

# Natural Image Denoising with Deep Neural Networks

Xuan Chen, Geng Ji, Shenglong Li, Zhonglei Wang

## 1 Objective

In this project, we aim to remove the additive white Gaussian noise from gray-scale natural images using end-to-end deep neural networks.

## 2 Related work

Natural image denoising is a well studied image processing task. Many non-deep-learning methods have reached success in this area. The K-SVD method (Elad and Aharon, 2006) generalizes K-means clustering to learn a dictionary for sparse coding of image patches. The state-of-the-art *learned simultaneous sparse coding* (LSSC) (Mairal et al., 2009) and *block matching and 3D filtering* (BM3D) (Dabov et al., 2008) methods integrate clustering, dictionary learning, and denoising to extract information directly from a single corrupted image. Alternatively, the *expected patch log-likelihood* (EPLL) (Zoran and Weiss, 2011) method maximizes the log-likelihood of overlapping image patches under a finite Gaussian mixture model learned from uncorrupted natural images.

A couple of deep neural networks have been developed for image denoising as well. Jain and Seung (2009) designed a four-layer convolutional neural network. Xie et al. (2012) combined the idea of stacked denoising auto-encoders with sparse coding, and came up with a fully connected network called SSDA on image patches. However, the performance of those methods are inferior the state-of-the-arts. The reason might be that comparing to other tasks such as image classification where deep learning has already done pretty well, image denoising requires per-pixel intensity estimation so that the dimension of this problem is much higher than just predicting a single label.

## 3 Our contributions

### 3.1 Short-term goals

First, we hope to replicate the work done by Jain and Seung (2009) and Xie et al. (2012) using tensorflow. The diagram for the CNN developed by Jain and Seung (2009) is shown in Fig. 1. It takes in noisy images, and tries to minimize the mean square error against their Tuncorrupted counterparts. The network contains four hidden layers, each with 24 feature maps that all share the same size with the input image. Note that it is a pure CNN without any pooling layer which would otherwise shrink the size of the feature maps or any fully connected layer which would result in too many parameters to learn.

The sparse stacked denoising auto-encoders (SSDA) uses denoising auto-encoders (DA) as building blocks, a popular pre-training framework, and stacks up four layers of them, as shown in Fig. 2. For each DA, it tries to learn a pair of weights  $W$  that encodes the noisy input  $x = \eta(y)$  and  $W'$

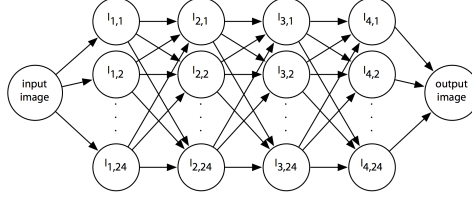


Figure 1: Convolutional neural network developed by Jain and Seung (2009).

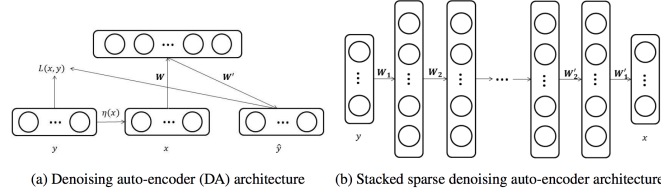


Figure 2: Sparse stacked denoising auto-encoder developed by Xie et al. (2012).

that decodes it into the clean counterpart  $y$ . To enhance sparsity, the loss function  $L(x, y)$  contains not only the mean square error between  $x$  and  $y$ , but also some regularization terms for the weight parameters in the network.

### 3.2 Long-term goals

Our long-term goal is to improve the two networks for better performance. For example, we may add more layers to each of them, and make use of dropout to prevent overfitting. If time allows, we also plan to try other types of recently developed networks such as ResNet and Highway networks.

## 4 Dataset Link

We use the Berkeley Segmentation Dataset for both training (400 images) and test (100 image). Its link is <https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/>. We've already started training the networks shown in Fig. 1 and Fig. 2 on this dataset using our machines, but it takes too much time. We've talked to the TA's about using the GPU version of tensorflow on department machines and got positive feedback. Hopefully packages like scipy could be correctly installed on those machines in a short time so that we may run our code there.

## References

- Michael Elad and Michal Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing*, 15(12):3736–3745, 2006.
- Julien Mairal, Francis Bach, Jean Ponce, Guillermo Sapiro, and Andrew Zisserman. Non-local sparse models for image restoration. In *International Conference on Computer Vision*, 2009.
- Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image restoration by sparse 3d transform-domain collaborative filtering. In *Electronic Imaging*, 2008.

- Daniel Zoran and Yair Weiss. From learning models of natural image patches to whole image restoration. In *International Conference on Computer Vision*, 2011.
- Viren Jain and Sebastian Seung. Natural image denoising with convolutional networks. In *Advances in Neural Information Processing Systems*, pages 769–776, 2009.
- Junyuan Xie, Linli Xu, and Enhong Chen. Image denoising and inpainting with deep neural networks. In *Advances in Neural Information Processing Systems*, pages 341–349, 2012.