Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 171 (2020) 1535–1541

Procedia Computer Science

www.elsevier.com/locate/procedia

Third International Conference on Computing and Network Communications (CoCoNet'19)

# Autoencoders Based Deep Learner for Image Denoising

Komal Bajaj[a], Dushyant Kumar Singh[b,*], Mohd. Aquib Ansari[c]

[a]*Department of Computer Science & Engineering, MNNIT Allahabad, Prayagraj, U.P., INDIA*
[b]*Department of Computer Science & Engineering, MNNIT Allahabad, Prayagraj, U.P., INDIA*
[c]*Department of Computer Science & Engineering, MNNIT Allahabad, Prayagraj, U.P., INDIA*

**Abstract**

Nowadays, digital images have a valuable role in our daily life, and can be used for various of applications like fingerprint recognition, video surveillance etc. Sometimes, images get infected with noise due to many reasons such as defects in camera sensors, transmission in noisy channel, faulty memory locations in the hardware etc. Processing a noisy image is not advisable because usually it yields erroneous outcomes. So, as to improve it for subsequence processing, the noise must be eliminated from the image in advance. Therefore, there is a need of an efficient image denoising technique that helps to deal with noisy image. Image denoising is a process to realign the original image from the degraded image. In this paper, autoencoders based deep learning model is proposed for image denoising. The autoencoders learns noise from the training images and then try to eliminate the noise for novel image. The experimental outcomes prove that this proposed model for PSNR has achieved higher result compared to the conventional models.



*Keywords:* Noise, Denoising, Deep learning, Convolutional Neural Network, Autoencoders;

## 1. Introduction

Image processing has numerous applications including image representation, image segmentation, object detection & classification, action recognition etc. [1] [14]. An image contains valuable information and this information can be effected by the noise. The noise is an undesirable signal, which generates a random variation of color or brightness values within an image. Noise can create inexpedient outcomes like unrealistic edges, ignored lines, corners, artifacts and blurry objects, which ultimately hinders the manipulation of the image. The noise may be accumulated within an image during the image acquisition or transmission process. There will always be some noise within an image that was captured by erroneous camera. Therefore, removing noise is an essential requirement to enhance and retrieve the valuable hidden details within an image. So, image denoising technique [2] [10] can be used to fulfil this purpose.

* Corresponding author. Tel.: +91-790-551-4182 ; fax: +0-000-000-0000.
*E-mail address:* dushyant@mnnit.ac.in


10.1016/j.procs.2020.04.164

Image denoising techniques are used to eliminate the noise and generates clean image from the noisy image [3]. It also preserves the essential details of an image while suppressing noise. For most of the applications, the image denoising is a pre-processing step. The objective of image pre-processing is to improve the interpretability of an image for further processing by suppressing reluctant distortions or enhancing the image features. If somehow an image gets disturbed during picture acquisition and transmission process, it should be re-established before going for further processing. The image denoising can be used in association of different image activities like image restoration, visual tracking, image registration, image segmentation, image classification etc.

Nowadays, neural network and deep neural network [10] [17] are getting popular as an intelligent solution to almost every type of computing problem. Object classification & result prediction are common of those. The effort here is to expedite these deep learning techniques for modelling & identifying noise in the infected image.

This paper uses the concept of autoencoder for removing the Gaussian noise from the degraded image. An autoencoder is a type of artificial neural network that aims to learn a representation (encoding) for a set of data. It is the autoencoder that learns noise from training images and then tries to generate a clean image very close to its original input. This autoencoder based proposed model uses convolution layer, convolutional layer and deconvolution layer to defuse the noise. At last, peak signal to noise ratio (PSNR) and Structural Similarity Index (SSIM) are used to measure the performance of this proposed model.

This paper is organized into five sections. Section one is an introduction of image denoising. Literature survey has been done in section two. Section three presented a proposed network architecture for image denoising. The experimental results and analysis have been depicted in section four. Section five includes the conclusion with some future perspectives of the proposed work.

## 2. Related Work

Efros and Leung [4] used non local self-similarities to synthesize textures and to fill the holes in the images. Their algorithm scans a vast portion of the image in search of all the pixels that resemble the pixel in restoration. The resemblance is evaluated by comparing a whole window around each pixel. Applying this idea to neighborhood filters creates a generalized neighborhood filter that is called non-local means [2] [5].

Another method called Block Matching and 3D Matching (BM3D), given by Dabov et al. [6], is a complex and an advance method for image denoising. This technique gathers comparative 2D image sections and utilizes inverse 3D transformations to accomplish fine details of denoising. The BM3D algorithm has been broadened (IDD-BM3D) to perform decoupled deblurring and denoising by using Nash Equilibrium Balance. It is moreover fascinating to see these papers [7] [9] that attempts to assess the characteristic limits of fix based denoising techniques. It guaranteed that BM3D is truly near those optimality limits.

Another variety of deep neural network model that gave outstanding results, known as denoising autoencoder [8]. It is used for image denoising that inputs noisy images and tries to produce denoisy version of it. A variety of denoising autoencoders called as stacked denoising auto-encoder are some amongst the deep neural system models that can be utilized for image denoising. In this network, the reconstruction error is minimized at each layer with respect to inputs.

Xie, Linli and Chen [3] presented an approach for low level vision problems that used sparse coding with deep autoencoders. They suggested the different training methods that adapts denoising autoencoders for image denoising and image inpainting. For better performance, they used KSVD (K-means Single Valued Decomposition), which is widely used sparse coding method. For image inpainting, there is no need of any prior information about the region requiring inpainting. Their proposed work encompasses the stacked denoising autoencoders with sparse representation.

Generally, the strategies based on deep neural network derive the necessary parameters from the training data. This becomes progressively compelling in real world image reconstruction applications.

## 3. Network Architecture

In this section, the proposed network architecture is briefly described, which takes degraded image as input and produces clean enhanced image as output. This model mainly consists of input layer, convolution layer, convolutional layers, deconvolution layer and output layer. The several small and linearly connected convolutional denoising autoencoder (CDA) [8] [11] blocks are used in convolutional layer. CDA learns the degradation from the training images and tries to remove this degradation from input image. The structure of convolutional denoising autoencoder (CDA) is illustrated in figure 1.

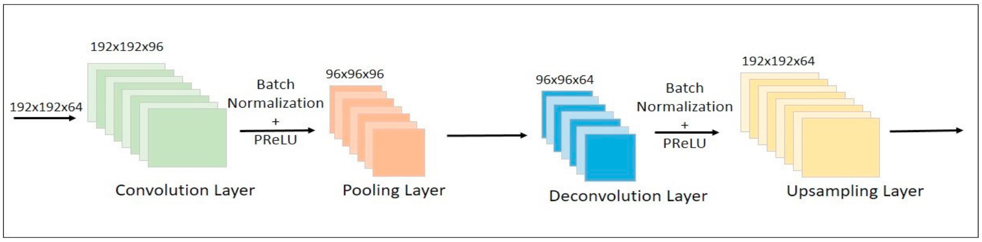


Fig. 1. Structure of Convolutional Denoising Autoencoder block

The CDA block includes four internal layers, named as convolution layer, pooling layer, deconvolution layer and upsampling layer. In this model, the batch normalization and Parametric Rectified Linear Unit (PRelu) activation function are deployed between convolution layer and pooling layer, and deconvolution layer and upsampling layer respectively. The output of CDA is connected with its input. The convolution layers are connected here to upsampling layers by direct connections. This helps in speeding up the process of learning.

The figure 2 shows the overall proposed model architecture. It comprises of input layer followed by convolution layer [1] [17]. The series of convolutional denoising autoencoder are used to process the image. At the end, two deconvolutional layers are accustomed to bring back the denoised image. Skip connections are utilized to link the input and output of same denoiser autoencoder. Here, ten convolutional denoising autoencoder blocks are used.

There are two purpose behind utilizing such skip connections within our proposed model. First, a more profound branch smoothes the color information more feasibly while some fine details of an image might be lost. Due to this, it can raise some difficulty for the deconvolution layer to reconstruct the image. So, skip connection helps to recuperate some finer image details. Second, the skip connections likewise accomplish benefits on backpropagation, which can make training much easier for deep neural network. The feature images passed by skip connections convey much

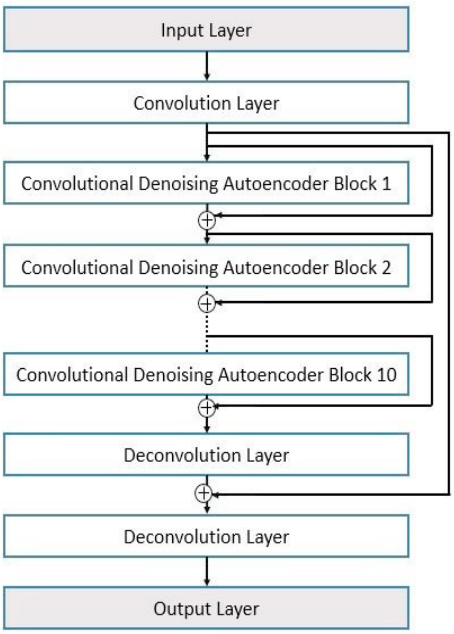


Fig. 2. Visualization of proposed Deep Neural Architecture

image details that encourages deconvolution to reconstruct an improved clean image. From deep branches, the overall model can cognize color smoothing efficiently with the help of skip connection.

## 4. Experimental Results and Analysis

The Python programming language is used here to perform the experiments using Numpy, Tensorflow and keras libraries. For accelarating the training of proposed model, the explicit Graphics Processing Unit (GPU) of NVIDIA GEFORCE 920 MX is used. It helps in dealing with deep structure and complex type of neural network. CUDA 10 toolkit is used in connection with GPU.

STL-10 dataset is used to train this proposed model. It is a image recognition dataset and it can be used for self taught learning algorithms, deep learning and unsupervised feature learning. This dataset contains 3 categories like training image set, testing image set and unlabeled image set. The training and testing image sets contain 5000 images and 8000 images respectively. The unlabelled image set contains 100000 unlabelled images, which are used for unsupervised learning. Each image is 96X96 pixels in size. This proposed denoising network used unsupervised learning based model. To train the model, 40000 unlabelled images are used for training and validation purpose.

The SET5 standard image dataset is used for testing purpose, that having set of 5 standard images. Here, we only focused to deal with gaussian noise. First, the gaussian noise is added to an image with different standard deviations $\sigma = 10, 30, 50, 70$. Then, the original image is recovered from its degradation using our proposed model.

As performance metrics, PSNR (peak signal to noise ratio) [15] [18] and SSIM (structural similarity Index) [16] are used to measure the performance of the proposed system, which are shown in equation 1, 2 and 3. The PSNR evaluates the peak signal to noise ratio value between the original image and the reconstructed image, where high PSNR value indicates good quality of the reconstructed image. SSIM is used to measure the perceptual difference between the original image and the reconstructed/cleaned image, where the SSIM value '1' indicates that the original image is identical to the reconstruction image and '< 1' indicates that the original image is far from reconstructed image. PSNR and SSIM values for different standard variations are illustrated in table 1.

$$PSNR = 10.log_{10}(\frac{MAXIMUM_I^2}{MeanSquareError}) \tag{1}$$

Where, $MAXIMUM_I$ is the largest probablistic value of image.

$$MeanSquareError = \frac{1}{mn}\sum_{p=0}^{m-1}\sum_{q=0}^{n-1}[I(p,q) - R(p,q)]^2 \tag{2}$$

Table 1. PSNR and SSIM results of Proposed Model on SET5 dataset.

| | | | PSNR | | | |
|---|---|---|---|---|---|---|
| S.D. | Baby | Bird | Butterfly | Head | Woman | Average |
| $\sigma = 10$ | 25.70 | 26.07 | 24.28 | 24.23 | 24.95 | 25.046 |
| $\sigma = 30$ | 26.05 | 26.36 | 24.77 | 24.64 | 25.48 | 25.464 |
| $\sigma = 50$ | 27.85 | 27.17 | 26.27 | 26.44 | 27.45 | 27.036 |
| $\sigma = 70$ | 29.12 | 27.28 | 26.26 | 27.68 | 28.01 | 27.671 |

| | | | SSIM | | | |
|---|---|---|---|---|---|---|
| S.D. | Baby | Bird | Butterfly | Head | Woman | Average |
| $\sigma = 10$ | 0.766 | 0.620 | 0.881 | 0.628 | 0.749 | 0.728 |
| $\sigma = 30$ | 0.764 | 0.628 | 0.880 | 0.633 | 0.752 | 0.731 |
| $\sigma = 50$ | 0.774 | 0.669 | 0.880 | 0.700 | 0.789 | 0.762 |
| $\sigma = 70$ | 0.789 | 0.783 | 0.869 | 0.734 | 0.858 | 0.806 |

Where, I(p, q) is the original image and R(p, q) is restored image of dimension mxn.

$$SSIM(p,q) = \frac{(2\mu_p\mu_q + c_1)(2\sigma_{pq} + c_2)}{(\mu_p^2 + \mu_q^2 + c_1)(\sigma_p^2 + \sigma_q^2 + c_2)} \tag{3}$$

Where, $\mu_p$ and $\mu_q$ are the mean of p and mean of q. $\sigma_p$ and $\sigma_q$ are variance of p and variance of q. $\sigma_{pq}$ is covariance of p and q. $c_1$ and $c_2$ are two variable for stablization of division with diminutive denominator.

Table 2 shows the comparison between the two baseline models and our proposed denoising model. These two baseline methods are Convolutional Denoising Deep Neural Network (CDDNN) [12] and Residual Encoder Decoder with 30 layer (RED30) [13].

Table 2. Comparison against Baseline Methods on SET5 dataset for different noise densities

| PSNR | | | |
|---|---|---|---|
| S.D. | RED30 | CDNN | Butterfly |
| $\sigma = 10$ | 25.74 | 24.80 | 25.046 |
| $\sigma = 30$ | 26.13 | 25.11 | 25.464 |
| $\sigma = 50$ | 26.40 | 24.63 | 27.036 |
| $\sigma = 70$ | 25.43 | 27.33 | 27.671 |

| SSIM | | | |
|---|---|---|---|
| S.D. | RED30 | CDNN | Butterfly |
| $\sigma = 10$ | 0.842 | 0.799 | 0.728 |
| $\sigma = 30$ | 0.819 | 0.773 | 0.731 |
| $\sigma = 50$ | 0.749 | 0.702 | 0.762 |
| $\sigma = 70$ | 0.670 | 0.623 | 0.806 |

As indicated in Table 2, the improvement of PSNR's values are significant enough for the proposed methods. Our proposed model performed better than CDNN model in terms of PSNR. For RED30 model, it is performing well only for the higher rate of standard deviation. In some cases, the value of SSIM is also decreased. Accordingly, it can very well be estimated that our proposed technique exhibits better than CDNN and RED30 as far as PSNR values are considered. The baseline methods are preferable sometimes over our proposed strategies while considering SSIM values. This implies that our proposed model keeps up higher PSNR values than the fundamental strategies, yet compromise with the structural similarity.

Figure 3 illustrates the generated loss function curve (MSE curve) and PSNR curve for our proposed model. This curve is generated for 50 iterations. Figure 4 shows the visual results that are evaluated from different image denoising methods. First, gaussian noise is added within the original image to make it noisy. Then, RED30, CDNN and our proposed model are applied to reconstruct the original image from noisy image.

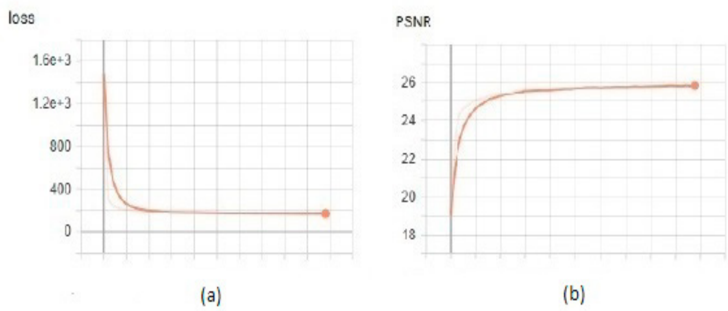


Fig. 3. (a) Loss Function Curve (b) PSNR Value Curve

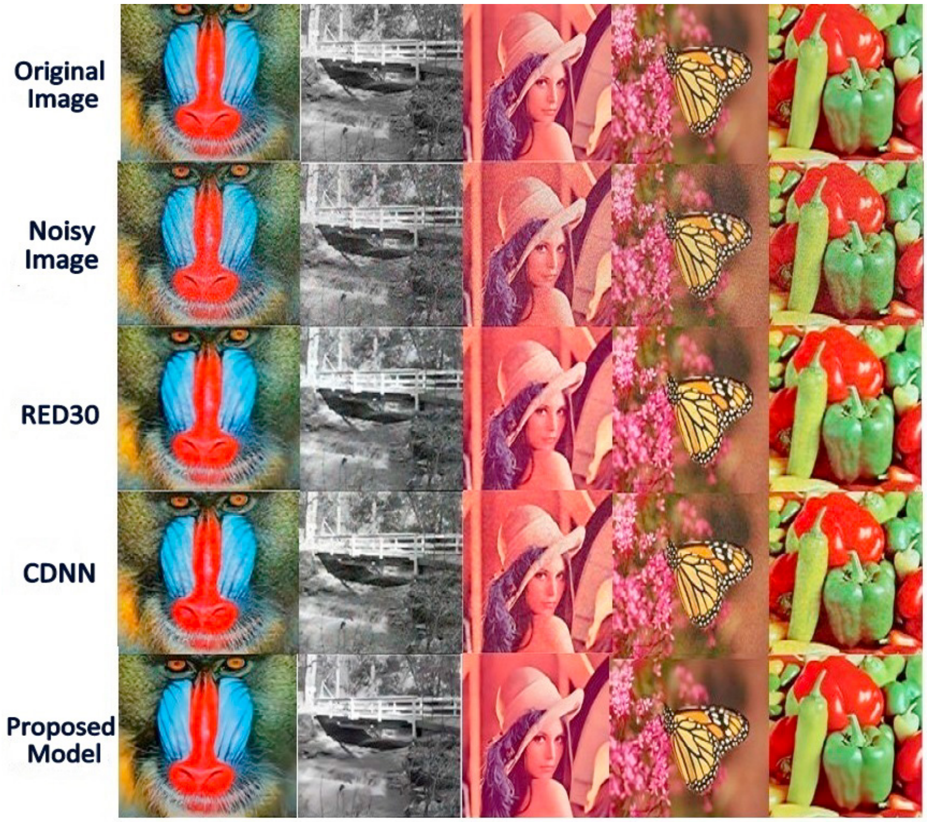


Fig. 4. Visual results of image denoising. Images from top to bottom are: Original Image, Noisy Image, the recovered image from RED30, CDNN and Proposed Model.

## 5. Conclusion

This paper has presented a better approach for image denoising based on deep convolutional denoising autoencoder framework. This proposed framework helps to remove the gaussian noise. Here, first traditional methods for image denoising have been delineated and then improvement of deep learning for image denoising has been presented. Forby, the skip connections are utilized here to solve the classic task of image denoising, which not only helps to reproduce clean images but also helps to handle the optimization issue i.e. gradient fading. The result of the experiment shows that our proposed model has performed better than conventional technologies in terms of PSNR. In future, we can deal with other type of noises like Salt and Pepper noise, Poisson noise etc. This proposed method can also be incorporated with different applications like image inpainting, image super resolution and denoising of sound and video.

## References

[1] Dwivedi N., Singh D.K. "Review of Deep Learning Techniques for Gender Classification in Images." *In: Advances in Intelligent Systems and Computing*, vol 741. Springer, Singapore.

[2] Buades, Antoni, Bartomeu Coll, and Jean-Michel Morel. "Nonlocal image and movie denoising." *International journal of computer vision* 76.2 (2008): 123-139.

[3] Xie, Junyuan, Linli Xu, and Enhong Chen. "Image denoising and inpainting with deep neural networks." *Advances in neural information processing systems.* 2012.

[4] A. Efros and T. Leung. "Texture synthesis by non parametric sampling." *In Proc. Int. Conf. Computer Vision (ICCV 99)*, Vol. 2, pp. 1033-1038, 1999.

[5] Buades, Antoni, Bartomeu Coll, and Jean Michel Morel. "On image denoising methods." *CMLA Preprint* 5 (2004).

[6] Dabov, Kostadin, et al. ”BM3D image denoising with shape-adaptive principal component analysis.” *SPARS’09-Signal Processing with Adaptive Sparse Structured Representations.* 2009.
[7] Mahmoudi, Mona, and Guillermo Sapiro. ”Fast image and video denoising via nonlocal means of similar neighborhoods.” *IEEE signal processing letters* 12.12 (2005): 839-842.
[8] Larrue, Tara, et al. ”Denoising Videos with Convolutional Autoencoders.” (2018).
[9] Koziarski, Michal, and Boguslaw Cyganek. ”Deep neural image denoising.” *International Conference on Computer Vision and Graphics.* Springer, Cham, 2016.
[10] Lehtinen et al. ”Noise2noise: Learning image restoration without clean data.” arXiv preprint arXiv:1803.04189 (2018)
[11] Jain, Viren, and Sebastian Seung. ”Natural image denoising with convolutional networks.” *Advances in neural information processing systems.* 2009.
[12] Liu, Po-Yu, and Edmund Y. Lam. ”Image Reconstruction Using Deep Learning.” arXiv preprint arXiv:1809.10410 (2018).
[13] Mao, Xiao-Jiao, Chunhua Shen, and Yu-Bin Yang. ”Image restoration using convolutional auto-encoders with symmetric skip connections.” arXiv preprint arXiv:1606.08921 (2016).
[14] Ansari M.A., Singh D.K. ”Review of Deep Learning Techniques for Object Detection and Classification.” *In: Verma S., Tomar R., Chaurasia B., Singh V., Abawajy J. (eds) Communication, Networks and Computing. CNC 2018. Communications in Computer and Information Science.*, vol 839. Springer, Singapore
[15] Huynh-Thu, Quan, and Mohammed Ghanbari. ”Scope of validity of PSNR in image/video quality assessment.” *Electronics letters.* 44.13 (2008): 800-801.
[16] Hasan, M. and El-Sakka. ”Improved BM3D image denoising using SSIM-optimized Wiener filter.” *M. J Image Video Proc.* 2018: 25.
[17] U. Ojha, U. Adhikari and D. K. Singh, ”Image annotation using deep learning: A review.” *In: International Conference on Intelligent Computing and Control (I2C2)*, Coimbatore, 2017, pp. 1-5.
[18] Ansari, Mohd & Kurchaniya, Diksha & Dixit, Manish. (2017). ”A Comprehensive Analysis of Image Edge Detection Techniques.” International Journal of Multimedia and Ubiquitous Engineering. 12. 1-12.