Computer Vision Project: Paper Review

Pulkit Madaan (pmadaan2), Parthvi Mehta (pmehta14), Rishabh Nanawati (rnanawa1), Allan Raposo(Araposo1)

Paper: Single Image Reflection Removal Using Deep Encoder-Decoder Network (1)

Authors: Zhixiang Chi, Xiaolin Wu, Xiao Shu and Jinjin Gu

DOI: https://doi.org/10.48550/arXiv.1802.00094

Goal of the Paper

The goal of the paper is to tackle and propose a solution to removing reflections from a single image using deep convolutional neural networks. Reflections are phenomena in images caused by taking pictures through reflective mediums, such as glass, which often are unwanted by users. While humans can easily distinguish reflections from intended objects in images, it still proves to be a challenge for some of the state-of-the-art approaches published. Some of the approaches rely on making assumptions such as reflections being smoother on an image or requiring multiple shots in different angles of a scene. These approaches may prove to work under those conditions but fail when required to remove reflection from a single image. This paper proposes to solve this problem by leveraging machine learning with a synthetically created dataset of images which would allow them to create an end-to-end mapping between images with and without reflections to learn from.

Data

In the paper, the researchers steered from creating real image pairs for the end-to-end mapping due to the many difficulties it presented. Thus, their research depended on a creative approach of forming reflections to create photo-realistic reflection-tainted images from images without reflections from collected publicly available images. To create these images the researchers had to manipulate the images by applying blur to fix misalignment and double reflections, and inverse gamma correction to transform JPEG into raw images. The researchers collected 2303 images from the indoor scene recognition dataset and 2622 street snap images. These images were then used to create a collection of synthesized images split into a training set of 66,540 images and a testing set of 22,100 images. While the data their approach was tested in was synthetic, the researchers believe and continue to prove in their paper why their approach will work on real images.

Algorithm

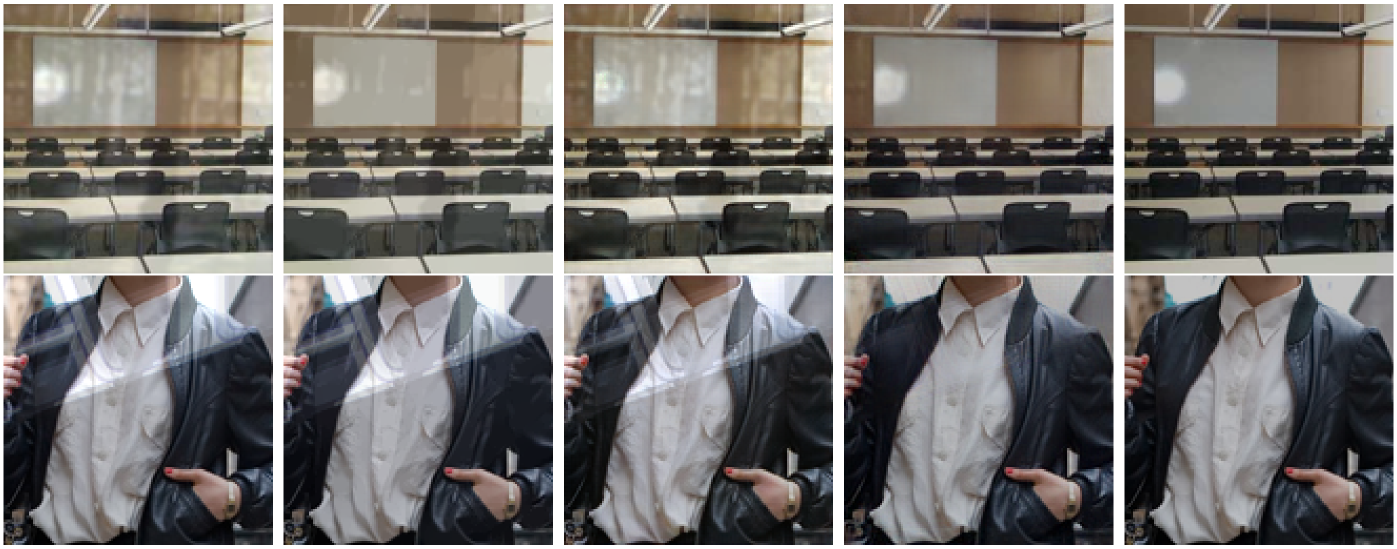
The work uses a deep learning paradigm where an encoder-decoder approach is used to reconstruct a reflection-free image from an image that has a reflection in it. The method involves training an end-to-end model that extracts features, recovers & removes reflections, and restores & refines the image to produce the final reflection-free output. To achieve this the work uses Convolutional Neural Networks, the first part of the network extracts features with convolutions, and the second part of the network uses residual networks with addition to recovering reflections followed by a residual network to subtract reflections from the recovered features. These features with reflections removed are refined by the third part of the network with transpose convolutions to finally yield a reflection-free image. The model is optimized to minimize the L2 distance between the ground-truth reflection-free image and reconstructed reflection-free image along with the L2 distance between the layer representations from VGG (2) network of the ground-truth and reconstructed images to avoid only focusing on low-frequency features.

Statistical Results

It is quite difficult to solve the challenge of reducing reflection interference from a single image. They propose a brand-new reflection creation model that takes the physics of digital camera imaging into account and uses it in a data-driven deep convolutional encoder-decoder network technique. Their experimental results reveal, despite the neural network learning only from synthetic data, that the proposed method is effective on real-world photos, outperforming the other tested state-of-the-art techniques significantly. They have used Peak Signal-to-Noise Ratio (PSNR) results to compare their results with the results of the two state-of-the-art reflection removal techniques (3) and (4) using synthetic data and a benchmark dataset (5). The assumption behind the method in (4) is that the reflection will be smooth. However, if this assumption is incorrect, as shown in the sample photos, (4) might even make the reflections worse than

	(3)	(4)	Results of Proposed Solution
Synthetic Images	19.72	19.82	29.08
Benchmark Set (5)	16.85	18.29	18.70

Table 1: PSNR results of tested techniques using synthetic images and a benchmark dataset.



Synthetic input (3) (4) Current Work (1) Ground truth

Figure 1: Comparison of reflection removal algorithms using synthetic images.

they were before. The approach suggested in this research does not rely on such an assumption; it functions effectively regardless of how smoothly the reflections are rendered. The output of (3) has a serious loss of details, creating an unnatural-looking image as a result. None of the evaluated algorithms can produce good results when the reflections are substantially stronger than the transmission layer. Their method achieves the highest average PSNR value in these tests. The proposed method takes around 0.6 s to process a 128 × 128 image and 2 s to process a 512 × 512 image.

Interpretation

This work uses a deep learning paradigm with encoder-decoder architecture to train the model to remove reflections from images. The data is synthetically generated to train the model. The methodology used by the authors and the reasoning provided for it justify the fact that the experiment produces better results when compared to two similar state-of-the–art approaches and the results translate well with real-world images too. Additionally, these compared approaches though have some shortcomings as described in the previous section. However, when compared to methods proposed by other publications that do not use CNNs, this algorithm may not be better. It would have helped to have a comparison drawn with LSTM-based approaches like (6) and Generative Adversarial Network approaches like (7) which solve the same task.

References

[1] Z. Chi, X. Wu, X. Shu, and J. Gu, "Single image reflection removal using deep encoder-decoder network," *CoRR*, vol. abs/1802.00094, 2018.

[2] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings* (Y. Bengio and Y. LeCun, eds.), 2015.

[3] N. Arvanitopoulos Darginis, R. Achanta, and S. Süsstrunk, "Single image reflection suppression," IEEE Conference on Computer Vision and Pattern Recognition, (New York), pp. 9. 1752–1760, Ieee, 2017.

[4] Q. Fan, J. Yang, G. Hua, B. Chen, and D. P. Wipf, "A generic deep architecture for single image reflection removal and image smoothing," *CoRR*, vol. abs/1708.03474, 2017.

[5] R. Wan, B. Shi, L.-Y. Duan, A.-H. Tan, and A. Kot, "Benchmarking single-image reflection removal algorithms," 10 2017.

[6] C. Li, Y. Yang, K. He, S. Lin, and J. E. Hopcroft, "Single image reflection removal through cascaded refinement," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.

[7] R. Abiko and M. Ikehara, “Single image reflection removal based on gan with gradient constraint,” in *Pattern Recognition* (S. Palaiahnakote, G. Sanniti di Baja, L. Wang, and W. Q. Yan, eds.), (Cham), pp. 609–624, Springer International Publishing, 2020.