Style Augmentation: Data Augmentation via Style Randomization

Philip T. Jackson, Amir Atapour-Abarghouei, Stephen Bonner, Toby Breckon, Boguslaw Obara

Durham University

Randomized style transfer can be used to remove the correlation between texture and class label, resulting in increased robustness to domain shift.

Motivation:

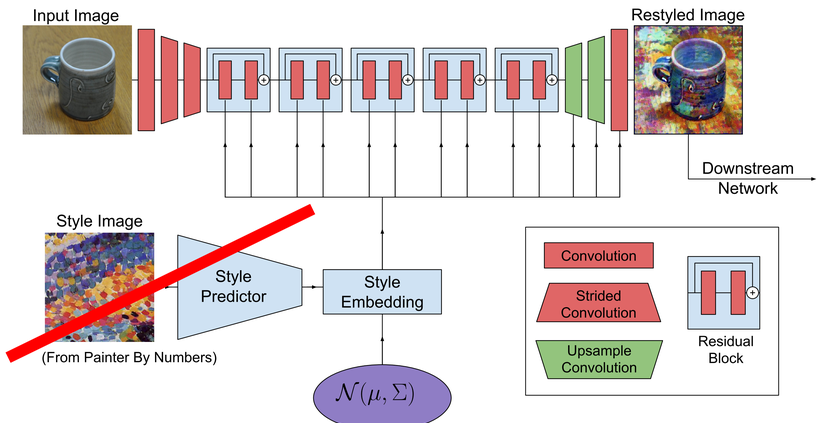
- **Domain randomization** has been shown to improve generalization between domains [1].
- **Neural style transfer** has been been linked theoretically to **domain shift** [2] and has been used successfully for **domain adaptation** [3].
- ImageNet-trained CNNs have been shown to rely heavily on texture at the expense of shape [4].

What if instead of using domain adaptation, we randomize the style of the source domain, so that the model becomes robust against domain shift in the first place?

[1] Tobin et al., 'Domain Randomization for Transferring Deep Neural Networks from Suimulation to the Real World', 2017
[2] Li et al., 'Demystifying neural style transfer.' arXiv preprint arXiv:1701.01036, 2017.
[3] Atapour et al., 'Real-Time Monocular Depth Estimation using Synthetic Data with Domain Adaptation via Image Style Transfer'
[4] Geirhos et al., 'ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness'
[5] Saenko et al., 'Adapting visual category models to new domains', 2015

Proposed Approach:

We use a modified **arbitrary style transfer** network to randomly restyle training images on the fly, before passing them to a downstream network for training.

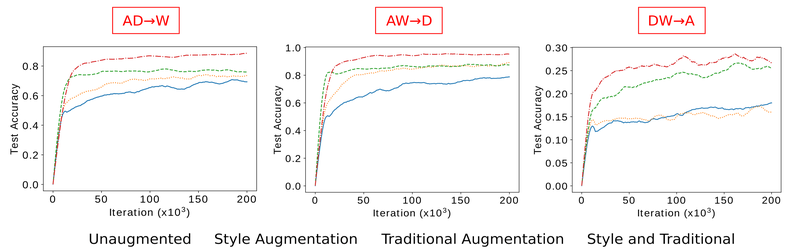


Normal distribution with mean and covariance matched with those of PBN embeddings

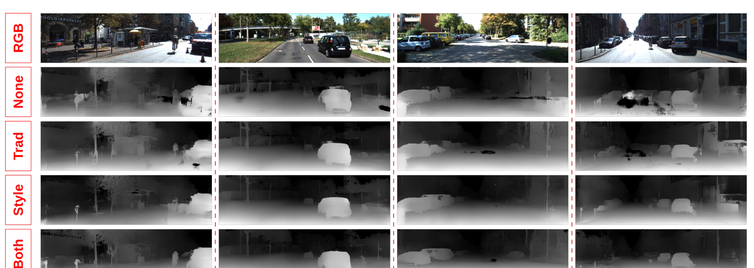
- To randomize style transfer, we need only randomize the style embedding.
- Instead of computing style embeddings from random style images, we simulate this process by sampling embeddings directly from a normal distribution. This is faster and ensures a unique embedding every time.

Results:

Our approach produces substantial accuracy boosts in both classification and monocular depth estimation tasks when domain shift is present.



Test accuracy on the Office dataset [5], training on two domains, testing on the other. A,D,W refer to the Amazon, DSLR and Webcam domains, respectively. ``Traditional Augmentation'' refers to a combination of seven standard augmentation techniques. Network is InceptionV3.



Style augmentation reduces artifacts in monocular depth estimation when training on GTA-V data and testing on KITTI.

Code and models available here:

https://github.com/philipjackson/style-augmentation