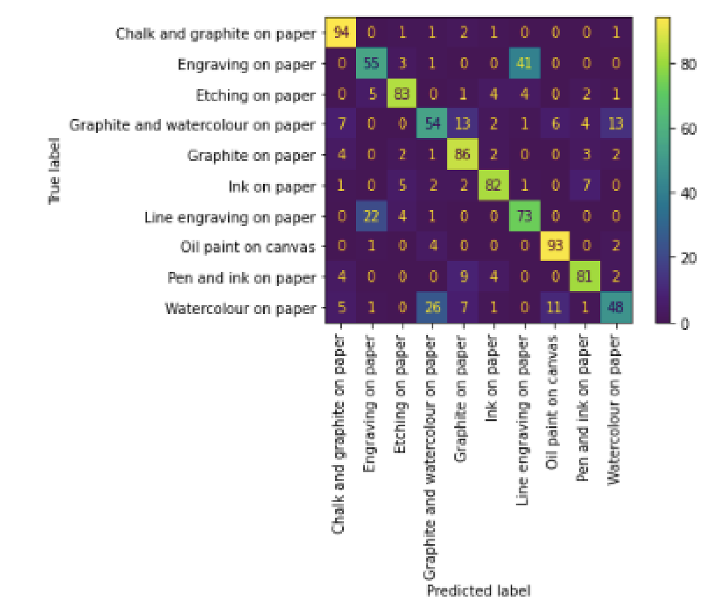

Figure 7. Confusion Matrix for 10 classes

- Classes like pen and ink on paper and etching on paper, saliency (red markings) is found in the blank area of the artwork
- Most likely due to these classifications being differentiated by their lack of color and sparsity of marking on paper
- Red in the saliency map is most present where there is filled in color in the watercolor and oil painting images

Conclusions and Future Work

- ResNet18 with data augmentation and frozen gradients yielded the highest test accuracy due to its robust layer architecture with skip connections
- Binary classification performed well for individual tag classification
- We were able to pinpoint where misclassification occurred as well as understand the reasonings for valid classifications.
- With machine classification, museum curators can automate the tagging of pieces with their associated metadata.
- Future work: engage more robust models and tackle style transfer, extending the per-pixel loss function from the saliency maps.

References

Poster template provided by Rylan Schaeffer (Stanford-LaTeX-Poster-Template. https://github.com/RylanSchaeffer/Stanford-LaTeX-Poster-Template)

[1] Ramzan, F., Khan, M. U., Rehmat, A., Iqbal, S., Saba, T., Rehman, A., amp; Mehmood, Z. (2019). A deep learning approach for automated diagnosis and Multi-class classification of alzheimer's disease stages using resting-state fmri and residual neural networks. Journal of Medical Systems, 44(2). https://doi.org/10.1007/s10916-019-1475-2

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