Learning More Robust Neural Network Representations with Paintings of Materials and Objects

Hubert Lin[1], Mitchell Van Zuijlen[2], Maarten W.A. Wijntjes[2], Sylvia C. Pont[2], Kavita Bala[1]
[1]*Cornell University* [2]*Delft University of Technology*

Motivation

Style transfer can create painting-like images, but real artist-created paintings are not simply a style filter applied to photos. Recent work has shown neural networks trained on stylized images can be more robust. Do real paintings affect network robustness similarly or differently?

Summary of Findings

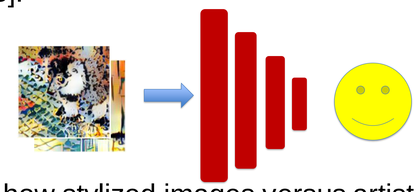
- Paintings and stylized images capture complementary invariances with respect to common image corruptions and novel viewpoints.
- Artforms like sketches or cartoons are unable to improve robustness to same extent as paintings.
- Stylization relies on high-frequency signals to improve robustness towards noise.
- Stylization does not require painting-like styles to be transferred to improve robustness (not shown here, see [1]).

Background

Photos captured in real-world settings often contain noise and other corruptions which reduce the quality of the final image. Examples of corruptions like noise, blur, weather, and digital manipulations [2]:

While denoising techniques or improved image capture systems may be used to improve the quality of images for downstream tasks, an alternative approach is to improve the robustness of the computer vision systems applied to such images.

Recent work has shown image stylization as a form of data augmentation can improve the robustness of neural network based computer vision systems [3].



In this work, we compared how stylized images versus artist-created paintings improve network robustness.

Stylized Photo of Giraffe versus Artist-Created Painting of Giraffe

Experimental Setup

We trained ResNet classifiers on photos, and explored how replacing a proportion of photos with stylized photos or artist-created paintings instead impacts model robustness.

Robustness is measured by accuracy on:
(a) Photos corrupted by common corruptions [2] (see figure on left)
(b) Photos from a different dataset, which represents changes in distribution over viewpoints or background context.

Experiments on two datasets (material classification and object classification) were conducted.

Findings

For full details, please see our paper [1]. Here, we briefly highlight a small number of our findings.

Paintings and stylized images are complementary.

Training Data	MEAN	Corruption Robustness	View/Background Robustness
Photos-Only	48.03 | 79.37	54.73 | 76.16	41.33 | 82.57
+ Stylized Images	48.61 | 82.35	62.67 | 87.27	34.54 | 77.43
+ Paintings	50.92 | 82.21	57.92 | 78.99	43.92 | 85.43
+ Both	**51.49 | 86.32**	61.47 | 87.31	41.50 | 85.33

"*x* | *y*" indicates results for material classification and object classification respectively (different datasets).

Paintings improve robustness to both corruptions and viewpoint shifts; stylized images greatly improve robustness to corruptions but harms viewpoint robustness. Using both results in greater robustness overall over either alone.

Art forms like sketches and cartoons do not improve robustness to the same extent as paintings.

Training Data (fixed total number)	Corruption Robustness
Photos-Only	54.73 | 76.16
+ Paintings	**56.31 | 79.83**
+ Cartoons	- | 75.51
+ Sketches	- | 73.78

Paintings offer a fine balance of realism and abstraction. Sketches and cartoons are too abstract / stylized, which harms the model's ability to learn cues that are useful for recognizing objects in real photos.

Stylized images improve robustness to noise through invisible high-frequency signals.

Training Data	Noise Robustness	
Photos-Only	43.71 | 62.64	
+ Stylized Images	61.87 | 85.98	-16.05 | -8.43
+ Stylized Images (low freq)	45.82 | 77.55	
+ Paintings	49.82 | 68.04	-4.87 | +3.12
+ Paintings (low freq)	44.95 | 71.16	

Image stylization injects imperceptible high-frequency signals that greatly improve noise robustness; removing these signals with a low-pass filter impacts the effect of image stylization more than paintings.

References

[1] Lin et al, *What Can Style Transfer and Paintings Do For Model Robustness?*
[2] Hendrycks and Ditterich, *Benchmarking Neural Network Robustness to Common Corruptions and Perturbations*
[3] Geirhos et al, *ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness*

Acknowledgements

This work was supported in part by NSF (CHS-1617861 and CHS-1513967), NSERC (PGS-D 516803 2018), and the Netherlands Organization for Scientific Research (NWO) project 276-54-001