

Image Enhancement and Denoising in Extreme Low-Light Conditions

Utsav Krishnan, Ayush Agarwal, Avinash Senthil, Pratik Chattopadhyay

Abstract: Image noise refers to the specks of false colors or artifacts that diminish the visual quality of the captured image. It has become our daily experience that with affordable smart-phone cameras we can capture high clarity photos in a brightly illuminated scene. But using the same camera in a poorly lit environment with high ISO settings results in images that are noisy with irrelevant specks of colors. Noise removal and contrast enhancement in images have been extensively studied by researchers over the past few decades. But most of these techniques fail to perform satisfactorily if the images are captured in an extremely dark environment. In recent years, computer vision researchers have started developing neural network-based algorithms to perform automated de-noising of images captured in a low-light environment. Although these methods are reasonably successful in providing the desired de-noised image, the transformation operation tends to distort the structure of the image contents to a certain extent. We propose an improved algorithm for image enhancement and de-noising using the camera's raw image data by employing a deep U-Net generator. The network is trained in an end-to-end manner on a large training set with suitable loss functions. To preserve the image content structures at a higher resolution compared to the existing approaches, we make use of an edge loss term in addition to PSNR loss and structural similarity loss during the training phase. Qualitative and quantitative results in terms of PSNR and SSIM values emphasize the effectiveness of our approach.

Keywords: Image Noise, PSNR, ISO, Illumination, Network based Algorithms.

I. INTRODUCTION

In the past, the quality of digital photographs was mostly dependent on the availability of high-end camera sensors and optics. Thus, for capturing good quality digital photographs, it was mandatory to use cameras with sophisticated hardware which was cost-intensive. With the gradual advancement of digital photography technology, nowadays we no longer rely on cameras with high-end hardware support to capture good photographs. Rather, we are more interested in procuring cameras with good software features, since these are capable of providing good quality photographs within an affordable price range.

This is also evident from the fact that Google's Pixel phone, which consists of only a 12.3-megapixel camera, is capable of

producing high-quality images of even a fast-moving object. The quality of these images is comparable with that of snapshots captured by a high-end digital single-lens reflex camera (DSLR). The HDR+ burst photography technology is the reason behind the outstanding performance of Google's Pixel camera [1]. It takes as input a sequence of images captured in rapid succession to generate a de-blurred and enhanced image. Apart from this, a number of de-noising techniques have been developed that work on a single noisy image, e.g., (i) threshold-based de-noising in wavelet domain [2, 3], (ii) Bayesian denoising in wavelet domain [4], (iii) directional transformation based denoising [5], (iv) non-local means [6, 7, 8], etc. Block Matching 3D (BM3D) [9] is one of the state-of-the-art denoising algorithms developed for removing additive white Gaussian noise from images. However, each of the above methods has its own limitation in the sense that it is successful in eliminating only specific types of noise. Since, in a real environment, the noise distribution is not known, the above algorithms fail to perform satisfactorily always.

Deep neural networks are popular for their strong generalization ability in learning the underlying function if an adequate number of input-output patterns are available, and these networks have been extensively used in the recent past to perform a number of image-to-image translation tasks. In the present work, we consider a data set of ground truth images is available that consists of (a) noisy images captured in extreme low-light conditions and (b) corresponding denoised images. We first extract important color channels from the raw sensor data and next employ an end-to-end neural network that is trained with three different loss functions to learn the mapping from noisy to denoised images: (i) the PSNR loss that minimizes the differences between the pixel intensities of the two images, (ii) the SSIM loss that minimizes the differences between the luminance, contrast and structure between the generated de-noised image and the ground truth denoised image, and (iii) the Edge loss that preserves the edges and other high frequency components present in the ground truth denoised image. The trained model is made available at <https://github.com>, which readers can access by clicking here.

Our model does not take into account any prior information regarding the type and distribution of noise, and hence can be deployed to perform image denoising in any real-world environment. The ground truth data used for training our model has been captured in extreme low-light conditions, and hence the trained network attains the capability to denoise images captured in dark. Neural network-based prediction is also much more time-efficient than traditional image processing-based approaches like BM3D, and the software developed out of this work can be easily integrated with

Revised Manuscript Received on November 08, 2019.

* Correspondence Author

Utsav Krishnan, Department of Computer Science and Engineering, Indian Institute of Technology (Banaras Hindu University), Varanasi

Ayush Agarwal, Department of Electronics and Communication Engineering, Motilal Nehru National Institute of Technology Allahabad

Avinash Senthil, Department of Electrical and Electronics Engineering, National Institute of Technology, Tiruchirappalli

Pratik Chattopadhyay, Department of Computer Science and Engineering, Indian Institute of Technology (Banaras Hindu University), Varanasi

smart-phone cameras to improve the quality of captured images in a cost-effective manner. Experimental results show that the proposed approach also outperforms the state-of-the-art denoising algorithms when deployed for recovering images captured in extreme low-light and noisy environments.

II. RELATED WORK

The present problem can be viewed as a combination of two important subproblems of image processing: (i) image contrast enhancement, and (ii) denoising in a low-light environment. In this section, we review the existing literature pertaining to each of the above sub-problems. Traditional contrast enhancement techniques such as histogram equalization, power-law transformation, log-transformation, etc., are used to improve the contrast of low-light images [10]. While power-law and log-transformation techniques require manual selection of transformation parameters based on the intensity distribution of the image, histogram equalization automatically spreads out the most frequently occurring intensity values to improve the overall contrast. Other similar image processing-based contrast enhancement methods include brightness preserving bi-histogram equalization (BBHE), and quantized bi-histogram equalization (QBHE) [11], dark channel inverse [12, 13], wavelet transform based enhancement [14], Retinex model-based enhancement [15], and low-light image enhancement by illumination map estimation [16]. The above-mentioned techniques are suitable for improving the contrast of an image by applying an intensity transformation operation on every pixel. However, these are not effective enough in solving the problem considered in this work, i.e., de-noising images captured in a low-light environment. Hence, we have provided only the relevant citations without discussing the above methods in further depth.

Existing work on image de-noising can be categorized as either non-neural network-based [9, 17, 18, 19, 20, 21], or neural network-based [22, 23, 24]. Among these, the neural network-based methods have gained more popularity due to their high generalization power and robust performance against various types of noise, including natural noise. Before discussing the recent neural network-based de-noising techniques, we briefly highlight some widely used primitive non-neural based approaches to image de-noising. One such algorithm is the block-matching and 3D filtering (BM3D) algorithm [9], which is till-date regarded as one of the most popular algorithms for image de-noising due to its capability to effectively eliminate additive white Gaussian noise and speckle noise from images. The algorithm works by finding groups of similar patches of the same size in the image and stacks each group together to construct a 3D cylinder. A linear transform followed by Wiener filtering is next applied on each fragment group followed by inversion of the linear transform to filter out the noise. Finally, the image is transformed back to its two-dimensional form. One drawback of this approach is that it is unable to perform blind denoising, i.e., the noise level in the image must be mentioned while performing the algorithm. Moreover, it is not successful in handling real noise. The work in [17] describes an adaptive total variation model based on a new orientation information metric for image de-noising. Other popular non-neural network-based approaches include denoising via wavelet-domain processing

[18], sparse coding [19], nuclear norm minimization [20], and K-SVD algorithm [21], etc.

In recent years, several neural network-based image de-noising and enhancement algorithms have been proposed. For example, in [22] Zhang et al. proposed a deep network (abbreviated as DnCNN), which is a residual learning-based convolutional neural network for de-noising images corrupted at different noise levels. This network is capable of handling additive white Gaussian noise more effectively than any other non-neural network-based approaches, but it performs poorly in de-noising images degraded with real noise. In other words, DnCNN cannot de-noise images effectively if the noise distribution is unknown. Moreover, this network is unable to improve the image contrast, and hence cannot be potentially used for low-light image enhancement. A stacked sparse denoising auto-encoder (abbreviated as LLNet) for image enhancement has been proposed by Lore et al. in [23] which has been shown to work well in de-noising low-light gray-scale images. However, we observe that the performance of this method reduces drastically if multi-channel images are given as input, and hence it is not suitable enough for practical application.

Inspired by the classical BM3D denoising technique [9] described earlier, in [25] Yang et al. proposed a learning-based technique termed as BM3D-Net that maps the entire computational pipeline of the BM3D approach into the convolutional neural network architecture. BM3D-Net has been seen to achieve better performance in image de-noising as compared to the traditional BM3D method. This approach works satisfactorily on gray-scale or depth-images, corrupted with Gaussian noise. But its applicability in removing real noise has not been tested yet. The method described in [24] has an objective similar to our work, i.e., it also attempts to enhance and remove noise from images captured in a low-light environment. The authors use a U-Net generator for de-noising and enhancing an input noisy image, but the training of the network is accomplished by considering the mean squared error (MSE) metric as the only loss function. Results given in [24] show that this technique is able to perform satisfactorily for scene images corrupted with natural noise. However, due to the use of a very simplistic loss function, the application of the algorithm on images with high-frequency components (such as edges and textures) causes a loss in structural information along with de-noising. The present work can be considered as an extension of the approach given in [24], in which we improve the network training phase by employing more appropriate loss functions to perform de-noising, contrast enhancement, and retain structural information simultaneously. The rest of the paper is organized as follows. In Section 4, we explain the steps of our algorithm in detail with the help of a block diagram, Section 5 presents the experimental studies conducted by our team along with the results obtained. Finally, Section 6 concludes the paper with insights into possible future directions of work.

III. PROPOSED APPROACH

A block diagram of the proposed approach is shown in Figure 1.

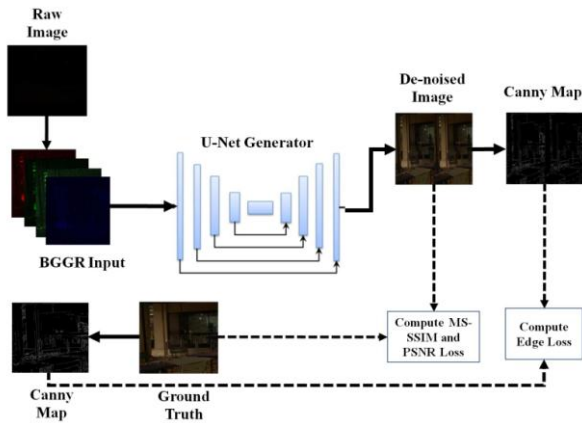


Figure 1: A block diagram of the proposed approach

With reference to the figure, the first step is to unpack the one-dimensional sensor data into four-dimensional BGR color space based on the Bayer Color Filter Array (CFA) [26]. For Bayer arrays, we pack the input into four channels (two green, one blue, and one red) and accordingly reduce the dimension of channels by a factor of two, following which we subtract the black level. Next, the data is scaled by the desired amplification ratio (we have used x100 and x300) to enhance the brightness of the image. This packed and amplified data is finally passed through a deep network to perform de-noising in an end-to-end manner, as discussed next.

A U-Net architecture has been used as the denoising network in the present study. U-Net is a type of auto-encoder network with skip connections from the encoding layers to the decoding layers. These skip connections help in preserving the content of the input image at a higher resolution than the traditional auto-encoders. The proposed approach is completely supervised, and it involves training the network by adjusting its weights in multiple epochs using Adam optimizer. The dual objective of image enhancement and de-noising is met by considering three different loss functions, namely (a) the PSNR Loss, (b) the MS-SSIM Loss and (c) the Edge Loss (refer to Figure 1). The PSNR and the MS-SSIM losses optimize the structural, luminance, color and contrast related features of the generated image with respect to the ground truth thereby helping in effective de-noising. On the other hand, the Edge Loss helps in image enhancement by retaining the high-frequency components of the image such as edges and textures. The computation of the individual loss functions is discussed in detail next.

PSNR Loss

The peak signal-to-noise ratio metric between two images has been extensively used in several studies in the past to measure the reconstruction quality of an image [27]. This value also reveals the amount of noise present in a target image with respect to a reference image. Higher the value of the metric, lesser is the amount of noise present in the target image. Considering the network generated de-noised image at each epoch as the target, and the corresponding ground truth denoised image as the reference, the peak signal-to-noise ratio metric (PSNR) is computed as per (1):

$$PSNR = 10 \log_{10} \left(\frac{MAX^2}{MSE} \right) \quad (1)$$

where MAX represents the maximum value of any pixel in the image, and MSE denotes the mean squared error between the output and the corresponding ground truth image. Since our ultimate objective is to tune the network parameters to match the generated image with the ground truth denoised image, we will try to increase the value of $PSNR$ in successive epochs. (1) can be re-written as:

$$PSNR = 20 \log_{10}(MAX) - 10 \log_{10}(MSE) \quad (2)$$

Since the value of MAX for both the generated and ground-truth denoised images will be comparable, increasing the value of $PSNR$ can be said to be equivalent to decreasing the value of MSE . The network is trained in multiple epochs by considering the value of MSE as the PSNR loss (denoted by L_{PSNR}). Mathematically,

$$L_{PSNR} = \frac{1}{N} \sum_{p \in P} (X(p) - Y(p))^2 \quad (3)$$

where X and Y respectively denote the generated and ground-truth denoised images, while $X(p)$ and $Y(p)$ represent the intensities at a particular pixel p corresponding to the two images X and Y , and N denotes the total number of pixels in the image.

MS-SSIM Loss

The above loss term alone is not sufficient to retain the luminance, contrast and structural similarity of the denoised image with respect to the input image. Hence, a second loss term is considered which is termed as the *MS-SSIM Loss*. It takes into account the structural and contrast related aspects of an image properly in order to generate an aesthetically pleasing de-noised version of the input noisy image. Pixel-wise computation of the SSIM value [28] between the generated image and the ground-truth denoised image is done according to (4), by considering windows x and y of the same size around each pixel. For a particular pixel p , $SSIM(p)$ is defined as:

$$SSIM(p) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (4)$$

where C_1 and C_2 are stabilization parameters, μ_x and σ_x respectively represent the average and variance of pixel intensities in the patch x , whereas σ_{xy} denotes the covariance between x and y . The notations μ_y and σ_y can also be defined similarly. We have set the value of both C_1 and C_2 to $0.03L$, where L represents the maximum possible intensity level in an image (i.e., for an 8-bit image, $L=255$). Once the SSIM value for each pixel is obtained following (4), the SSIM Loss ($L_{MS-SSIM}$) is computed as follows:

$$L_{MS-SSIM} = \frac{1}{N} \sum_{p \in P} (1 - SSIM(p)) \quad (5)$$

In the present work, we have used the Multi-scale SSIM (MS-SSIM) [29] which typically computes SSIM at multiple scales through a process of sub-sampling, and is highly effective in finding a similarity index between any two given images [29, 30, 31].

Edge Loss

Both the generated de-noised image and the ground truth denoised image are in sRGB space. We propose a new loss metric called the Edge loss in order to preserve minute artifacts in the image. First, we compute a Canny edge map from each of the two images. and then compare these edge maps by means of L2 loss. If X_e and Y_e respectively denote the edge maps corresponding to images X and Y , then the resulting Edge loss function L_{Edge} is computed pixel-wise as follows:

$$L_{Edge} = \frac{1}{N} \sum_{p \in P} (X_e(p) - Y_e(p))^2, \quad (6)$$

where p denotes a particular pixel location.

Consider at each epoch during training, M image pairs are used, and the average of these M errors is used to adjust the weights before the successive epoch. If L_{PSNR}^m , $L_{MS-SSIM}^m$, and L_{Edge}^m respectively represent the PSNR loss, MS-SSIM loss and Edge loss for image m at a particular epoch, and L_{avg} denotes the average instantaneous error that is to be back-propagated, then

$$L_{avg} = \frac{1}{M} \sum_{m=1}^M (\alpha L_{PSNR}^m + \beta L_{MS-SSIM}^m + \gamma L_{Edge}^m), \quad (7)$$

where α , β , and γ are user-defined constants. Training of the deep U-Net generator is carried out for a maximum of 4000 epochs, or till the value of L_{avg} in two consecutive epochs does not undergo significant variation.

IV. EXPERIMENTAL RESULTS

The proposed algorithm has been implemented on a system having 64 GB RAM, one i9-18 core processor, along with three GPUs: one Titan Xp with 12 GB RAM, 12 GB frame-buffer memory and 256 MB BAR1 memory and two GeForce GTX 1080 Ti with 11 GB RAM, 11 GB frame-buffer memory and 256 MB of BAR1 memory. For the evaluation of our algorithm, we use the *See in the Dark* (SID) data set [24]. This data consists of 5094 raw short-exposure images, each with a corresponding long-exposure image. The short-exposure images were captured under different shutter speeds, so a single long-exposure will be ground truth for several short-exposure images. The number of distinct long exposure images is 424. The data set consists of images for both indoor and outdoor scenes. The indoor images appear to be darker than that of the outdoor images. Figure 2(a) shows five sample noisy images from the SID data set, while Figure 2(b) shows the corresponding de-noised versions. In the original data set, the noisy images are very dark since they are captured in an extreme low-light environment. But for ease of visualization, here we show the noisy images by amplifying their intensity by a factor 300.

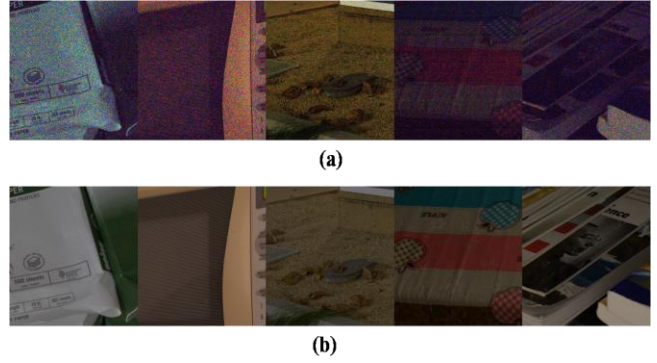


Figure 2: Sample images from the SID data set: (a) noisy input and (b) ground truth.

Training of the U-Net generator is done by considering raw-sensor data of short exposure images as input and the corresponding long-exposure image in the sRGB space as the desired output. This network considers the 512x512-dimensional image as input. Since, the dimensions of the images present in the SID data is significantly high, at each epoch during training, we randomly crop a 512x512 patch from each image in the gallery set and apply random flipping and rotation for data augmentation. During the deployment scenario, each test image is resized to the same dimensions before passing through the trained network. The value of the learning rate used in training the generator is initially set to 10^{-4} , which is then gradually reduced to 10^{-5} after completion of 2000 epochs.

In the first experiment, we study the effect of using different combinations of the loss terms, L_{PSNR} , $L_{MS-SSIM}$, and L_{Edge} (refer to Section 4) in generating the final denoised image. Quantitative results are shown in Table 1 for different combinations of the loss terms by means of PSNR and SSIM values, computed by using (1) and (5), respectively. The test set for these experiments consists of 100 raw short-exposure images. It can be seen from the table that aggregation

Terms in Loss Function	PSNR	SSIM
$L_{MS-SSIM}, L_{PSNR}, L_{Edge}$	32.11	0.8537
$L_{MS-SSIM}, L_{PSNR}$	31.95	0.8533
$L_{MS-SSIM}, L_{Edge}$	30.78	0.8524
L_{PSNR}, L_{Edge}	31.83	0.8474

Table 1: PSNR and SSIM values obtained for different combinations of the loss terms

of all the three-loss terms provides the best denoising performance in terms of PSNR and SSIM. We also observed that the best results were obtained on assigning equal weight to all loss terms (i.e., by assigning $\alpha=\beta=\gamma=1$). Giving more weight to $L_{MS-SSIM}$ and L_{PSNR} fails to preserve the artifacts present in an image properly. On the other hand, imposing a higher weight on L_{Edge} fails to preserve contrast and color at a sufficiently high resolution.

We next present a comparative performance analysis of our proposed approach with some other popular image enhancement and de-noising techniques, namely, [9, 23, 24]. In Figure 3, we show the denoised images provided by our approach and each of the above techniques on a sample set of six noisy images from the SID data. In this figure, the first column (i.e., column a) shows the ground-truth denoised image, the second column (i.e., column (b)) shows the results

obtained by applying our proposed approach, while the third, fourth and fifth columns (i.e., columns (c), (d) and (e)), respectively represent the denoised images obtained by applying [24], [9], and [23], respectively. It can be visually seen from the figure that LLNet has the worst performance among the compared methods. This is due to the fact that this network is specifically designed for contrast enhancement in images and so it is ineffective in eliminating noise from images. Since a scaled-down version of each of the images has been presented in the figure, it is difficult to visually judge which among our proposed approach, [24], and [9] is performing the best. For the purpose of critical examination, we crop out certain portions (shown by blue rectangles) from each image in Figure 3, and present them in Figure 4 by zooming into these images. These blue rectangles consist of some high-frequency information, e.g., textual regions, fine textures, etc. It can be clearly observed from the zoomed-in versions of the images that in terms of preserving high-frequency details, such as text and texture, the proposed approach performs the best among all the techniques used in the comparative study. Also, from visual observation, it appears that the de-noising quality of our approach is comparable with the work of [24] only. It outperforms each of the other de-noising approaches to a considerable extent.

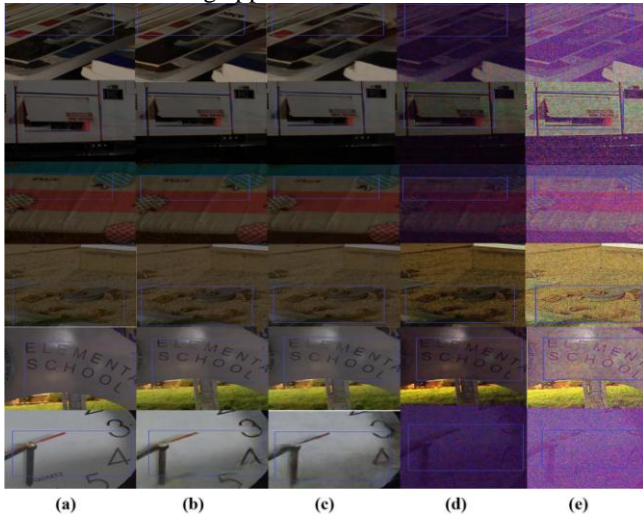


Figure 3: Comparative analysis: (a) ground truth, denoised images obtained by applying (b) proposed approach, (c) [24], (d) [9], (e) [23]

De-noising Approach	PSNR	SSIM	Response Time (secs)
Proposed approach	32.11	0.8537	0.18
[24]	31.69	0.8458	0.18
[9]	19.14	0.6041	21.12
[23]	10.97	0.2567	26.47

Table 2: Quantitative performance analysis of the proposed work and the other compared denoising approaches

Finally, we present a quantitative comparative performance analysis of our approach with each of the compared methods by means of average PSNR and average SSIM values between the generated and the ground-truth denoised images. Results are shown in Table 2 along with the corresponding average response times. The analysis was carried out on a randomly selected sample of 100 test images from the SID data. As

expected, it can be seen from the table that here also our approach outperforms each of the other denoising techniques in terms of both PSNR and SSIM values. The method that performs closest to our approach is the work in [24], which also exploits the generalization capability of deep neural networks to generate de-noised images with high PSNR and SSIM values. However, as also shown in Figure 4, due to the absence of the edge-preserving loss term during the training of the network, it is unable to retain high-frequency details in an image as good as our approach. The average response times for both the proposed approach and that of [24] are 0.18 seconds, which is significantly less when compared with the other methods used in the study. We have also studied the de-noising capability of the BM3D-Net [25] algorithm using the implementation given in [32]. The obtained results were highly noisy and even worse than the output of [23]. Hence, we have not presented its results in the paper. Based on the above set of experiments, it can be concluded that the proposed work performs the best in recovering images captured in an extreme low-light environment and corrupted with natural noise.

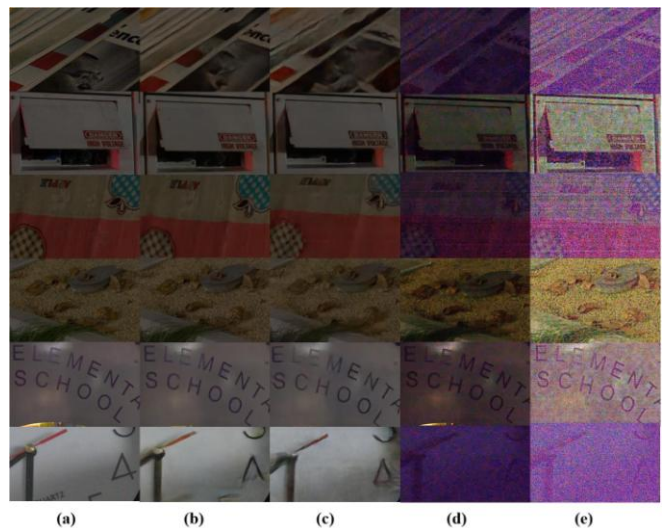


Figure 4: Zoomed-in versions of the cropped regions corresponding to: (a) ground truth, (b) proposed approach, (c) [24], (d) [9], (e) [23] of Figure 3

V. CONCLUSION

In this paper, we present an effective approach for de-noising and enhancement of low-light images using a deep neural network. With the inclusion of an edge loss term during the network training phase, we could successfully preserve detailed edge and texture related details in the denoised images that existing de-noising approaches failed to achieve. Although the use of the other loss terms is also important in generating de-noised images of high quality, the use of edge maps to preserve fine edges in an image is a novel contribution of this work, which has definitely improved the visual quality of the denoised images. The proposed work substantially improves the state-of-the-art research on image de-noising and enhancement which is evident from the high PSNR and SSIM values of our approach obtained during experimentation. In this work, we have worked on low-light image enhancement and de-noising.

But the same approach can also be extended to perform de-noising in videos, which may be studied in the future.

ACKNOWLEDGMENTS

The authors would like to thank the Head of the Department of Computer Science and Engineering Department, Indian Institute of Technology (Banaras Hindu University), Varanasi, the Head of the Department of Electronics and Communication Engineering, Motilal Nehru National Institute of Technology, Allahabad and the Head of the Department of Electrical and Electronics Engineering, National Institute of Technology, Tiruchirappalli for their kind support. The authors also express their sincere gratitude to NVIDIA for supporting their research with a Titan Xp GPU.

REFERENCES

1. Samuel W Hasinoff, Dillon Sharlet, Ryan Geiss, Andrew Adams, Jonathan T Barron, Florian Kainz, Jiawen Chen, and Marc Levoy. Burst Photography for High Dynamic Range and Low-Light Imaging on Mobile Cameras. *ACM Transactions on Graphics*, 35(6):192, 2016.
2. Wei-Qiang Zhang and Guo-xiang Song. Signal De-noising in Wavelet Domain based on a New Kind of Thresholding Function [J]. *Journal of Xidian University*, 2:030, 2004.
3. Tie Cai and Jie Zhu. Adaptive Selection of Optimal Decomposition Level in Threshold De-noising Algorithm based on Wavelet. *Control and Decision*, 21(2):217, 2006.
4. Eero P Simoncelli. Bayesian Denoising of Visual Images in the Wavelet Domain. In *Bayesian Inference in Wavelet-Based Models*, pages 291–308. Springer, 1999.
5. Ramin Eslami and Hayder Radha. The Contourlet Transform for Image Denoising using Cycle Spinning. In *37th Asilomar Conference on Signals, Systems & Computers*, volume 2, pages 1982–1986. IEEE, 2003.
6. Antoni Buades, Bartomeu Coll, and J-M Morel. A Non-Local Algorithm for Image Denoising. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 60–65, 2005.
7. Buades, Antoni and Coll, Bartomeu and Morel, Jean-Michel. Non-Local Means Denoising. *Image Processing On Line*, 1:208–212, 2011.
8. Pierrick Coup'ée, Pierre Yger, and Christian Barillot. Fast Non Local Means Denoising for 3D MR Images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 33–40. Springer, 2006.
9. Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. BM3D Image Denoising with Shape-Adaptive Principal Component Analysis. In *Workshop on Signal Processing with Adaptive Sparse Structures Representations*, 2009.
10. Rafael C Gonzalez and E Richard. *Digital Image Processing*. ed: Prentice Hall Press, ISBN 0-201-18075, 8, 2002.
11. Manpreet Kaur and Jasdeep Kaur. Survey of Contrast Enhancement Techniques based on Histogram Equalization. 2(7):138–141, 2011.
12. Kaiming He, Jian Sun, and Xiaoou Tang. Single Image Haze Removal Using Dark Channel Prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(12):2341–2353, 2011.
13. Bin Xie, Fan Guo, and Zixing Cai. Improved Single Image Dehazing Using Dark Channel Prior and Multi-scale Retinex. In *International Conference on Intelligent System Design and Engineering Application*, volume 1, pages 848–851, 2010.
14. Sachin D Ruikar and Dharmal D Doye. Wavelet Based Image Denoising Technique. *International Journal of Advanced Computer Science and Applications*, 2(3):49–53, 2011.
15. Mading Li, Jiaying Liu, Wenhan Yang, Xiaoyan Sun, and Zongming Guo. Structure-Revealing Low-Light Image Enhancement via Robust Retinex Model. *IEEE Transactions on Image Processing*, 27(6):2828–2841, 2018.
16. Xiaojie Guo, Yu Li, and Haibin Ling. LIME: Low-Light Image Enhancement via Illumination Map Estimation. *IEEE Transactions on Image Processing*, 26(2):982–993, 2017.
17. C. Wu, W. Liu, and X. Guo. Adaptive Total Variation Model for Image Denoising based on Modified Orientation Information Measure. In *3rd International Congress on Image and Signal Processing*, volume 2, pages 616–620, 2010.
18. Javier Portilla, Vasily Strela, Martin J Wainwright, and Eero P Simoncelli. Adaptive Wiener Denoising Using A Gaussian Scale Mixture Model In The Wavelet Domain. In *International Conference on Image Processing*, volume 2, pages 37–40. IEEE, 2001.
19. Yibin Tang, Ying Chen, Ning Xu, Aimin Jiang, and Lin Zhou. Image Denoising via Sparse Coding Using Eigenvectors of Graph Laplacian. *Digital Signal Processing*, 50:114–122, 2016.
20. S. Gu, L. Zhang, W. Zuo, and X. Feng. Weighted Nuclear Norm Minimization with Application to Image Denoising. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2862–2869, June 2014.
21. Songyuan Tang and Jian Yang. Image Denoising using K-SVD and NonLocal Means. In *IEEE Workshop on Electronics, Computer and Applications*, pages 886–889, 2014.
22. Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. *CoRR*, abs/1608.03981, 2016.
23. Kin Gwn Lore, Adedotun Akintayo, and Soumik Sarkar. LLNet: A Deep Autoencoder Approach to Natural Low-light Image Enhancement. *CoRR*, abs/1511.03995, 2015.
24. Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to See in the Dark. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3291–3300, 2018.
25. D. Yang and J. Sun. BM3D-Net: A Convolutional Neural Network for Transform-Domain Collaborative Filtering. *IEEE Signal Processing Letters*, 25(1):55–59, 2018.
26. Bryce E Bayer. Color Imaging Array, July 1976. US Patent 3,971,065.
27. Wikipedia contributors. Peak Signal-to-Noise Ratio - Wikipedia, The Free Encyclopedia, 2019. [Online; accessed 31-July-2019].
28. Wikipedia contributors. Structural Similarity - Wikipedia, The Free Encyclopedia, 2019. [Online; accessed 31-July-2019].
29. Zhou Wang, Eero P Simoncelli, and Alan C Bovik. Multiscale Structural Similarity for Image Quality Assessment. In *37th Asilomar Conference on Signals, Systems & Computers*, 2003, volume 2, pages 1398–1402. IEEE, 2003.
30. Jacob Søgaard, Luk'a's Krasula, Muhammad Shahid, Dogancan Temel, Kjell Brunnstr'om, and Manzoor Razaak. Applicability of Existing Objective Metrics of Perceptual Quality for Adaptive Video Streaming. *Electronic Imaging*, 2016(13):1–7, 2016.
31. Richard Dosselmann and Xue Dong Yang. A Comprehensive Assessment of the Structural Similarity Index. *Signal, Image and Video Processing*, 5(1):81–91, 2011.
32. legendongary. BM3D-Net <https://github.com/legendongary/BM3D-Net>, 2018.