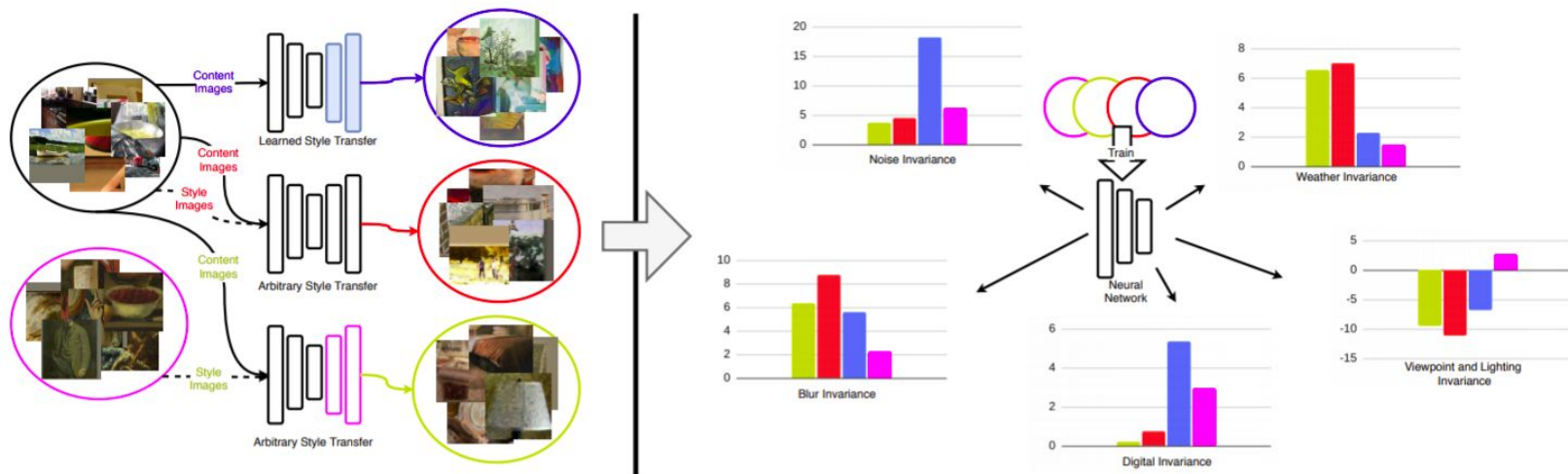


What Can Style Transfer and Paintings Do For Model Robustness?

Hubert Lin¹, Mitchell Van Zuijlen², Maarten W.A. Wijntjes², Sylvia C. Pont², Kavita Bala¹



Background

“What Can **Style Transfer and Paintings** Do For **Model Robustness**?”

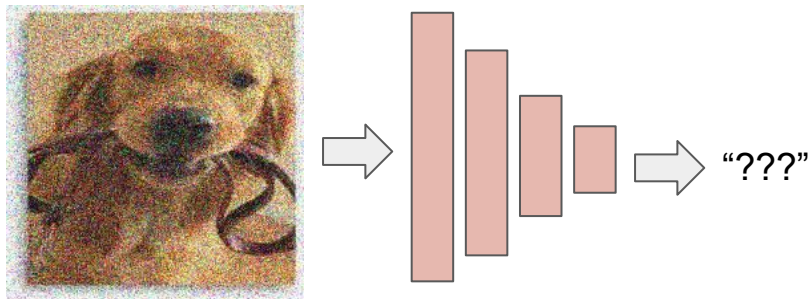
What does it mean for a neural network to be robust?

Why should we expect paintings to be useful for learning robust models?

Background

In real world settings, images may be noisy, blurry, or digitally altered, or be taken from different viewpoints unlike those during training.

- Humans can still recognize objects¹ and materials² in such images.
- However, neural networks struggle in these settings.



¹Geirhos et al, Generalisation in Humans and DNNs

²Sharan et al, Accuracy and speed of material categorization in real-world images

Background

One solution is “Data Augmentation”:

- Carefully chosen transformations can encourage the network to ignore certain factors in the data.

E.g., left-right reflection to ignore left-right orientation of animals:



Background

In reality, it is infeasible to list all of the transformations to which a visual recognition model should be invariant.

Observation:

Perceptually realistic artworks implicitly encode invariances of the human visual perception system.¹

Artworks can be considered a form of data augmentation which corresponds to (some) invariances of the human visual system.

¹ c.f. Cavanagh, Gombrich, Mamassian, etc.

Background

What does it mean for a neural network to be robust?

Generalization to noisy images or images from novel viewpoints, where humans are relatively robust.

Why should we expect paintings to be useful for learning robust models?

Implicit encoding of invariances of the robust human visual perception system.

Research Question

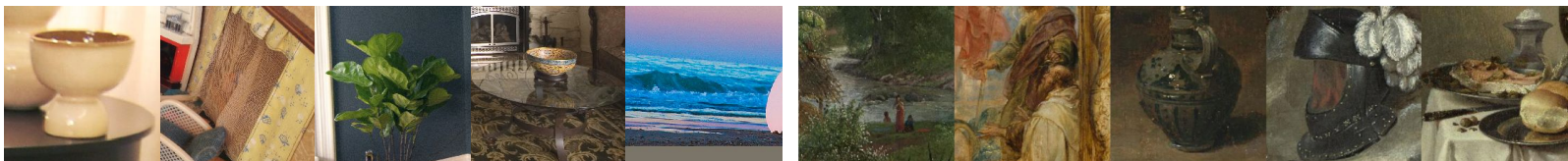
Instead of real paintings, recent work has explored the use of ‘fake paintings’ created via style transfer.¹ In this work, we explore:

- To what extent does style transfer capture the invariances encoded in real paintings?
- What are the different invariances learned by models trained on stylized images versus real paintings?

Datasets

Materials:

- Photographs of materials from existing datasets (MINC¹, COCO²)
- Paintings of materials from Materials in Paintings (MIP³)



Objects:

- Existing dataset of photos, paintings, cartoons, and sketches (PACS⁴).



Evaluating Robustness

Accuracy with respect to common image corruptions¹:



¹ Hendrycks and Dietterich

Evaluating Robustness

Accuracy with respect to out-of-distribution photos (different viewpoint, lighting):

- Materials → FMD¹



- PACS → Subset of YFCC100M²



¹ Sharan et al ² Thomee et al

Experiment: Role of Painting Styles in Style Transfer



Painting styles transferred onto Photos via AdaIN¹

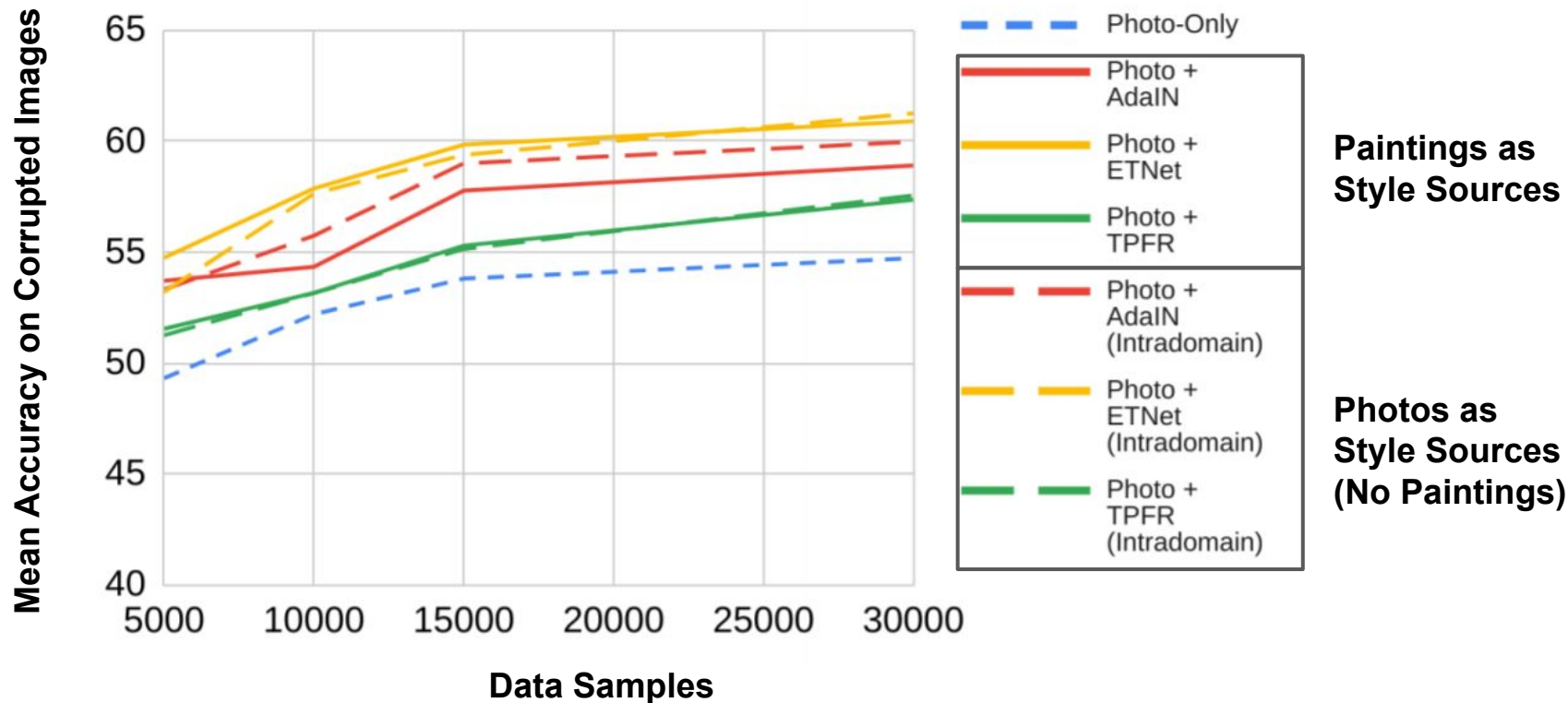
Hypothesis:

Stylization improves model robustness by capturing styles found in paintings.

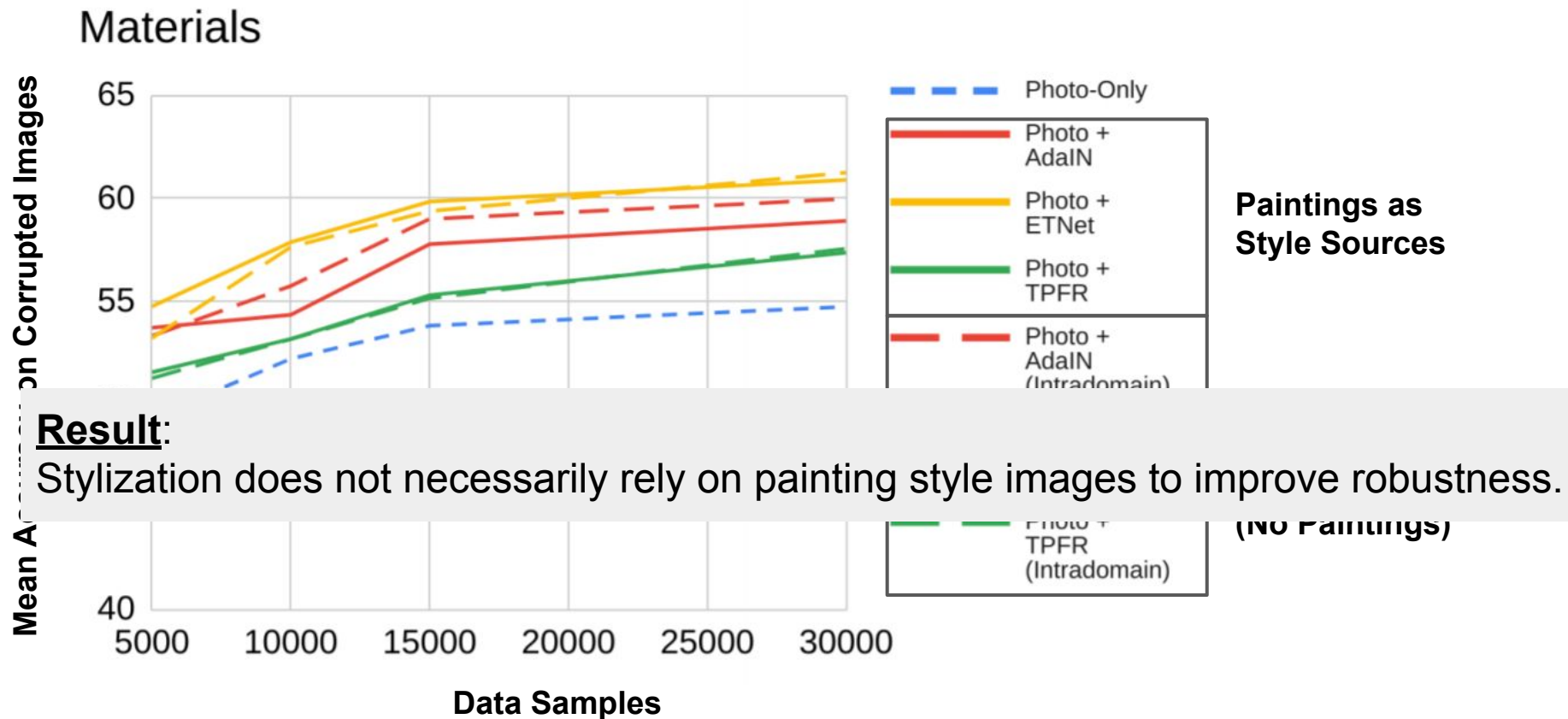
¹ Huang and Belongie

Experiment: Role of Painting Styles in Style Transfer

Materials

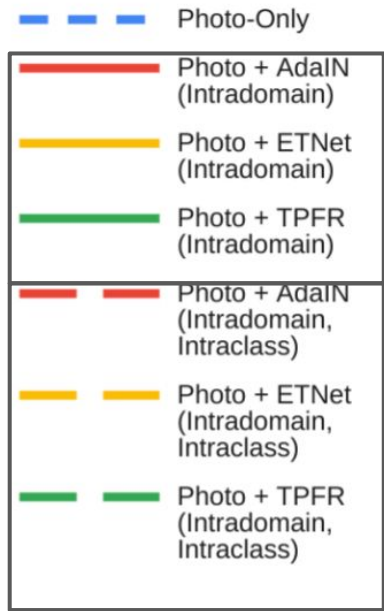
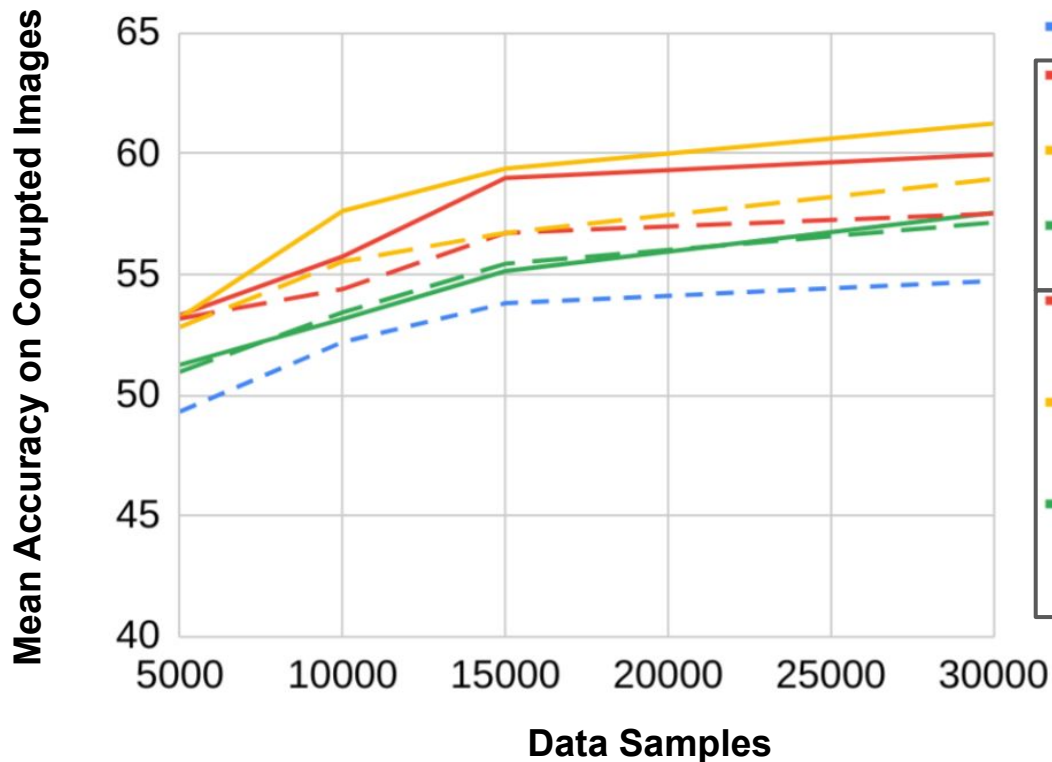


Experiment: Role of Painting Styles in Style Transfer



Experiment: Role of Painting Styles in Style Transfer

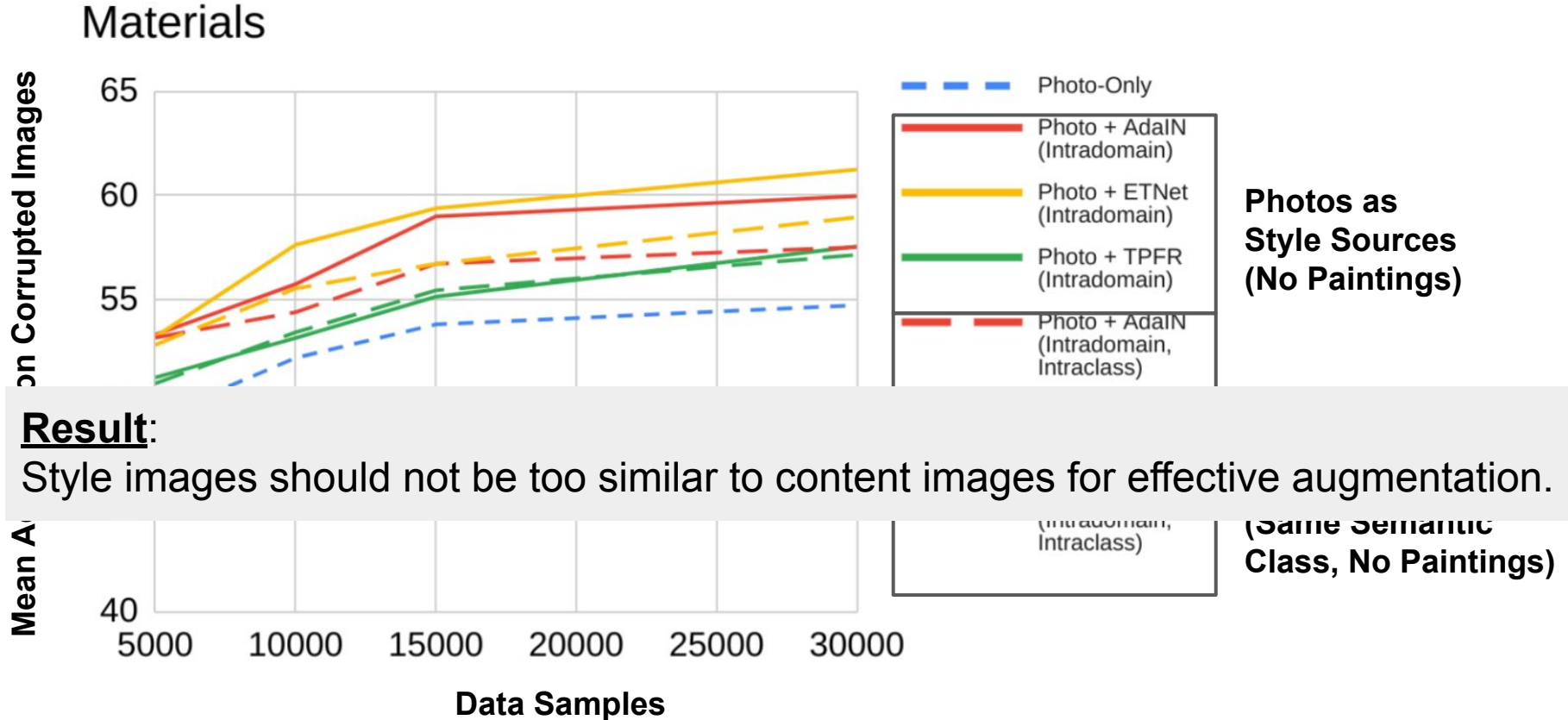
Materials



Photos as Style Sources (No Paintings)

Photos as Style Sources (Same Semantic Class, No Paintings)

Experiment: Role of Painting Styles in Style Transfer



Experiment: Role of Painting Styles in Style Transfer

- Both photo and paintings as sources of styles can result in similar stylized images.
- These stylized images can improve robustness in similar ways.
- However, if the style photos are “too similar” to the content photos, the gains in robustness are far smaller.
 - “Too similar” = “Same semantic content”
 - A more quantitative measure of similarity that correlates to robustness would be interesting.
 - Similarity measured via style distances such as Gram Matrix distance did not correlate with gains in robustness.

Experiment: Learning Directly from Paintings

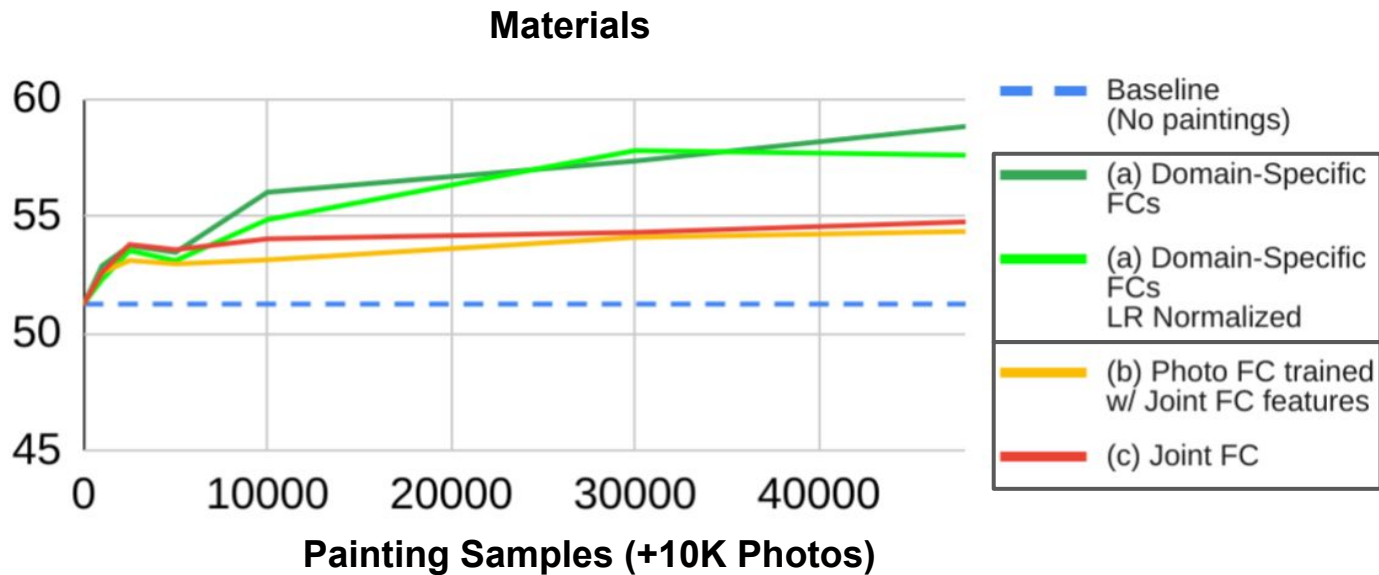


Hypothesis:

Learning directly from paintings can improve robustness, but domain shift and annotation costs may be issues.

Experiment: Learning Directly from Paintings

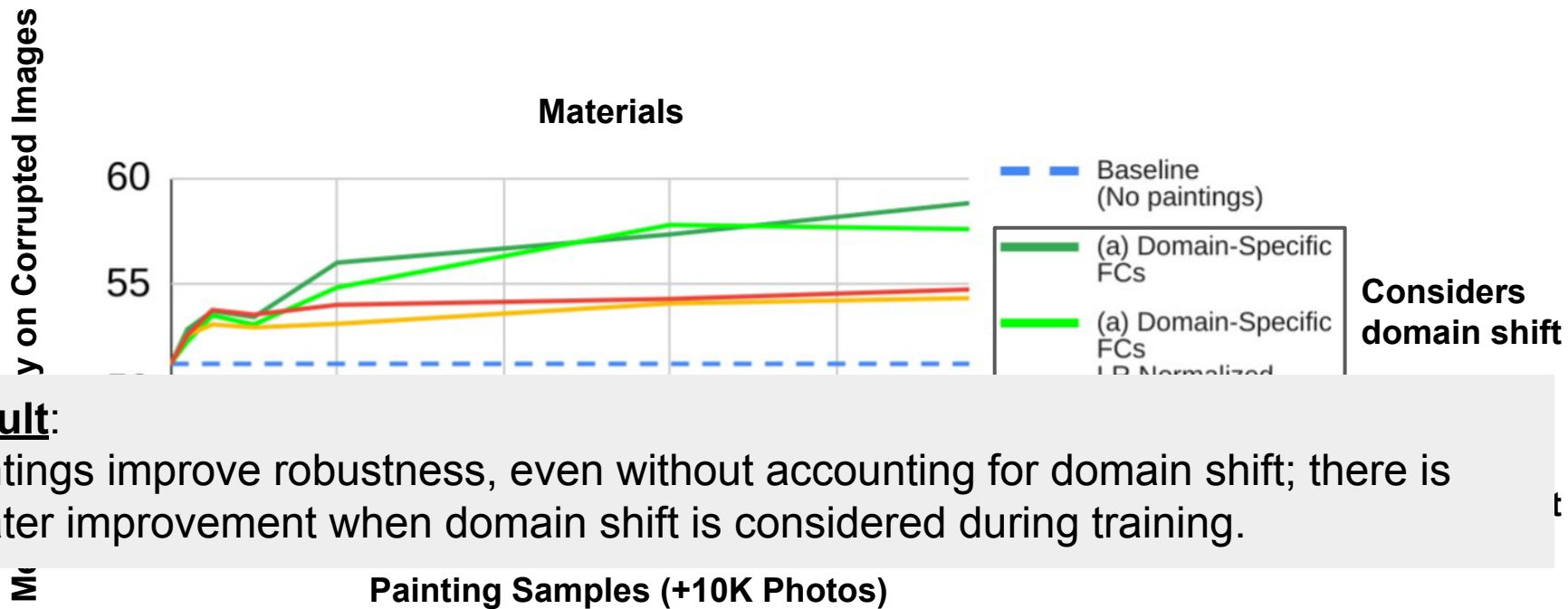
Mean Accuracy on Corrupted Images



Considers domain shift

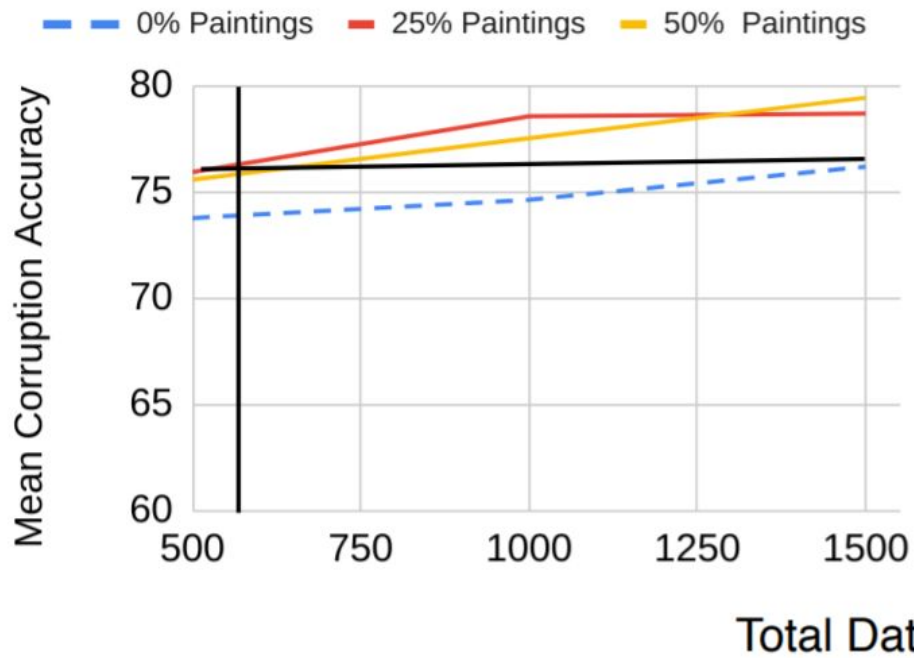
Does not consider domain shift

Experiment: Learning Directly from Paintings

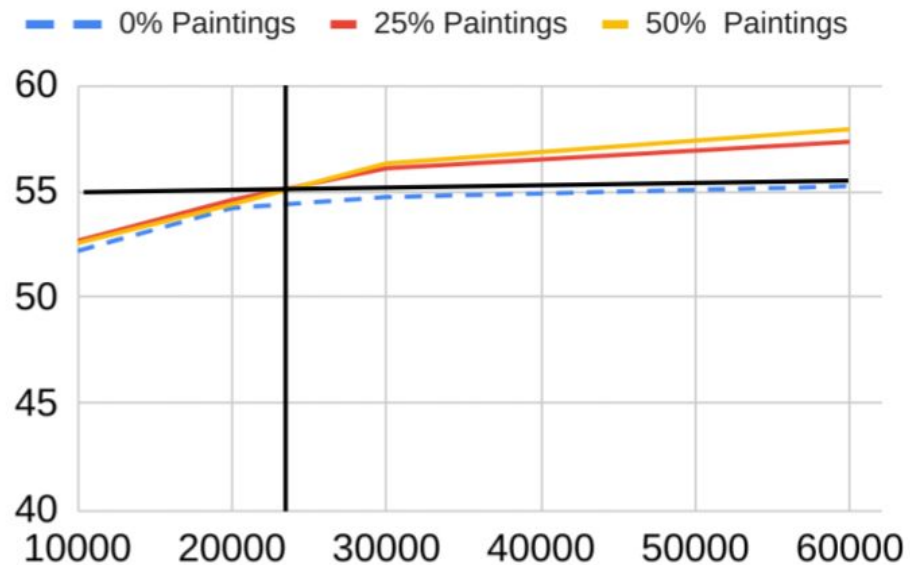


Experiment: Learning Directly from Paintings

PACS

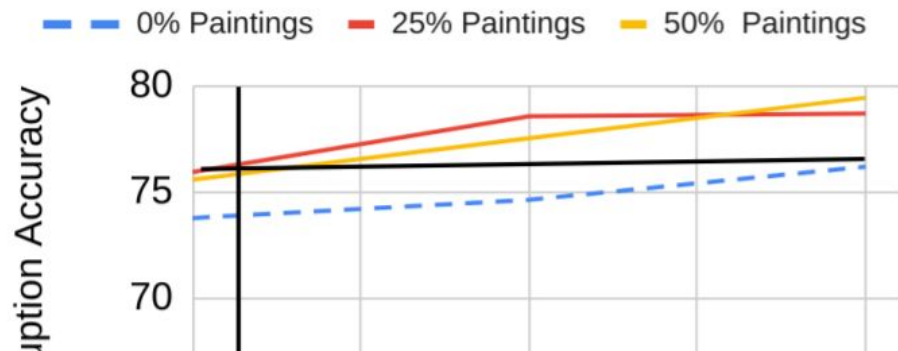


Materials

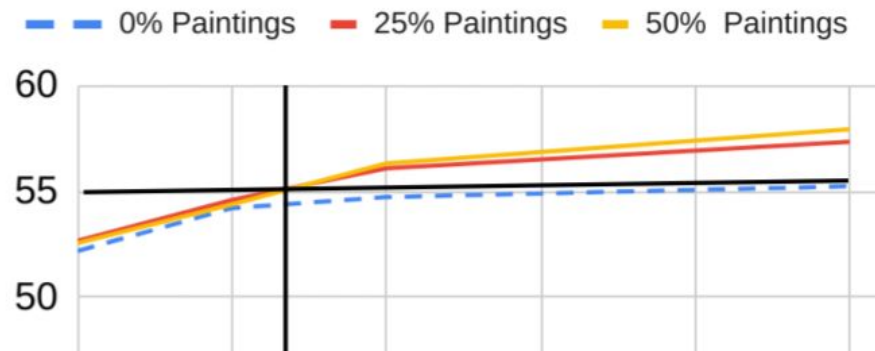


Experiment: Learning Directly from Paintings

PACS



Materials



Result:

Paintings are cost-effective for improving robustness; it is good to mix both photos and paintings.

Total Data Samples

Experiment: Learning Directly from Paintings

- Learning from paintings directly improves robustness, even without accounting for domain shift during training.
- Paintings improve robustness more so than an equivalent number of photos (assuming a sufficiently large number of training photos already exist).

Experiment: Stylized Images vs Paintings



Learned Style Transfer

Paintings

Hypothesis:

Paintings and stylized photos can improve robustness in different ways.

Experiment: Stylized Images vs Paintings

Corr. = Image Corruptions

OOD = Novel environment (photos from different dataset)

Method	MEAN	Corr.	OOD	
<i>Materials (30K Samples/Domain)</i>				
Photo-Only	48.03±0.21	54.73±0.25	41.33±0.62	Photos (No stylization)
Photo + SACL	48.56±0.45	62.67±0.03	34.54±0.91	Learned Style Transfer
Photo + Painting	50.92±0.22	57.92±0.09	43.92±0.47	
Photo + SACL + Painting	51.49±0.69	61.47±0.50	41.50±1.38	
<i>PACS (1.5K Samples/Domain)</i>				
Photo-Only	79.37±0.17	76.16±0.34	82.57±0.00	Photos (No stylization)
Photo + SACL	82.35±0.37	87.27±0.10	77.43±0.84	Learned Style Transfer
Photo + Painting	82.54±0.59	79.65±0.49	85.43±0.70	
Photo + SACL + Painting	85.42±0.18	87.31±0.30	83.52±0.27	

Experiment: Stylized Images vs Paintings

Corr. = Image Corruptions

OOD = Novel environment (photos from different dataset)

Method	MEAN	Corr.	OOD
<i>Materials (30K Samples/Domain)</i>			
Photo-Only	48.03±0.21	54.73±0.25	41.33±0.62
Photo + SACL	48.56±0.45	62.67±0.03	34.54±0.91
Photo + Painting	50.92±0.22	57.92±0.09	43.92±0.47
Photo + SACL + Painting	51.49±0.69	61.47±0.50	41.50±1.38

Photos (No stylization)

Learned Style Transfer

Result:

Stylization improves robustness to image corruptions, but hurts viewpoint generalization.

Photo + SACL + Painting	85.42±0.18	87.31±0.30	83.52±0.27
-------------------------	-------------------	-------------------	------------

Experiment: Stylized Images vs Paintings

Corr. = Image Corruptions

OOD = Novel environment (photos from different dataset)

Method	MEAN	Corr.	OOD	
<i>Materials (30K Samples/Domain)</i>				
Photo-Only	48.03±0.21	54.73±0.25	41.33±0.62	Photos (No stylization)
Photo + SACL	48.56±0.45	62.67±0.03	34.54±0.91	
Photo + Painting	50.92±0.22	57.92±0.09	43.92±0.47	Paintings
Photo + SACL + Painting	51.49±0.69	61.47±0.50	41.50±1.38	
<i>PACS (1.5K Samples/Domain)</i>				
Photo-Only	79.37±0.17	76.16±0.34	82.57±0.00	Photos (No stylization)
Photo + SACL	82.35±0.37	87.27±0.10	77.43±0.84	
Photo + Painting	82.54±0.59	79.65±0.49	85.43±0.70	Paintings
Photo + SACL + Painting	85.42±0.18	87.31±0.30	83.52±0.27	

Experiment: Stylized Images vs Paintings

Corr. = Image Corruptions

OOD = Novel environment (photos from different dataset)

Method	MEAN	Corr.	OOD	
<i>Materials (30K Samples/Domain)</i>				
Photo-Only	48.03±0.21	54.73±0.25	41.33±0.62	Photos (No stylization)
Photo + SACL	48.56±0.45	62.67±0.03	34.54±0.91	
Photo + Painting	50.92±0.22	57.92±0.09	43.92±0.47	Paintings
Photo + SACL + Painting	51.49±0.69	61.47±0.50	41.50±1.38	

Result:

Paintings boost both image corruption and viewpoint generalization.

Photo + Painting	82.54±0.59	79.65±0.49	85.43±0.70	Paintings
Photo + SACL + Painting	85.42±0.18	87.31±0.30	83.52±0.27	

Experiment: Stylized Images vs Paintings

Corr. = Image Corruptions

OOD = Novel environment (photos from different dataset)

Method	MEAN	Corr.	OOD	
<i>Materials (30K Samples/Domain)</i>				
Photo-Only	48.03±0.21	54.73±0.25	41.33±0.62	Photos (No stylization)
Photo + SACL	48.56±0.45	62.67 ±0.03	34.54±0.91	
Photo + Painting	50.92±0.22	57.92±0.09	43.92 ±0.47	
Photo + SACL + Painting	51.49 ±0.69	61.47±0.50	41.50±1.38	Paintings + Stylized Imgs
<i>PACS (1.5K Samples/Domain)</i>				
Photo-Only	79.37±0.17	76.16±0.34	82.57±0.00	Photos (No stylization)
Photo + SACL	82.35±0.37	87.27±0.10	77.43±0.84	
Photo + Painting	82.54±0.59	79.65±0.49	85.43 ±0.70	
Photo + SACL + Painting	85.42 ±0.18	87.31 ±0.30	83.52±0.27	Paintings + Stylized Imgs

Experiment: Stylized Images vs Paintings

Corr. = Image Corruptions

OOD = Novel environment (photos from different dataset)

Method	MEAN	Corr.	OOD	
<i>Materials (30K Samples/Domain)</i>				
Photo-Only	48.03±0.21	54.73±0.25	41.33±0.62	Photos (No stylization)
Photo + SACL	48.56±0.45	62.67±0.03	34.54±0.91	
Photo + Painting	50.92±0.22	57.92±0.09	43.92±0.47	
Photo + SACL + Painting	51.49±0.69	61.47±0.50	41.50±1.38	Paintings + Stylized Imgs

Result:

Stylized images and paintings improve robustness in a complementary manner.

Photo + Painting	82.54±0.59	79.65±0.49	85.43±0.70	
Photo + SACL + Painting	85.42±0.18	87.31±0.30	83.52±0.27	Paintings + Stylized Imgs

Experiment: Stylized Images vs Paintings

LF = low frequency images, i.e., high frequency signals filtered out.



Method	Noise
<i>Materials (30I)</i>	
Photo-Only	43.70
Photo+SACL	61.87
Photo+Painting	49.82
Photo+SACL (LF)	45.82
Photo+Painting (LF)	44.95
<i>PACS (1.5K)</i>	
Photo-Only	62.64
Photo+SACL	85.98
Photo+Painting	68.04
Photo+SACL (LF)	77.55
Photo+Painting (LF)	71.16

Experiment: Stylized Images vs Paintings

LF = low frequency images, i.e., high frequency signals filtered out.



Method	Noise
<i>Materials (30I)</i>	
Photo-Only	43.70
Photo+SACL	61.87
Photo+Painting	49.82
Photo+SACL (LF)	45.82
Photo+Painting (LF)	44.95
<i>PACS (1.5K)</i>	
Photo-Only	62.64
Photo+SACL	85.98
Photo+Painting	68.04

Result:

Stylized images rely on high-frequency signals to improve noise robustness.

Experiment: Stylized Images vs Paintings

- Models learn complementary invariances from stylized images and paintings.
- Paintings improve robustness to both corruptions and novel viewpoints.
- Stylization greatly improves robustness to corruptions while harming generalization to new views.
- Stylization improves robustness to noise corruptions through injecting visually imperceptible signals

Conclusions

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

Thank you!

For details, please check out our paper and our poster.

What Can Style Transfer and Paintings Do For Model Robustness? CVPR 2021

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

