What Can Style Transfer and Paintings Do For Model Robustness?

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

CVPR VIRTUAL JUNE 19-25

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 model robustness similarly or differently?

Summary of Findings

- Style transfer improves model robustness, but not necessarily through real painting styles.
- Real paintings and stylized images capture complementary invariances.
  - Real paintings improve robustness to corruptions and novel viewpoints
  - Stylization greatly improves robustness to corruptions but harms novel viewpoint generalization
- Real paintings improve robustness cost-effectively.
- Other art forms, such as sketches, cartoons, and untextured renderings are unable to improve model robustness to the same extent as paintings.

Background

CNNs struggle to generalize well in settings where humans are relatively capable. For example, noisy images or images taken in new environments can cause high CNN misclassification rates.

Paintings are perceptually realistic but not physically realistic – artworks encode some invariances of the human perception system, which can be useful for networks to learn.

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

Stylized Photo of Giraffe versus Artist-Created Painting of Giraffe

Experimental Setup

Model robustness is measured via accuracy on:
(a) Photos corrupted by common image corruptions
(b) Photos from a different dataset, which encapsulates changes in distribution over viewpoints or background context.

Experiments on two datasets (material classification [Materials] and object classification [PACS]) were conducted.

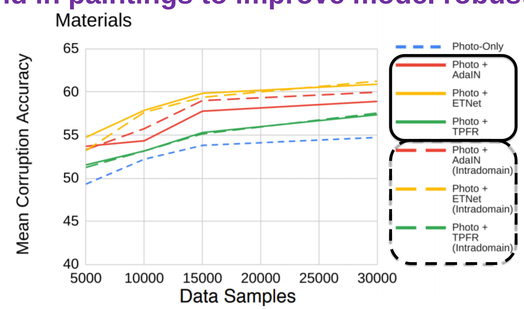
Acknowledgements

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

Findings

Here, we present a selection of our findings.
[H#] corresponds to hypothesis # in the paper.
For full results, please see our paper!

Style transfer does not necessarily rely on styles found in paintings to improve model robustness. [H1]



Images stylized with other photos (dashed lines) results in similar robustness to images stylized with paintings (solid lines). This suggests that style transfer acts independently from painting styles to improve model robustness.

Stylized images improve robustness to noise through invisible high-frequency signals. [H6]

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	

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

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.

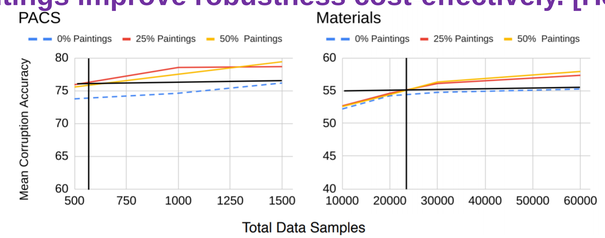
Paintings and stylized images are complementary. [H5]

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.

Paintings improve robustness cost-effectively. [H3]



It is better to annotate some paintings and photos instead of only photos.

Art forms like sketches and cartoons do not improve robustness to the same extent as paintings. [H4]

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

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

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.