DOI: https://doi.org/10.24113/ijoscience.v8i5.486

# Sparse Convolution Neural Network for Image Denoising

Dheeraj Kumar Chaturvedi
Department of Electronics and Communication Engineering
Technocrat Institute of Technology, Excellence
Bhopal, MP, India

Saima Khan
Department of Electronics and Communication Engineering
Technocrat Institute of Technology, Excellence
Bhopal, MP, India

***Abstract:*** **In this paper, Convolution neural network (CNN) is used to enhance the efficiency and adaptability of image noise reduction. To increase de-noising efficacy, the component enlarges the receiving area of the region employing dilated and traditional convolutions in the sparse representation stage. The feature augmentation phase uses system global and regional features to increase image de-noising interpretation. The Sparse representation comprised of dilated and generalized convolutions is given to improve de-noising acceleration and efficiency. The reconstructive stage is then employed to collect noise information completely. Finally, employing the residue learning approach, this phase manages to generate the noise free image. The result analysis shows the efficiency of proposed framework over existing works.**

Keywords: Image Denoising, Convolution Neural Network, Residual Learning, Sparse Representation, Image Reconstruction.

## I. INTRODUCTION

Images become prone to the development of certain random noise during image acquisition due to the inherent physical constraints of various recording technologies. Noise is a type of signal distortion that obstructs the observation and extraction of information from images. Image noise suppression is a fundamental component of imaging processing and analysis, thus any evolution in the image denoising area aids our knowledge of fundamental image statistics and processing. Image restoration methods have become indispensable tools in the current era of computer aided analysis, with the tremendous growth in the output of digital images typically acquired in poor atmospheric/illumination conditions. Additive White Gaussian Noise (AWGN), impulse noise (salt and pepper), Poisson noise, quantization noise and speckle noise are the most common types of noise found in digital photographs. AWGN arises predominantly in analog circuits during transmission of images and the acquisition [1]. Other types of disturbances, including as impulse noise, speckle noise, Poisson noise and quantization noise are more common as a result of poor manufacturing, bit error, and insufficient photon count during image capture.

Remote sensing, military and surveillance, robotics, medical imaging and artificial intelligence are just few of the sectors where digital images can be used to give vital information. Image interpretability is irreversibly lost as a result of the pollution. As a result, image denoising techniques are widely used in remote sensing, military and surveillance, biometrics and forensics, industrial and agricultural automation, medical imaging, and biometrics and forensics. In its digital version, an image can be thought of as an encoded form of a matrix whose members are grey-level or color pixel intensity values. In the case of a video, this matrix, of course, has the third dimension of time.

Individual recognition and remote sensing are only two examples of where digital image devices have been used. The acquired image is a degraded version of the latent observation, with elements like lighting and noise corruption influencing the degradation processing. The noise is produced by the unknown latent observation during the transmission and compression procedures. To reduce the noise and recover the latent observation from a degraded image, image denoising techniques must be used. In the last 50 years, image denoising techniques have gotten a lot of attention [10].

## II. LITERATURE REVIEW

Thakur et al. [2] discussed different CNN image denoising models. It is explained how CNN models with application-oriented noise specifications are built. On BSD-68 and Set-12, the performance of each of the CNN noise models and other state-of-the-art methods is compared in terms of PSNR. Denoising performance is improved when CNN is used in conjunction with denoising filters and other iterative optimization methods.Because it unfolds the observation model of image degradation into the CNN with iterative optimization technique and back-propagation modules, PDNN outperforms all other networks for both BSD-68 and Set-12. When running on GPU, the FFDNet model is the fastest in terms of average running time. SCNN, FFDNet, NN3D, PDNN and DnCNN-S, are found to be more useful in denoising synthetic images that have been manually distorted by users.

Pardo [3] investigated the relationship between non-local denoising methods and spectral features of the graph of image neighbors was investigated in this paper. They presented a variant of NLM that calculates the parameter r automatically based on these findings. They used the connection between NLM and graph cuts to justify the suggested approach. Researchers found that this technique performs as well as the best feasible NLM outcome based on simulations. They intend to address the reason for the optimal values for h discovered in the trials with real photos in future work. The value of h was consistent across all photos, and it appears to be (2K + 1).

Knaus & Zwicker [4] discovered a modest image denoising algorithm that outperforms sophisticated state-of-the-art approaches. The approach effectively denoises 1D and 2D data, and it is expected that it will be simple to extend to greater dimension. The algorithm also has some noteworthy qualities, such as a signal hierarchy, parallelism between the noisy and guiding images and spatial and frequency domain symmetry.The researchers claim that by investigating basic methods, they will be able to develop fresh insights and make progress. The method's simplicity invites more investigation, finally resulting in greater knowledge of image denoising.

Gondara[5]demonstrated that a denoising auto-encoder built with convolutional layers may be utilized to denoise medical photos efficiently. Contrary to popular opinion, they have demonstrated that good denoising efficiency may be obtained with short training datasets, with training samples as low as 300.

Liu et al.[6] introduced image denoising algorithms based on a linear CNN model in this work. They discovered that for removing Gaussian noise, the filtering approach based on the linear CNN model performs best; for salt-and-pepper noise, the linear CNN model outperforms two classical filters and median filtering. Traditional image filters can be significantly improved by using the linear CNN model.

Shahdoosti & Hazavei [7] combined the complete variation approach and the ripplet transform to create a new hybrid denoising method. To identify textured regions from smooth ones, the suggested technique makes use of the spatial regularity idea. We picked the training examples for the classifier using ripplet, which decomposed the image into several scales and orientations. Following that, the TSVM classifier was used to separate the coefficients of high frequency subbands into two classes: textured regions and smooth regions. The mean and standard deviation of the image rebuilt from the smooth regions-labeled coefficients was then minimized while maintaining the noise-free image's resemblance to the original image.

Zhang et al. [8] introduced the IDCNN, an image denoising method based on a deep convolutional neural network. When a contaminated image is sent into the network, the developed DCNN predicts the noise image, unlike other state-of-the-art learning approaches. Separating the anticipated noisy image from the polluted image yields the latent clear image. We ran some tests to see what the designed DCNN's properties were like. We discovered that the greater network depth there is, the better the proposed denoising approach performs. Furthermore, the suggested denoising technique with an unified framework can reduce many noises with varied levels of noise at the same time.

By keeping the moving frame representing the graph of a scaled version of the image, Ghimpeteanu et al.[9] have built a framework that allows any denoising approach to take more into consideration the local geometry of the image to be denoised. Investigations on both gray-level and color images using NLM, BM3D and the VTV-based denoising method on the Kodak database revealed that our approach enhances the denoising method it is given to in terms of SSIM metrics and PSNR.The consistency of our methodology is demonstrated by the fact that they were able to increase the performance of three different types of denoising methods: a local variational method, a patch-based method, and a method combining a patch-based approach with a filtering in spectral domain approach.

## III. NETWORK ARCHITECTURE

As illustrated in Fig. 1, 17-layer denoising convolutional neural network is proposed that is divided into three processing steps: sparse representation, feature enhancement, and reconstruction. The 12-layer sparse representation is used to improve image noise removal productivity and flexibility. In sparse representation step, the module enlarges the reception size of the field using dilated and conventional convolutions to improve de-noising effectiveness. The feature enhancement step incorporates network's local and global characteristics to improve expressiveness abilities in image de-noising. For complicated noisy applications including such actual chaotic image and blind de-noising, the proposed network can swiftly capture the important noisy characteristics. To enhance de-noising speed and economy, a Sparse Block composed of dilated and common convolutions is presented. It is often used to decrease the depth. The feature enhancement step employs a lengthy route to combine information from surface and subsurface layers in order to improve the de-noising model's expressiveness capabilities. Then reconstruction step is used to thoroughly extract noise information. Lastly, this step attempts to create the de-noised image using the residual learning method.

DOI: https://doi.org/10.24113/ijoscience.v8i5.486

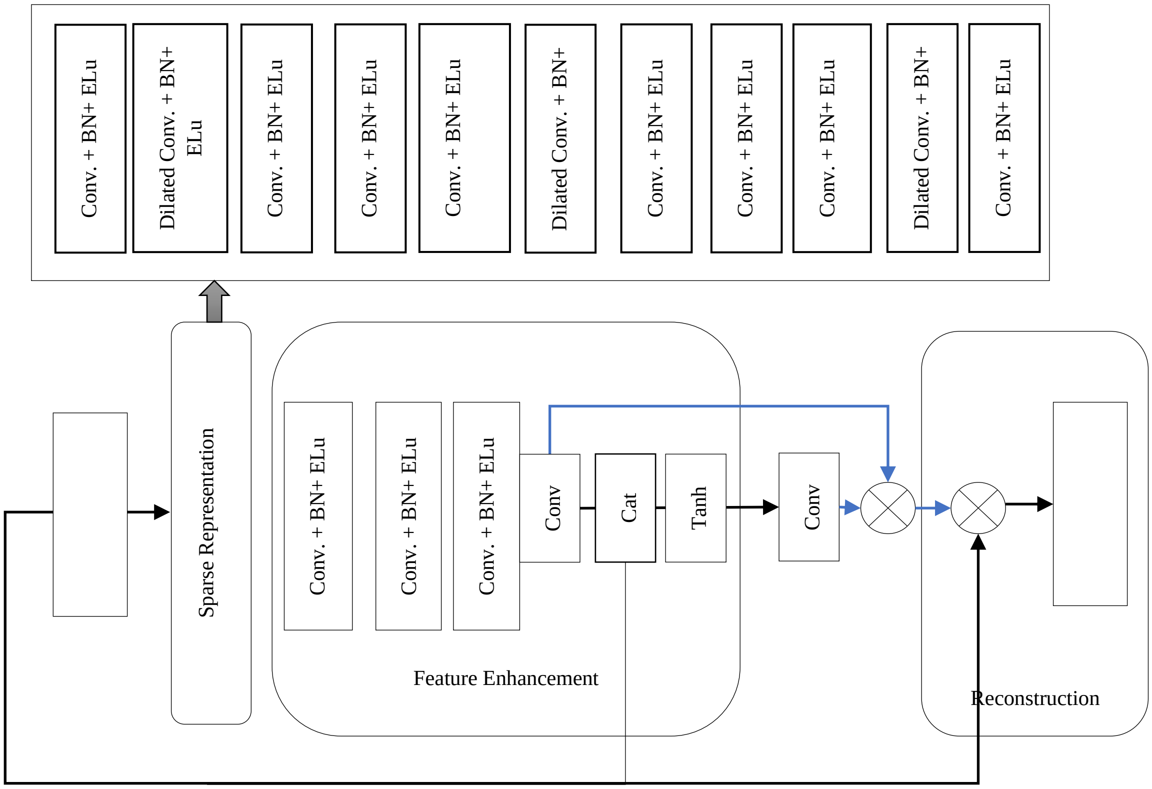


Fig. 1. Network Architecture

*Loss Function:*To make training of the proposed network easier, the paper include scale-specific losses. We observe that scale-specific losses can act as intermediary supervision on hidden layers, enhancing model openness and reducing training complexity. In proposed sparse CNN, the L2-norm reconstructive loss is also used.

$$\mathcal{L}_R = \frac{1}{2}\left\|\hat{X} - X\right\|^2 \qquad \text{(ii)}$$

Where, $\mathcal{L}_R$= loss function, $\hat{X}$ = $denoised\ image$, $X$ = noisy image

## IV. RESULT ANALYSIS

This part outlines a theoretical and simulation explanation of the suggested image noise removal technique, which will then be tested to determine efficiency. Training lasts 100 iterations, with the learning rate decreasing gradually from 1e-1 to 1e-4. To train the system for experimenting, the MatConvNet program in Matlab (R2020a) is employed.

*A. Performance Indicators*

PSNR (Peak Signal to Noise Ratio): PSNR is a highly essential technique for improving graphics performance. It is rated as follows:

$$\text{PSNR}=10\log 10\,\frac{(X*Y)}{MSE} \qquad \text{(iii)}$$

Here X and Y are the image's height and breadth, accordingly. MSE is the mean square error among the recovered image and the test image.

SSIM (Structured Similarity Index): Structural similarity has become one of the qualitative evaluation techniques used in digital signal processing (SSIM). This indexing is entirely term referring and uses certain reference images to detect chaotic or distorted images. SSIM is intended to outperform existing approaches including maximal signal to noise ratio (PSNR) and mean squared error (MSE) (MSE).

$$\text{SSIM}(x,y)=\frac{\left(2\mu_x\mu_y+c_1\right)\left(2\sigma_{xy}+c_2\right)}{\left(\mu_x^2+\mu_y^2+c_1\right)\left(\sigma_x^2+\sigma_y^2+c_2\right)} \qquad \text{(iv)}$$

When $\mu_x$=mean of x, $\mu_y$ =mean of y, $\mu_x^2$=variance of x, $\mu_y^2$=variance of y, $\mu_{xy}$ = co-variance of x and y, and c1 and c2 are factors to normalize the partition with a low denominator.

Fig. 2. Denoising Results for Proposed Methodology

Table 1: Performance Evaluation

| | Set5 | | Set12 | | BSD68 | | Biomedical Images | | SAR | |
|---|---|---|---|---|---|---|---|---|---|---|
| | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 15 | 33.48 | 0.918 | 35.25 | 0.931 | 34.491 | 0.935 | 37.85 | 0.9508 | 33.207 | 0.952 |
| 25 | 31.04 | 0.880 | 32.74 | 0.90 | 31.692 | 0.890 | 35.27 | 0.9218 | 30.361 | 0.918 |
| 30 | 30.07 | 0.859 | 31.84 | 0.886 | 30.734 | 0.87 | 34.33 | 0.9079 | 29.372 | 0.895 |
| 50 | 27.78 | 0.804 | 29.40 | 0.839 | 28.310 | 0.80 | 31.92 | 0.8673 | 27.060 | 0.840 |

Table 2: Comparative Analysis

| BSD68 | | | | Set5 | | | |
|---|---|---|---|---|---|---|---|
| | 15 | 25 | 50 | | 15 | 25 | 50 |
| DnCNN [1] | 31.72 | 29.23 | 26.23 | DnCNN | 32.86 | 30.43 | 27.18 |
| IRCNN [2] | 31.63 | 29.15 | 26.19 | IRCNN | 32.77 | 30.38 | 27.14 |
| FFDNet [3] | 31.62 | 29.19 | 26.3 | FFDNet | 32.77 | 30.44 | 27.32 |
| BRDNet[4] | 31.79 | 29.29 | 26.36 | BRDNet | 33.03 | 30.61 | 27.45 |
| SCNN[5] | 34.491 | 31.692 | 28.310 | SCNN | 33.484 | 31.045 | 27.78 |

The performance evaluation of different datasets i.e; Set5, Set12, BSD68, Biomedical Images and SAR are given in table 1 below with respect to peak signal to noise ratio and Structure Similarity Index with 15, 25, 30 and 50 noise levels. This table demonstrates that the average values of Set5 for PSNR is 30.59, similarly for Set12 is 32.30, for BSD68 is 31.30, for biomedical images is 34.84 and for SAR is 30. Likewise the average value of structure similarity index for Set5 is 0.865, for Set12 is 0.889, for BSD68 is 0.873, for Biomedical images 0.911 and for SAR is 0.901. Table 2 shows the comparative analysis between BSD68 and Set5 with different networks like DnCNN, IRCNN, FFDNet, BRDNet, SCNN with 15, 25,50 Noise levels. In BDS68, the greater values are found to be 34.491 at 15, 31.692 at 25 and 28.310 at 50 on employing SCNN and the lowest values are found to be 31.62 at 15, 29.19 at 25 and 26.30 at 50 on employing FFDNet. In Set5 the greater values are found to be 33.484,31.045 and 27.78 at 15, 25 and 50 on employing SCNN and the lowest values are found to be 32.77, 30.38 and 27.14 at 15,25 and 50 noise levels on employing IRCNN. And average values on using DnCNN is 29.06, IRCNN is 28.99, FFDNet is 29.03, BRDNet is 29.14, SCNN is 31.497 for BSD68. For Set5, on using DnCNN is 30.15, IRCNN is 30.09, FFDNet is 30.17, BRDNet is 30.36 and SCNN is 30.76.

## V. CONCLUSION

This work demonstrates, a noise reducing convolutional network to enhance the productivity of noise removal and the complete simulation takes place on the MatConvNet program in MATLAB (R2020a) with 100 iterations. The comparison with different sets of images, with respect of peak signal to noise ratio and structure similarity index is also done to evaluate the performance efficiency. The performance evaluation is measured by different datasets on 15, 25, 30 and 50 noise levels. Further comparative analysis on BSD68 and Set5 dataset are also done and shows the improvement of proposed model over existing works.

## References

[1] Goyal, B., Dogra, A., Agrawal, S., Sohi, B. S., & Sharma, A. (2020). Image denoising review: From classical to state-of-the-art approaches. *Information Fusion*, *55*(September 2019), 220–244. https://doi.org/10.1016/j.inffus.2019.09.003

[2] Thakur, R. S., Yadav, R. N., & Gupta, L. (2019). State-of-art analysis of image denoising methods using convolutional neural networks. IET Image Processing, 13(13), 2367–2380. https://doi.org/10.1049/iet-ipr.2019.0157

[3] Pardo, A. (2011). Analysis of non-local image denoising methods. Pattern Recognition Letters, 32(16), 2145–2149. https://doi.org/10.1016/j.patrec.2011.06.022

[4] Knaus, C., & Zwicker, M. (2013). DUAL-DOMAIN IMAGE DENOISING Claude Knaus Matthias Zwicker. 4, 440–444.

[5] Gondara, L. (2016). Medical Image Denoising Using Convolutional Denoising Autoencoders. IEEE International Conference on Data Mining Workshops, ICDMW, 0, 241–246. https://doi.org/10.1109/ICDMW.2016.0041

[6] Liu, Z., Yan, W. Q., & Yang, M. L. (2018). Image denoising based on a CNN model. Proceedings - 2018 4th International Conference on Control, Automation and Robotics, ICCAR 2018, 389–393. https://doi.org/10.1109/ICCAR.2018.8384706

[7] Shahdoosti, H. R., & Hazavei, S. M. (2018). Combined ripplet and total variation image denoising methods using twin support vector machines. Multimedia Tools and Applications, 77(6), 7013–7031. https://doi.org/10.1007/s11042-017-4618-9

[8] Zhang, F., Cai, N., Wu, J., Cen, G., Wang, H., & Chen, X. (2018). Image denoising method based on a deep convolution neural network. IET Image Processing, 12(4), 485–493. https://doi.org/10.1049/iet-ipr.2017.0389

[9] Ghimpeteanu, G., Batard, T., Bertalmio, M., & Levine, S. (2016). A decomposition framework for image denoising algorithms. IEEE Transactions on Image Processing, 25(1), 388–399. https://doi.org/10.1109/TIP.2015.2498413

[10] Zhang, F., Cai, N., Wu, J., Cen, G., Wang, H., & Chen, X. (2018). Image denoising method based on a deep convolution neural network. IET Image Processing, 12(4), 485–493. https://doi.org/10.1049/iet-ipr.2017.0389