Image Denoising Method based on Deep Learning using Improved U-net

Jaewook Han, Jinwon Choi, and Changwoo Lee*

School of Information, Communications and Electronics Engineering, The Catholic University of Korea / Pucheon-City, Korea

* Corresponding Author: Changwoo Lee, changwoo@catholic.ac.kr

Abstract: Various methods, including block-matching and 3D filtering (BM3D), have been proposed for image denoising. Recently, studies on deep learning methods for image denoising have been on the rise. In this paper, we propose a new structure for a deep neural network that improves image denoising performance. Among the existing deep neural networks, we improve U-net, which is widely used for image restoration, through the inclusion of pre-processing and post-processing and by modifying each of its stages. Extensive simulations show that the proposed structure performs very well for a wide range of noise levels with a single trained parameter, and it exhibits superior image denoising performance compared to conventional deep neural networks.

Keywords: Image denoising, Deep learning, U-net, New structure, Improved U-net

1. Introduction

Various image denoising methods have been studied, for example, when images are damaged by Gaussian noise, impulse noise, and speckle noise [1-9]. The nonlocal mean (NLM) technique and the block-matching and 3D filtering (BM3D) technique, which eliminate noise by calculating the weighted sum using similarity for each image patch in an entire image, show very good denoising performance [3, 4]. In recent years, deep learning methods, which have excellent performance in various image processing fields, have been studied for application to image denoising, showing performance superior to conventional image denoising techniques [5-9].

In this paper, we propose an efficient deep neural network structure to improve image denoising performance by improving the structure of U-net, which is widely used for image restoration. The proposed structure adds pre-processing and post-processing to the conventional U-net structure while also adding a convolution layer in addition to a shortcut for each stage of U-net. Since the proposed structure improves the convergence performance of the deep neural network when generating the target image, it can be used not only for denoising but also for various image restoration applications. By training the proposed structure using images with various noise intensities, noise at various intensities can be removed with a single trained parameter.

Extensive computer simulations show that the proposed method yields superior denoising performance compared to BM3D and other deep learning methods.

2. Image Denoising Method

To remove noise such as Gaussian noise, impulse noise and speckle noise [1, 2], starting with a median filter, various denoising methods using the characteristics of the frequency bands of the image and noise have been studied. A denoising technique that utilizes high-frequency characteristics of noise has a problem in that high-frequency components of the original image are also lost. The NLM method shows very good performance in image denoising by calculating the weighted sum from the entire image using the local similarity of each patch [3]. In particular, the BM3D technique, which groups image patches into a 3D structure and precisely calculates weights, was identified as the technique that showed state-of-the-art denoising performance prior to using a deep learning technique [4].

Since deep learning methods have shown excellent performance in various image processing fields, a lot of research has been conducted into applying deep learning to image denoising [5-9]. Zhang et al. proved that the deep convolutional neural network (CNN) structure can be applied to image denoising to achieve excellent denoising

performance [5]. A CNN using a variable split technique was proposed to reduce the number of computations for image denoising without degrading performance [6], and FFDNet was proposed, which can handle a wide range of noise levels and can improve convergence speed using a noise level map as input [7]. Tian et al. proposed ADNet using an attention-guided CNN [8]. These CNN methods have been proven to show superior denoising performance, compared to the BM3D technique [5-9].

3. Improving U-net for Image Denoising

Because deep learning has shown excellent performance in various fields of image processing, many studies on deep learning are being conducted. Research on deep learning is being conducted from various aspects, such as structures and training methods for deep neural networks. Among the various structures of deep neural networks, U-net, shown in Fig. 1, was proposed for medical image processing, but it has been used in various image processing fields, including image restoration [10, 11]. U-net improves convergence performance by adding skip connection to the autoencoder structure. The U-net encoder consists of a contractive path that extracts feature vectors from the input image, and the decoder consists of an expansive path that restores the image from the extracted feature vectors. In the deep learning process, the feature vectors extracted from the contractive path are trained so they are as close as possible to the feature vectors of the target image. For the expansive path, U-net is trained to restore the image as closely as possible to the target image using the extracted feature vectors. Image characteristics that may be lost in the process of reducing the size of the feature vectors in the contractive path are transferred to the expansive path through the skip connection, and are used in the image restoration process, thus improving convergence performance compared to the autoencoder.

In this paper, we propose improved structures for U-net, and we prove that it shows superior denoising performance, compared to the conventional deep neural networks. The improved U-net proposed in this paper can be used in various image restoration fields as well as for denoising. First, we propose the deep neural networks shown in Figs. 2 and 3. Convergence performance is enhanced by further processing the U-net input and output through pre-processing and post-processing, respectively. The input data from the pre-processing unit are transferred to the post-processing unit through an additional skip connection. After concatenation with the data processed in the expansive path, the image is restored through the final post-processing step. As shown in Fig. 2, pre-processing, the additional skip connection, and post-processing all compose a single module, and convergence performance can be further improved through cascaded connections of the modules. Also, as shown in Fig. 3, each stage of U-net can be modified by applying the so-called ResBlock structure that adds a convolution layer with a shortcut to each U-net stage. This structure can be used together with the pre-processing and post-processing structures
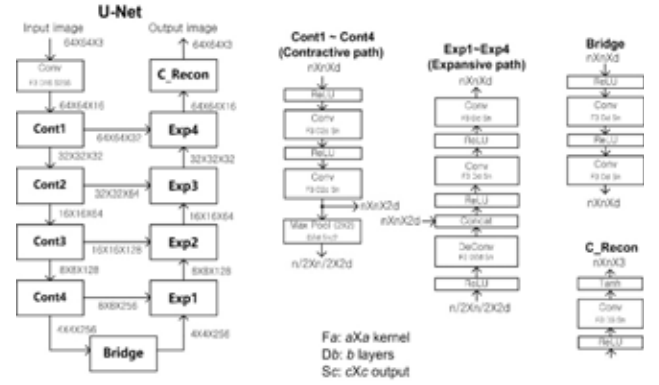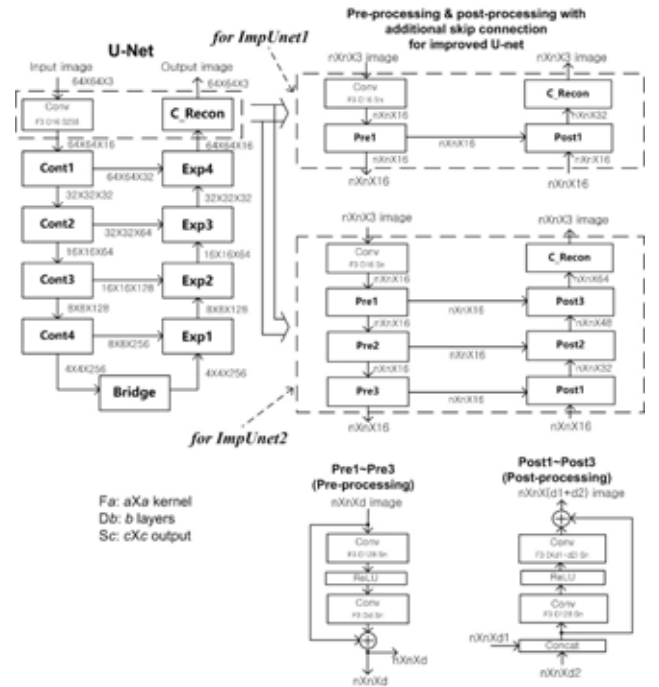

Fig. 1. U-net.

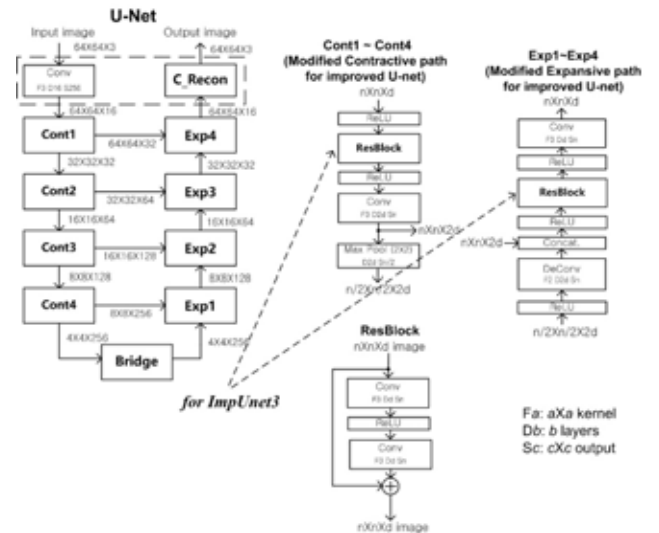**Fig. 2. Improved U-net (ImpUnet1 & ImpUnet2).**

Fig. 3. Improved U-net (ImpUnet3).

described above in order to maximize the overall performance. As is shown in Section 4, the convergence and denoising performance of the proposed structure are improved compared to the conventional U-net. Since the proposed structure can improve the overall convergence performance of a deep neural network that minimizes the difference between the target image and the degraded input image, it can be used in various image restoration fields as well as for image denoising.

4. Performance Evaluation

In order to evaluate the performance of the proposed method, extensive simulations were performed using a program based on TensorLayer [12]. Training images were generated using the DIV2K image database [13]. BSD68 images and Kodak images, which are the most widely used standard test images [14, 15], were used to measure performance. Image patches at 64×64 were extracted from the training images, and training was performed to minimize the mean square error (MSE) loss over a total of 20,000 epochs using the Adam optimizer [16]. The step size started at 10^{-4} and was decreased by 1/2 for every 4,000 epochs. Additive white Gaussian noise with standard deviation that varied between 5 and 50 was added to the input training images for the deep neural network, hence, training the deep neural network to operate regardless of the noise level.

Performance comparisons of the deep neural networks are presented in Figs. 4-9 and Tables 1 and 2, where ImpUnet1 to ImpUnet4 represent stages of the improved U-net as proposed in this paper. ImpUnet1 improves U-net by using only one pre-processing and post-processing unit, while ImpUnet2 improves U-net by using three pre-processing and post-processing units. ImpUnet3 improves U-net by using ResBlock, and ImpUnet4 improves U-net by using three pre-processing and post-processing units and ResBlock. First, to analyze the convergence performance of the deep neural network, the MSE convergence curves are presented in Fig. 4. We can see that the convergence performance of the proposed structure improves, compared to the conventional U-net. When pre-processing, post-processing, and ResBlock are used together, the convergence performance is at its best. Tables 1 and 2 show the average peak signal-to-noise ratio (PSNR) and the average structural similarity index measure (SSIM) [17] for 68 BSD68 test images and 24 Kodak images. For comparison with the proposed method, the denoising performance of the BM3D technique and deep neural networks that provide excellent performance from among the existing deep neural networks used for image denoising, was compared for various noise standard deviations, σ. We can see that the proposed deep neural network shows significant PSNR and SSIM gain, compared to BM3D and the existing deep neural networks, respectively. The proposed method outperforms the conventional U-net by up to 0.7 dB for PSNR, and shows better performance than BM3D and existing neural networks for all noise levels. As shown in Figs. 5-9, the

Table 1. Average PSNR and SSIM Results (Kodak images).

Method	PSNR (in dB)			SSIM		
	σ = 10	σ = 30	σ = 50	σ = 10	σ = 30	σ = 50
Noisy	28.21	18.85	14.78	0.6595	0.2744	0.1551
BM3D [4]	36.57	30.88	28.62	0.9435	0.8472	0.7788
DnCNN [5]	36.58	31.28	28.95	0.9447	0.8580	0.7917
IRCNN [6]	36.70	31.25	28.94	0.9448	0.8584	0.7943
FFDNet [7]	36.81	31.40	29.11	0.9462	0.8597	0.7952
ADNet [8]	36.73	31.28	28.93	0.9452	0.8576	0.7887
Unet [10]	36.19	31.29	28.98	0.9430	0.8622	0.7957
ImpUnet1	36.61	31.46	29.16	0.9461	0.8647	0.8025
ImpUnet2	36.72	31.56	29.27	0.9466	0.8677	0.8056
ImpUnet3	36.52	31.45	29.18	0.9452	0.8640	0.8027
ImpUnet4	36.88	31.63	29.30	0.9478	0.8688	0.8079

Table 2. Average PSNR and SSIM Results (BSD68 images).

Method	PSNR			SSIM		
	σ = 10	σ = 30	σ = 50	σ = 10	σ = 30	σ = 50
Noisy	28.30	19.03	14.99	0.7069	0.3299	0.1944
BM3D [4]	36.18	30.25	27.80	0.9541	0.8541	0.7776
DnCNN [5]	36.44	30.67	28.25	0.9562	0.8687	0.7987
IRCNN [6]	36.37	30.57	28.19	0.9557	0.8675	0.7985
FFDNet [7]	36.50	30.70	28.31	0.9567	0.8682	0.7984
ADNet [8]	36.38	30.56	28.13	0.9555	0.8660	0.7931
Unet [10]	35.84	30.56	28.22	0.9527	0.8690	0.8001
ImpUnet1	36.20	30.71	28.33	0.9557	0.8721	0.8050
ImpUnet2	36.30	30.75	28.39	0.9560	0.8741	0.8064
ImpUnet3	36.15	30.70	28.32	0.9549	0.8721	0.8043
ImpUnet4	36.39	30.79	28.38	0.9570	0.8749	0.8078

noise reduction performance of the proposed deep neural network is superior to that of the BM3D technique and the existing deep neural networks, and detailed characteristics of the image are restored well.

5. Conclusion

In this paper, a deep learning–based image denoising method using an improved U-net was proposed. The convergence and denoising performance of the proposed deep neural network is improved by adding pre-processing and post-processing to the conventional U-net. The performance is further enhanced by adding a convolution layer together with a shortcut in each stage of U-net. In particular, pre-processing and post-processing have a modular structure, and performance can be further improved through adopting a cascaded connection between modules. Extensive simulations confirmed that the proposed method has superior denoising performance compared to BM3D and existing deep learning methods. Since the proposed structure improves the overall
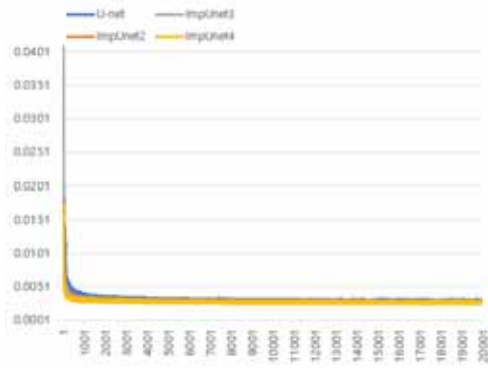
(a)

(b)

Fig. 4. MSE convergence (a) MSE for all 2,000 epochs, (b) MSE for the last 500 epochs.

(a)   (b)

(c)   (d)

Fig. 5. Test images for subjective comparison of denoising results (a) Kodak image 7, (b) BSD68 image 47, (c) Kodak image 1, (d) BSD68 image 18.

convergence performance of U-net, it can be used not only for image denoising but also for various image restoration applications.
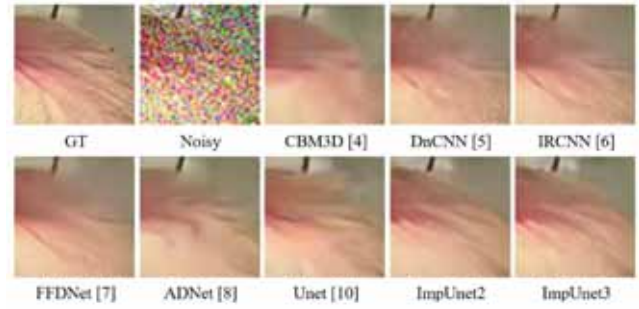

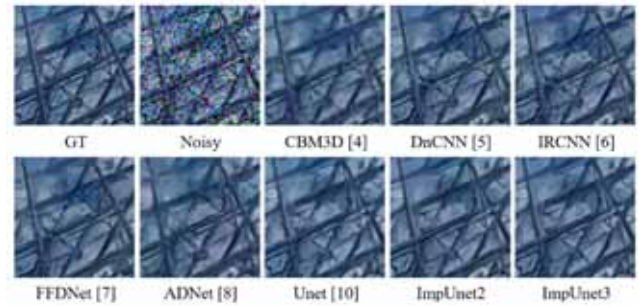**Fig. 6. Comparison of denoising results (Kodak image 7, σ=50).**

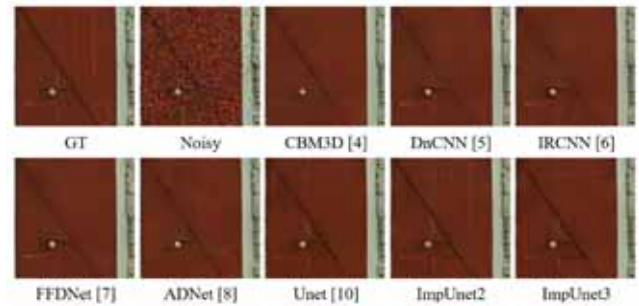**Fig. 7. Comparison of denoising results (BSD68 image 47, σ=50).**

Fig. 8. Comparison of denoising results (Kodak image 1, σ=30).

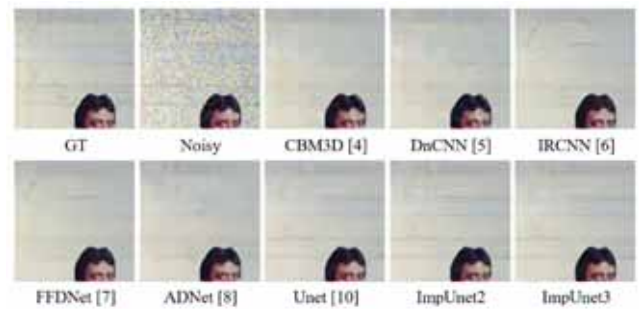
Fig. 9. Comparison of denoising results (BSD68 image 18, σ=30).

Acknowledgement

References

[1] M. Mafi, S. Tabarestani, M. Cabrerizo, A. Barreto, and M. Adjouadi, "Denoising of ultrasound images affected by combined speckle and Gaussian noise," *IET Image Processing,* vol. 12, np. 12, pp.2346–2351, 2018. Article (CrossRef Link)

[2] Y. Dong and S. Xu, "A new directional weighted median filter for removal of random-valued impulse noise," *IEEE Signal Processing Letters*, vol. 14, no. 3, pp. 193–196, 2007. Article (CrossRef Link)

[3] A. Buades, B. Coll and J.-M. Morel, "A non-local algorithm for image denoising," in *Proc. of Computer Vision and Pattern Recognition 2005 (CVPR 2005)*, pp. 60-65, June 2005. Article (CrossRef Link)

[4] K. Dabov, A. Foi and V. Katkovnik and K. Egiazarian, "Image denoising by sparse 3-D transform domain collaborative filtering," *IEEE Trans. on Image Processing*, vol. 16, no. 8, pp. 2080-2095, Aug. 2007. Article (CrossRef Link)

[5] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising," *IEEE Transactions on Image Processing*, 26(7): 3142–3155, 2017. Article (CrossRef Link)

[6] K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep CNN denoiser prior for image restoration," in *CVPR 2017*. Article (CrossRef Link)

[7] K. Zhang, W. Zuo, and L. Zhang, "FFDNet: Toward a fast and flexible solution for CNN-based image denoising," *IEEE Transactions on Image Processing*, vol. 27, no. 9, pp. 4608–4622, 2018. Article (CrossRef Link)

[8] C. Tian, Y. Xu, Z. Li, W. Zuo, L. Fei and H. Liu, "Attention-guided CNN for image denoising," *Neural networks*, vol. 124, pp. 117-129, Aprial 2020. Article (CrossRef Link)

[9] C. Tian, L. Fei, W. Zheng, Y. W. Zuo, C-W. Lin, "Deep learning on image denoising: An overview," *Neural networks*, vol. 131, pp. 251-275, Nov. 2020. Article (CrossRef Link)

[10] O. Ronneberger, P. Fischer and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," *MICCAI 2015: Medical Image Computing and Computer-Assisted Intervention 2015*, pp. 234-241, 2015. Article (CrossRef Link)

[11] Y. J. Kim and C. W. Lee, "Deep Learning Method for Extending Image Intensity Using Hybrid Log-Gamma," *IEIE Transactions on Smart Processing and Computing*, vol. 9, no. 4, pp. 312-316, August 2020. Article (CrossRef Link)

[12] H. Dong, A. Supratak, L. Mai, F. Liu, A. Oehmichen, S. Yu and Y. Guo, "TensorLayer: A versatile library for efficient deep learning development," in *Proc. ACM-MM* 2017, pp. 1201–1204, 2017. Article (CrossRef Link)

[13] E. Agustsson and R. Timofte, "NTIRE 2017 challenge on single image super-resolution: Dataset and study," in *CVPRW 2017*. Article (CrossRef Link)

[14] R. Franzen, "Kodak lossless true color image suite," source: http://r0k.us/graphics/kodak, vol. 4, 1999. Article (CrossRef Link)

[15] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *ICCV 2001*. Article (CrossRef Link)

[16] D. Kingma and J. B. Adam, "Adam: A method for stochastic optimization," *International Conference on Learning Representations*, 2015. Article (CrossRef Link)

[17] A. Horé and D. Ziou, "Image quality metrics: PSNR vs. SSIM," *20th International Conference on Pattern Recognition*, 2010. Article (CrossRef Link)

Jaewook Han is a student at the School of Information, Communications and Electronics Engineering, the Catholic University of Korea. His current interests lie in the area of image processing and deep learning.

Jinwon Choi is a student at the School of Information, Communications and Electronics Engineering, the Catholic University of Korea. His current interests lie in the area of image processing and deep learning.

Changwoo Lee received a BSc and an MSc in control and instrumentation engineering from Seoul National University. After receiving a PhD in image processing area from Seoul National University in 1996, he worked as a Senior Researcher with Samsung Electronics. He is currently a Professor at the School of Information, Communications and Electronics Engineering, the Catholic University of Korea. His current interests lie in the area of image processing and deep learning.