Image Restoration using Deep Neural Networks

Parul Gupta, Apurbaa Bhattacharjee, Bastin Joseph
{pgupta52,bhattachar26,bjoseph5}@wisc.edu
University of Wisconsin-Madison

Abstract

We present a study of the MemNet model which performs 3 tasks: image denoising, JPEG deblocking and super resolution image generation. This model tackles the image restoration problem by using a recursive block which has both long term and short term memory blocks. These blocks help in achieving persistent memory in the model. We also perform stress testing on the model. We will also test the model on different dataset to understand how well the model will generalize with data from other types of datasets.

1. Introduction

Image restoration[6] is a classical computer vision problem. It is the process of recovering original noise-free image from a corrupted one. Often when we capture an image, the image acquisition process can be affected by several external factors like motion-blur due to camera shake, motion due to scene objects, artifacts in low light scenes, noise, blur etc. This leads to corruption in the acquired image. With huge advancements in digital photography and rise in the number of digital images, there is an increasing need to have efficient techniques to obtain noise free and high resolution images. With extensive literature survey, it has been established that the process of image restoration can be divided into separate problem statements like image denoising, image super-resolution and JPEG deblocking.

Image denoising is the technique of removal of white Gaussian noise of standard deviation σ from a noise-corrupted image to obtain a latent clean image. Image super-resolution is the process of conversion of a low-resolution image to a high-resolution image. JPEG compression is a lossy image compression technique which results in loss of image information. When the compression factor is high, the information loss in the images will appear as blocking artifacts in the compressed image. JPEG deblocking is an algorithm used to remove these artifacts.

MemNet [1] model by Tai et al. presents a model that looks at all the three image restoration problems all at once. In the project we intend to understand and analyze this model in depth and look into future applications of this image restoration model.

2. Problem Statement

The problem statement for this project is to restore an image using using three image restoration tasks - image denoising, super-resolution and JPEG deblocking. Given a low quality corrupted image the proposed solution will perform a feature extraction on it, then pass on the extracted features to a densely connected structure to convert the image to a high resolution image. Then a reconstruction network based on a residual network is used to get the uncorrupted image.

3. Motivation

Images are captured from areas ranging from professional photography to astronomy, surveillance, remote sensing, biomedical imaging etc. Interference in camera, varying lighting conditions like low light, extreme low light, different exposure ranges and external factors can often cause blurring and corruption of captured images. Image reconstruction has been a widely explored and researched topic in the field of computer vision. This problem statement is critical because this forms the premise for any aforementioned applications which need to make use of these images. It is highly essential that we remove the image corruption and improve the resolution of these images so that they can be efficiently used in ongoing research, academia and industrial applications.

4. Current State-of-Art

Image Denoising : In [2], Cheng et al. proposed to use patch guided internal clustering algorithm for image denoising. It utilizes Gaussian mixture model learning to guide the clustering of noisy images followed by an approximation process to estimate the subspace for image recovery. Zhang et al. [3] introduced the concepts of short-term and restricted long-term memory by making use of skip connections to pass information through the layers of the network.

Single Image Super Resolution : Mao et al. [4] introduced symmetric skip connections into a 30-layer convolutional auto-encoder network for image denoising and single image super resolution.

JPEG Deblocking : In [5], Dong et al. introduced an extended convolution network called Artifacts Reduction Convolutional Neural Networks (ARCNN) for removing the JPEG compression artifacts effectively.

5. Steps Executed and Challenges Faced

5.1. Generate datasets

The pre-trained model[10] used PASCAL VOC2007 dataset for training and was only available for JPEG Deblocking task. We used the **PASCAL VOC2007**[13] for creating the synthetic datasets for each of the image restoration use-cases as follows -

1. Super Resolution
   Ground truth image is downscaled and then upscaled by 3x to generate train image. Nearest neighbour interpolation is used to scale the image. Here the goal is to increase the clarity of the output image. Train sample and ground truth image have same dimensions

2. Denoising
   Salt and pepper noise is applied on ground truth image

3. Combined - [Super Resolution + Denoising + JPEG Deblocking]
   All the three operations were applied on the ground truth image in the following order to produce images which all the 3 defects $\implies$ scaling, noise, JPEG compression.

In each of the above cases, ground truth image is the brightness channel of the images in PASCAL VOC2007 dataset

5.2. Environment Setup

We started with the training of baseline code[10] on CPU using small training data (1000 images which is $\frac{1}{5}$th the size of actual training set) for 1000 iterations only. The process was extremely slow, it was taking $\approx 283s/iteration$ which would have taken close to 8 days for training the entire dataset (5K images) for 100K iterations. The image results on the smaller dataset with fewer iteration were below average.

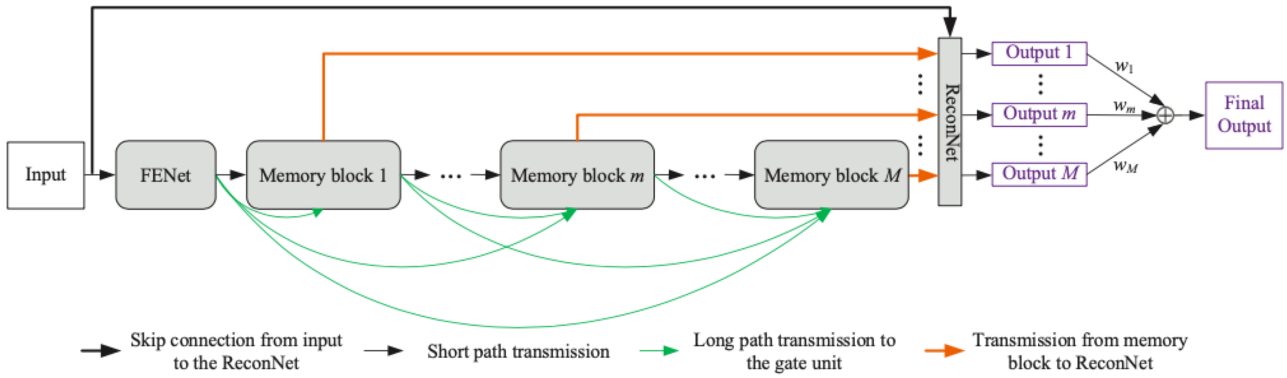

Figure 1. MemNet [2] architecture diagram

5.3. Porting to GPU using Google Cloud Platform (GCP)

Setting up GPU on GCP using a Deep Learning VM[14] instance with the following configurations -

- VM instance: 2 vCPUs + 13 GB memory (n1-highmem-2)
- Standard Persistent Disk: 100GB
- NVIDIA Tesla K80 GPU
- TensorFlow Enterprise 2.1 (CUDA 10.1)

The baseline code[10] was written on Tensorflow 1.x API. This was a major challenge we faced while deploying the model on the GPU. We had to port the entire codebase to Tensorflow 2.x APIs because the VM instance supported TensorFlow 2.1.

5.4. Model Training

We trained the model for 4 different image restoration tasks

1. JPEG deblocking
2. Denoising
3. Image Super-Resolution
4. Combined [JPEG deblocking + Denoising + Image Super-Resolution]

Each of the above models were trained for 100K epochs using 5K images from PASCAL VOC2007 dataset. On the GPU the training took $\approx 1.23s/iteration$, and based on our computational resources we could train each model in ≈ 2 days/model.

5.5. Model Testing

Each model has been tested on the same set of test images and compared visually for performance.

6. Revised Project Timeline

6.1. Milestone 1 : Mar 13, 2020

[Completed] Implementation/bring-up of the MemNet model

6.2. Milestone 2 : Mar 21, 2020

[Completed] Trained and tested several models based on the following use-cases :

1. JPEG Deblocking
2. Salt-n-Pepper Noise Removal
3. Super-Resolution
4. Combined model with 1,2,3

6.3. Milestone 3 : April 15, 2020

Model generalization by testing out new use-cases and how well they perform with this architecture

6.3.1 Robustness/Stress testing

(i) Test same image with different compression factor and see how well they perform on the JPEG Deblocking model.
(ii) Try out different types of noise for the Denoise model.
(iii) Generate new dataset with random scaling factor and check the Super-Resolution model performance.
(iv) Test images from the individual use-cases on the combined model.

6.3.2 Model generalization

(i) Train a model for **DeRain** task and compare with existing state-of-arts[14][9].
This task aims to remove rain droplets from images with rainfall
(ii) Train a model for **DeSnow** task and compare with existing state-of-arts[8].
The objective is to remove snow from images with snowfall.

6.3.3 Test RGB images

Current model support only grayscale images. We intend to extend it to color image restoration as well.

6.4. Milestone 4 : May 4, 2020

6.4.1 Performance testing

Test the performance of any state-of-art object detection model (Detectron2)[13] with and without image restoration (our model) as a preprocessing block in the image processing pipeline.

6.4.2 Ablation studies

Ablation studies on the current architecture such as changing the network depth, changing the position of the skip layers, removing CNN layers.

7. Intermediate Results

We have trained four separate models for the following image restoration tasks - Super Resolution, JPEG Deblocking, Salt-n-Pepper Denoising, combined model with all the three attacks. We have taken three images from our test set and demonstrated the result obtained from the respective trained models. The ground truth, attacked test image, recovered output image for each of the four models have been shown in Fig. [2-37]. Fig. [38-41] shows the variation of model loss wrt number of iterations.

8. Final Results & Performance Evaluation

The performance of our model will be tested using the following evaluation metrics -

Robustness - We will be using test images with random compression levels (JPEG deblocking model), random scaling factor(Super-Resolution model), different types of noise and/or different levels of noise(Denoise model).

Generalization - We will try to use this architecture for some other types of image noise and test how well it performs. For example - Train a model for **DeRain** and **DeSnow** tasks, compare the results with existing state-of-art models. Our goal is to make a qualitative analysis of how well the model generalizes to other datasets as well as other image restoration tasks.

Benchmark Model Output - We will be calculating the Average PSNR/SSIMs for each of the test images for varying noise levels, compression levels, scale factors for all the trained models.

Object Detection & Recognition - Test the performance of any state-of-art object detection model (Detectron2)[13] with and without image restoration (our model) as a preprocessing block in the image processing pipeline.

9. Applications

This model can be used as a preprocessor block for existing computer vision solutions such as autonomous vehicles, object detection and recognition.

References

[1] Tai, Ying, et al. ”Memnet: A persistent memory network for image restoration.” Proceedings of the IEEE international conference on computer vision. 2017.

[2] F. Chen, L. Zhang, and H. Yu. External patch prior guided internal clustering for image denoising. In ICCV, 2015.

[3] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang. Be- yond a gaussian denoiser: Residual learning of deep CNN for image denoising. IEEE Trans. on IP, 2017.

[4] X. Mao, C. Shen, and Y. Yang. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. In NIPS, 2016.

[5] C. Dong, Y. Deng, C. C. Loy, and X. Tang. Compression ar-tifacts reduction by a deep convolutional network. In ICCV,2015.

[6] P. Milanfar. A tour of modern image filtering: new insights and methods, both practical and theoretical. IEEE Signal Processing Magazine, 30(1):106–128, 2013.

[7] Redmon, Joseph and Divvala, Santosh and Girshick, Ross and Farhadi, Ali.You Only Look Once: Unified, Real-Time Object Detection. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[8] Y. Liu, D. Jaw, S. Huang and J. Hwang, ”DesnowNet: Context-Aware Deep Network for Snow Removal,” in IEEE Transactions on Image Processing, vol. 27, no. 6, pp. 3064-3073, June 2018.

[9] Xueyang Fu, Jiabin Huang, Xinghao Ding, Yinghao Liao, John Paisley, ”Clearing the Skies: A Deep Network Architecture for Single-Image Rain Removal” in IEEE Transactions on Image Processing, vol. 26, pp. 2944–2956, June 2017.

[10] https://github.com/lyatdawn/MemNet-Tensorflow

[11] https://github.com/facebookresearch/detectron2

[12] https://cloud.google.com/deep-learning-vm

[13] http://host.robots.ox.ac.uk/pascal/VOC/voc2007/index.html

[14] Fu, X., Huang, J., Ding, X., Liao, Y., Paisley, J. (2017). Clearing the skies: A deep network architecture for single-image rain removal. IEEE Transactions on Image Processing, 26(6), 2944-2956.

Figure 2. Ground Truth

Figure 3. Low resolution image

Figure 4. Resize model output

Figure 5. Ground Truth

Figure 6. Low resolution image

Figure 7. Resize model output

Figure 8. Ground Truth

Figure 9. Low resolution image

Figure 10. Resize model output

Figure 11. Ground Truth

Figure 12. Noisy Image

Figure 13. Denoise model output

Figure 14. Ground Truth

Figure 15. Noisy Image

Figure 16. Denoise model output

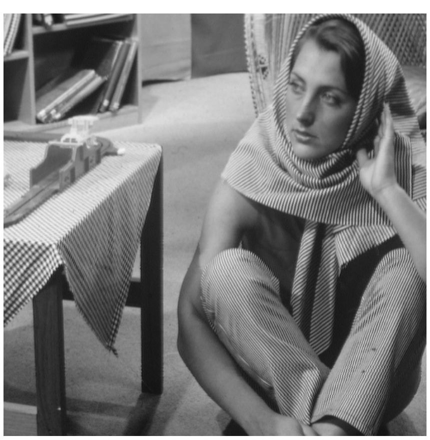

Figure 17. Ground Truth

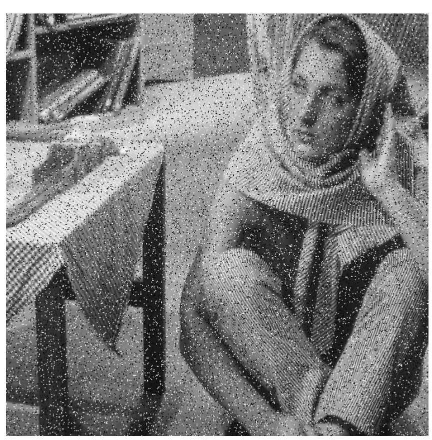

Figure 18. Noisy Image

Figure 19. Denoise model output

Figure 20. Ground Truth

Figure 21. JPEG compressed image

Figure 22. JPEG Deblock model output

Figure 23. Ground Truth

Figure 24. JPEG compressed image

Figure 25. JPEG Deblock model output

Figure 26. Ground Truth

Figure 27. JPEG compressed image

Figure 28. JPEG Deblock model output

Figure 29. Ground Truth

Figure 30. Combined attack image

Figure 31. Combined model output

Figure 32. Ground Truth

Figure 33. Combined attack image

Figure 34. Combined model output

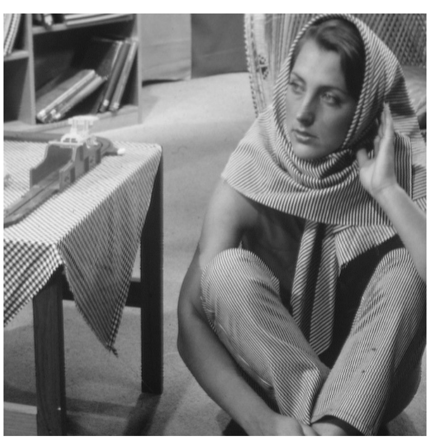

Figure 35. Ground Truth

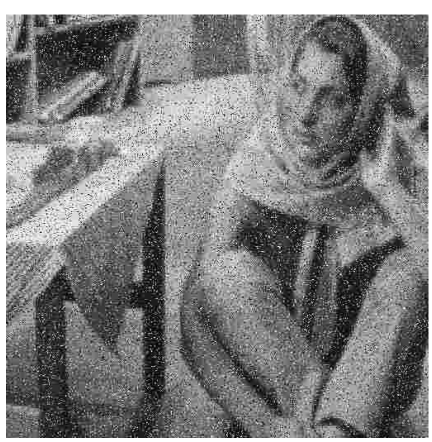

Figure 36. Combined attack image

Figure 37. Combined model output

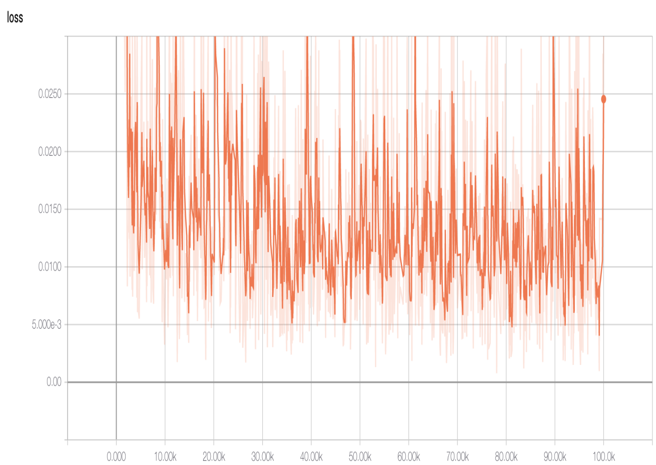

Figure 38. JPEG Deblocking iterations vs. training loss

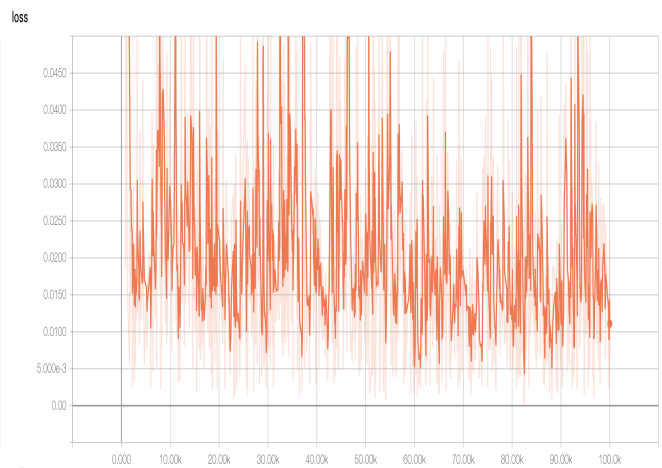

Figure 39. SuperResolution iterations vs. training loss

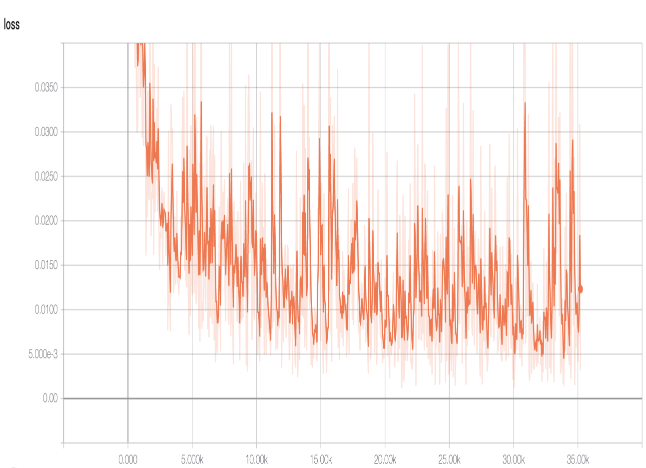

Figure 40. DeNoise iterations vs. training loss

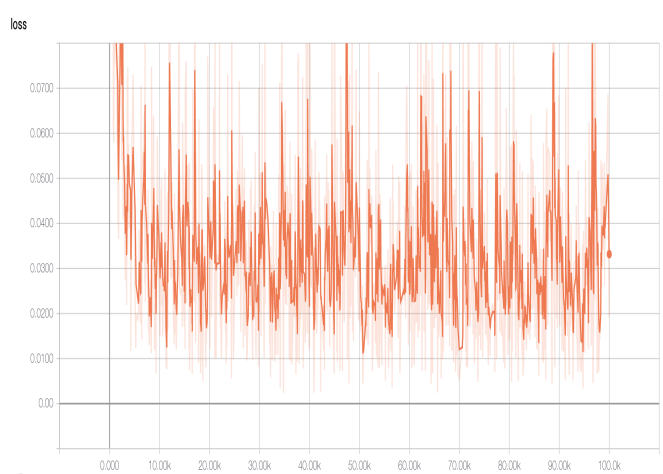

Figure 41. Combined MemNet iterations vs. training loss