SageRef: Single Image Reflection Removal

Pulkit Madaan (pmadaan2), Parthvi Mehta (pmehta14), Rishabh Nanawati (rnanawa1), Allan Raposo(araposo1)

1 Introduction

- Reflections in photographs can hinder image restoration tasks like segmentation and recognition, and humans can separate reflection and transmission images, but statistical models struggle with this task. Many of the existing techniques rely on multiple input images or specific photography techniques.
- The goal of this project is to implement a learning-based method for removing reflections from photographs taken through transparent media, such as glasses, using a single input image.
- The proposed method involves using a Denoising AutoEncoder with an encoder-decoder approach and Convolutional Neural Networks. Skip connections will be used, and the model will be optimized to minimize the L2 distance between the ground truth and reconstructed images and layer representations from the VGG network to avoid only focusing on low-frequency features.
- Codebase: https://github.com/madaanpulkit/SageRef

2 Methods

- Data: We utilized CEILNET dataset and leveraged its image pairs of mixture, reflection and background of images to train our model.
- Method & Algorithm: A Supervised Denoising AutoEncoder is used where the input is the image with reflection and the output is learned to be a reflection-free image called the label. The loss is L2 where the quantities are the Prediction and the Label (reflection-free image).
- Software: We use PyTorch Lightning software to reduce the training code overhead. We will use Google Colab for the demonstration of results.

3 Results

3.1 Metrics

We are using three metrics currently to compare our outputs with the ground truth images:

- MSE (Mean Squared Error): The average of the squared differences between the original image and the processed image.
- PSNR (Peak Signal-to-Noise Ratio): A measure of the peak error between the original and processed images. It is expressed in decibels (dB) and indicates how well the processed image preserves the details and colors of the original image.
- SSIM (Structured Similarity Index): The structural similarity between the original and processed images by comparing the luminance, contrast, and structure of the two images.

3.2 Current Progress and Results

We have currently 2 preliminary experiments. Figures 1, 2 and 3 represent the progress of training our model through the chosen metrics. Figure 4 represents one of the input images and how it gets segregated into its reflection layer and background layer.

4 Discussion

- (1) proposed the original CEILNet architecture. Researchers have experimented with a few more modifications of the architecture since 2017. Once ready, we plan on comparing our results with theirs.
- The following changes could be implemented to improve the model further:
  - Deeper/Larger Architecture: Due to hardware limitations, we are restricted in choosing our model's size.
  - More diverse training data: The dataset could include more realistic images, captured from a diverse set of scenarios. This would help train the model on different types of reflections.

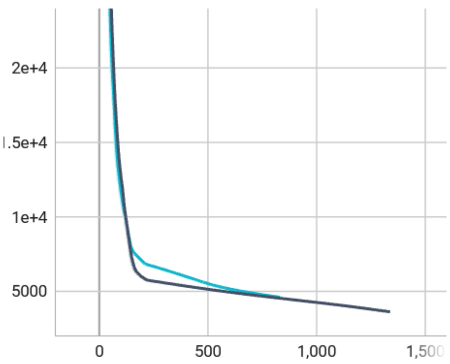

Figure 1: MSE Loss

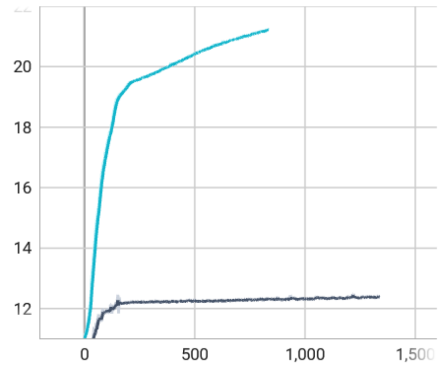

Figure 2: PSNR

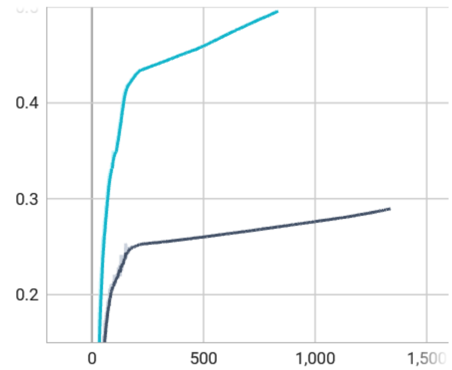

Figure 3: SSIM

Figure 4: The first image represents the input image, the second image shows the extracted reflection layer and the third image shows the output without the reflections.

5 Conclusion

To conclude, removing reflections from images is a difficult problem with many possible solutions. However, we have implemented a new model that considers the physics of digital camera imaging and used it in a deep convolutional encoder-decoder network based on a data-driven technique. Although the neural network only learns from synthetic data currently, we hope to find that our experiment is positively effective in removing reflections from real-world images, performing at-par other state-of-the-art techniques.

References

[1] Q. Fan, J. Yang, G. Hua, B. Chen, and D. Wipf, "A generic deep architecture for single image reflection removal and image smoothing," 2017.