Unsupervised Image Denoising using Deep Learning

Harshit Srivastava
hs3500@nyu.edu

Akshat Tyagi
at3761@nyu.edu

Anurag Marwah
am8482@nyu.edu

New York University - Fall 2018

1   Introduction

Recent advances in deep neural networks have sparked significant interest in denoising techniques without using any well-curated mathematical model to remove noise. Noisy images can be cleaned using clean images as target images and training a neural network on the (noisy, clean) pair but as we show in this project, similar results can be achieved by using (noisy, noisy) pairs as long as the underlying image remains the same. The loss functions for both methods can be defined as:

$$Supervised : \sum_i L(f(noisy_i|\theta), clean_i)$$

$$Unsupervised : \sum_i L(f(noisy_{1,i}|\theta), noisy_{2,i})$$

The image denoiser could have applications in a plethora of domains, including military and healthcare, where clean target images are not readily available or take significant amount of human intervention to denoise them.

2   Related work

Previous techniques have existed where deep neural networks were used to denoise images [4] but these methods have relied on clean images for training. Denoising the image using corrupted target images was performed recently where Lethtinen et al., 2018 [2] showed that clean images are not necessary for denoising images, which can be applied to a wide range of areas. As far as the architecture is concerned, we used a U-net which is essentially a combination of Convolutional neural networks, pooling, ReLu activation in a U-shaped architecture [5] with deconvolutional units which generate the output in the shape of an image and has been widely used for cases where outputs are required to be images.

3   Problem formulation

The objective is to apply synthetic noise on the image dataset, train the network, recover the images using the trained graph and observe the recovery loss. Different noise settings were used for source and target images but since the underlying image was the same, clean images were restored. We depreciated the size of the image to 128 x 128 to account for computing issues which may have been encountered on using too many features. At the same time, low-resolution images could not be used for this task as denoising might not be effective.

4   Methods

We have used a U-NET to perform Image denoising to remove Gaussian noise and Textual noise from an image. To do this, we used the IMAGENET dataset [1] and trained on 290 images from this dataset. For random impulse noise removal, we trained the network for 5 epochs and for textual and gaussian noise, we trained the network for 10 epochs. We measured our performance using PSNR (Peak Signal Noise Ratio). For each task we performed, we used a different loss function. We used L_2 loss for Gaussian noise removal, L_1 error for Textual Noise removal, and L_0 Loss for Random Impulse noise removal.

$$L_0 : \frac{1}{N} \sum_i \lim_{p \to 0} \sqrt[p]{(y_i)^p}$$

$$L_1 : \frac{1}{N} \sum_i |y - y_i|$$

$$L_2 : \frac{1}{N} \sum_i \sqrt{(y - y_i)^2}$$

4.1 Noise models

Gaussian noise model was obtained by sampling from a gaussian distribution with the following parameters:

source : $\mu - 0, \sigma - 50$
target : $\mu - 0, \sigma - 50$
validation : $\mu - 25, \sigma - 25$

Textual noise was added using Uniform distribution to decide on the amount of space to be occupied by random texts. The target images in this case were peppered with similar noise model but since the added text and the occupancy is randomly allocated, train and test images were distinct. The maximum number of words in the image was set at 50. For validation, the number of words were set at 25.

Random Valued Impulse noise was added using a mask to add on to the original image. The mask contained values generated from binomial distribution whose probability came from a uniform distribution. The maximum occupancy was set at 95. For validation, the occupancy was set at 70.

5 Architecture and design

We implemented U-Net to train the network in this part, we plan on using other architectures, specially SRRES-NET [3] in coming days. Figure (1) was generated using Tensorboard after training the network. The actual graph on tensorboard contained a lot of nodes as the network was very deep but this image shows a brief overview of our network. We experimented with different number of hidden layers and the same graph can be extended for deeper networks as well. Up-sampling was done by performing deconvolution which was implemented using Conv2DTranspose in keras.
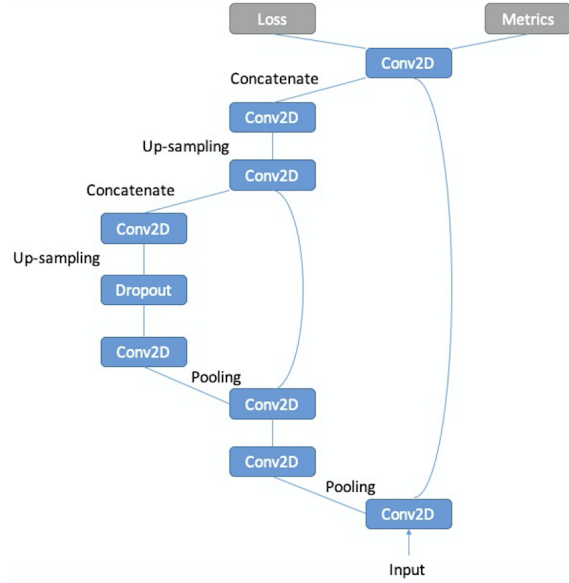

Figure (1)   U-shaped network used for training

6 Preliminary results

GitHub repository : github.com/harshit0511/Unsupervised-Image-denoising

We got PSNR of around 29.1 for all noise models when using noisy images as target. Interestingly, on using clean images as targets, PSNR of around 30.5 was obtained although visually there is no difference between the images generated by both networks. Figure (3) shows images which were denoised using our network. We would like to further experiment with our network and use some other well-known architectures in order to achieve higher PSNR with noisy target images.

7 Further steps

We would like to use another model (SRRESNET) to perform noise removal for gaussian, random and textual noise. We can then compare the performance of the two neural networks. We would also like to train the network for more epochs and for that, we will utilize NYU HPC Cluster environment. Ideally, we would like to train the network for at least 40 epochs. Finally, we will try to formulate our own network for performing noise removal

Gaussian Noise ($\mu = 0, \sigma = 50$)

Textual Noise

Random Impulse Noise

Figure (3)   Images on the left are the ground truth image, images in the middle are the generated noisy images which are used as source and target images, images on the right were the denoised images generated by the trained network.
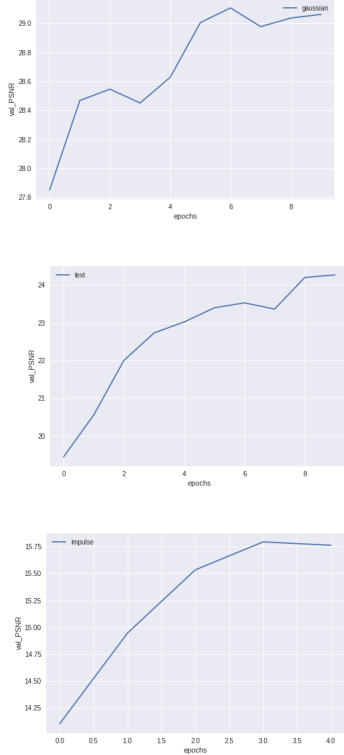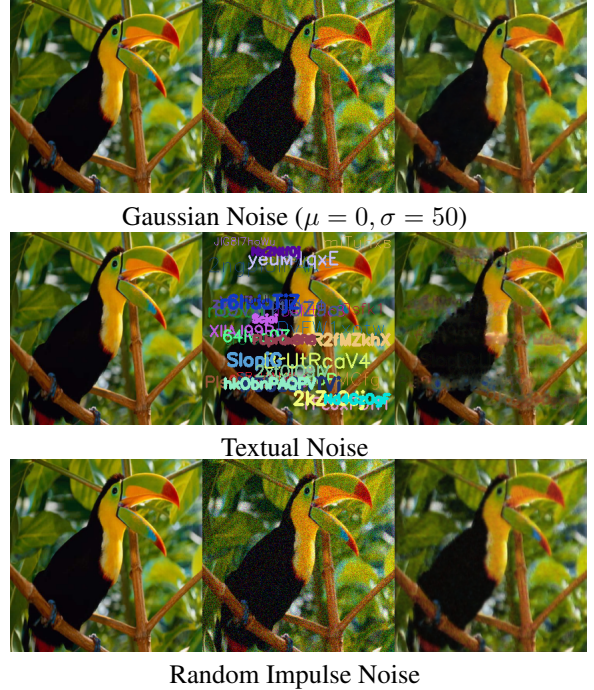
Figure (2)   Plots depicting the change in PSNR Values on the validation set while training the network on the 3 noise models (Gaussian, Textual, and Random impulse)

and compare the results with both SSRESNET and U-NET.

References

[1] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR09*, 2009.

[2] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila. Noise2noise: Learning image restoration without clean data. *CoRR*, abs/1803.04189, 2018.

[3] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution. *CoRR*, abs/1707.02921, 2017.

[4] Xiao-Jiao Mao, Chunhua Shen, and Yu-Bin Yang. Image restoration using convolutional auto-encoders with symmetric skip connections. *CoRR*, abs/1606.08921, 2016.

[5] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *CoRR*, abs/1505.04597, 2015.