Densely Self-guided Wavelet Network for Image Denoising

Wei Liu[1,2] Qiong Yan[1] Yuzhi Zhao[3]
[1]SenseTime Research [2]Harbin Institute of Technology [3]City University of Hong Kong
liujikun@hit.edu.cn | yanqiong@sensetime.com | yzzhao2-c@my.cityu.edu.hk

Abstract

During the past years, deep convolutional neural networks have achieved impressive success in image denoising. In this paper, we propose a densely self-guided wavelet network (DSWN) for real world image denoising. The basic structure of DSWN is a top-down self-guidance architecture which is able to efficiently incorporate multi-scale information and extract good local features to recover clean images. Moreover, such a structure requires a smaller number of parameters and enables us to achieve better effectiveness than Unet structure. To avoid information loss and achieve a better receptive field size, we embed wavelet transform into DSWN. In addition, we apply densely residual learning to convolution blocks to enhance the feature extraction capability of the proposed network. At the full resolution level of DSWN, we adopt a double branch structure to generate the final output. One branch of them tends to pay attention to dark areas and the other performs better on bright areas. Such a double branch strategy is able to handle the noise at different exposures. The proposed network is validated by BSD68, Kodak24 and SIDD+ benchmark. Additional experimental results show that the proposed network outperforms most state-of-the-art image denoising solutions.

1. Introduction

Image denoising is a fundamental task in low-level vision and an important pre-processing step in many other vision tasks. Traditional methods [1] usually address image denoising by domain transform [2], non-local algorithm [3], Markov Random Fields (MRFs) [4], etc. However, these methods need to manually set parameters and refer a complex optimization problem for the testing stage.

With the rapid development of deep learning technology, numerous advanced approaches [5-7] have been developed and achieve impressive success. By involving the strategy of residual learning and adding batch normalization (BN) and ReLU activate function into deep architectures, DnCNN [8] is proposed to handle Gaussian blind denoising and achieves a much higher peak signal to noise ratio (PSNR) than conventional state-of-the-art approaches [9]. For the pursuit of highly accurate denoising results, some follow-up works have been proposed to remove additive white Gaussian noise (AWGN) [10, 11]. To involve multi-scale information, some most advanced end-to-end methods [12-14] apply Unet [15] as their basic structure and add some dense residual block in each level. Although these methods obtain competitive performance on benchmark datasets, their heavy computation and memory footprint hinder their application.

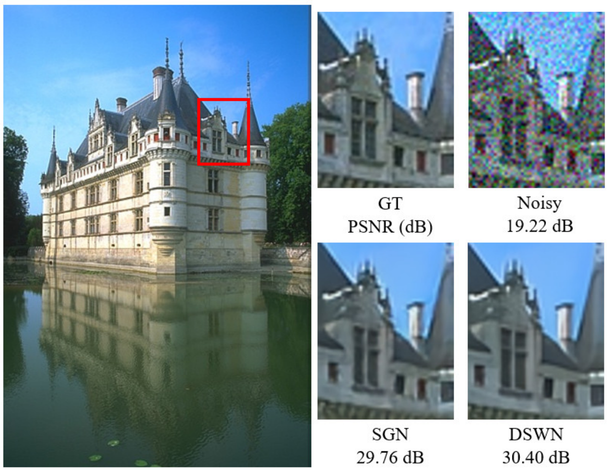


Figure 1: Example denoising results of a conventional method and the proposed method ($\sigma = 30$).

To seek a better trade-off between denoising performance and the consumption of computational resources, self-guided neural network (SGN) [16] is proposed for image denoising task by a top-down guidance strategy. SGN generates multi-resolution inputs with the PixelUnShuffle [17] before any convolutional operation. Large-scale contextual information extracted at low resolution is gradually propagated into the higher resolution sub-networks to guide the feature extraction processes at these scales. Using such a structure, SGN is able to achieve a better denoising performance than Unet with less runtime and GPU memory.

Inspired by SGN, we proposed a densely self-guided wavelet network (DSWN) (as shown in Figure 2) which is able to improve performance of SGN (Figure 1) and require less runtime than the state-of-the-art densely networks based on Unet structure [13]. In DSWN, we embed densely connected residual block (DCR) [13] (as shown in Figure 3) in each level. To achieve a better performance, we adopt more DCR blocks with skip connections at the full resolution level. We involve discrete wavelet transformation (DWT) and inverse discrete wavelet transform (IDWT) into DSWN to replace the shuffling operation in SGN. The motivation is that DWT / IDWT is able to not only avoid information loss and enlarge receptive field with better tradeoff between efficiency and restoration performance. In addition, wavelet has been applied to denoising task in traditional methods [2]. Utilizing wavelet transform to incorporate multi-scale information makes it possible for the network to have time-frequency analysis capabilities. At the full resolution level, we design a double branch structure including a residual learning branch and an end-to-end learning branch. Such a structure is able to help our network to deal with the denoising task at different exposures. Our main contributions are summarized as follows:

- We design a densely self-guided wavelet network which outperforms conventional methods and is more efficient than most state-of-the-art denoising networks with dense blocks.
- We replace PixelShuffle by wavelet transform to perform scale transformation and achieve a higher PSNR and preserve more details.
- We propose a double branch structure at the full resolution level. The residual learning branch tends to reserve more details in bright areas and end-to-end learning branch is able to supplement the dark area information and further improve denoising performance.

2. Related works

In this section, we briefly introduce some works related to our research. First, we review some deep learning based denoising networks. Then, we discuss some previous works for incorporation of multi-scale information.

2.1. Deep Neural Networks for Image Denoising

In recent years, researches have shown that deep learning technologies outperform traditional methods on image denoising by extracting more suitable image features [6]. Mao et. al. designed a convolutional encoder-decoder network with symmetric skip connections to perform image denoising [18]. By introducing a memory block to explicitly mine persistent memory through an adaptive learning process, MemNet [11] is able to learn multi-level representations of the current state under different receptive fields. By involving the strategy of residual learning and adding batch normalization and ReLU activate function into deep architectures, DnCNN [8] is proposed to handle Gaussian blend denoising. Using a tunable noise level map as the input, FFDNet [19] is able to handle a wide range of noise levels and remove spatially variant noise. Considering both Poisson-Gaussian model and in-camera processing pipeline, CBDNet [7] further improved the blind denoising ability by embedding a noise estimation network. To overcome the lack of paired training data, Chen et. al [20] simulated noise samples by a Generative Adversarial Network (GAN) and Noise2Noise [5] proposed a restoration learning strategy without any clean data. Finally, Zhang et al. [21] proposed a residual dense network (RDN) that uses both residual learning and dense connection as its basic structure, maximizing feature reuse and achieving a significant improvement in the performance of Gaussian noise image denoising.

2.2. Incorporation of multi-scale information

To extract multi-scale information for image denoising and single image super resolution (SISR) tasks, PixelShuffle and wavelet transform have been proposed to replace pooling and interpolation to avoid information loss. With self-guidance strategy and PixelShuffle, SGN [16] greatly improved the memory and runtime efficiency. Bae et al. proposed a wavelet residual network (WavResNet) [22] for image denoising and SISR and find wavelet subbands benefits learning convolutional neural network (CNN). Similarly, a deep wavelet super-resolution (DWSR) method [23] is propose to recover missing details on subbands. Both WavResNet and DWSR only consider single level wavelet decomposition. Multi-level wavelet transform is considered by MWCNN [24] to achieve better receptive field size and avoid down-sampling information loss by embedding wavelet transform into CNN architecture. MWCNN owns more power to model both spatial context and inter-subband dependency by embedding DWT and IDWT to CNN. In this paper, our proposed network adopts the same method as MWCNN to incorporate multi-scale information with a totally different architecture from MWCNN.

3. Densely Self-guided Wavelet Network (DSWN)

In this section, we firstly introduce the overall network structure and then introduce the details of DSWN.

3.1. Overall Structure of DSWN

Our proposed denoising network is shown in Figure 2. A top down self-guidance architecture is used to better exploit image multi-scale information. Information extracted at low resolution is gradually propagated into the higher resolution sub-networks to guide the feature extraction processes. Instead of PixelShuffle and PixelUnShuffle, DWT and IDWT are used to generate multi-scale inputs. Before any convolution operation, DSWN uses wavelet transform to transform the input image to three smaller

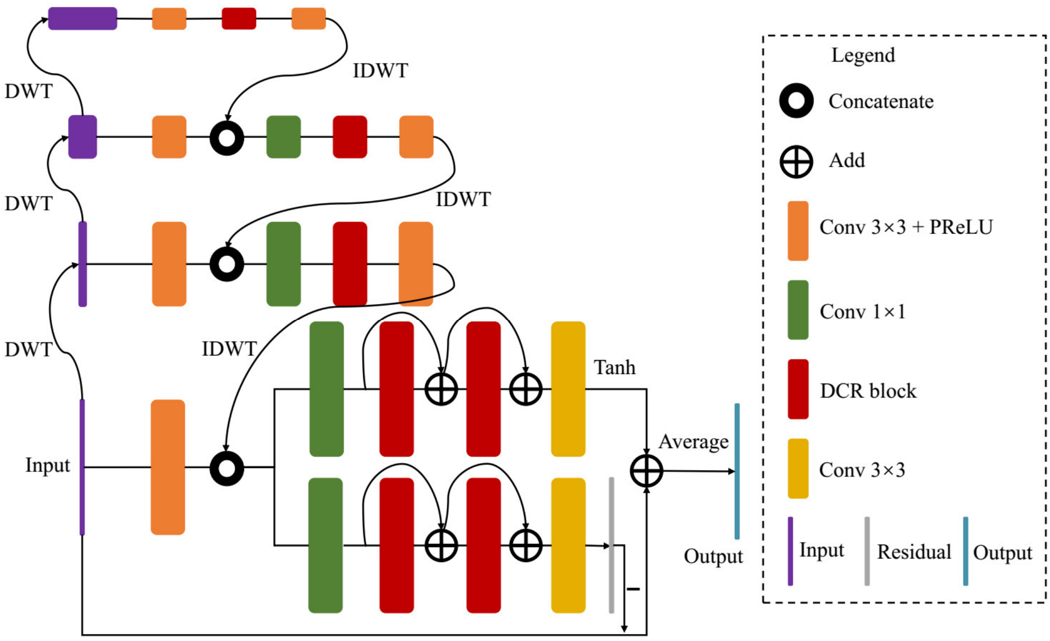


Figure 2: An illustration of our proposed network

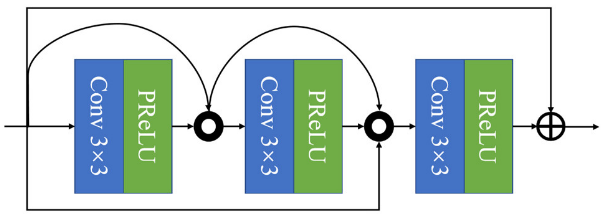


Figure 3: Diagram of densely connected residual (DCR) block.

scales. At the full resolution layer, we adopt a double branch structure consists of a residual learning branch and an end-to-end learning branch. In the rest of this paper, we simply refer to these two branches as residual branch and end2end branch. For our network, we observed that the residual branch focuses on bright areas and the end2end branch focuses on dark areas. Therefore, we use both of these two branches at the full resolution level to further improve the performance, especially when the network needs to work on noisy images with different ISO at the same time. In addition, we find batch normalization is harmful for the denoising performance and do not use any normalization layer in this network. For each level, we add densely connected residual (DCR) [13] one or two blocks as shown in Figure 3.

3.2. Detail Structure of DSWN

The top level of DSWN works on the smallest spatial resolution to extract large scale information. The top sub-network contains two Conv+PReLU layers (the orange box in Figure 2) and a DCR block (the red box in Figure 2). The DCR block simultaneously applies dense connectivity and residual learning to remove the noise of input images accurately and solve the vanishing-gradient problem.

At the middle two levels, 1×1 Convolutional kernel layers are used to merge information extracted from different resolution. The network structure of the middle sub-networks is similar to the structure of the top sub-network. As for the full resolution level, we add more DCR blocks with skip connections to enhance the feature extraction capability of DSWN after merging information from all the scales. For the residual branch, DSWN has a global residual connection between the input image and the final estimation. We add a Tanh activation function at the end of the end2end branch. The final output is the simple average result of the residual branch and the end2end branch. By adding gradient loss, our network is able to achieve better retention of details without reducing PSNR. In order to ensure the fairness of the comparative experiment, in the experimental part of this paper, our network only uses L1 loss for training.

4. Experimental Results

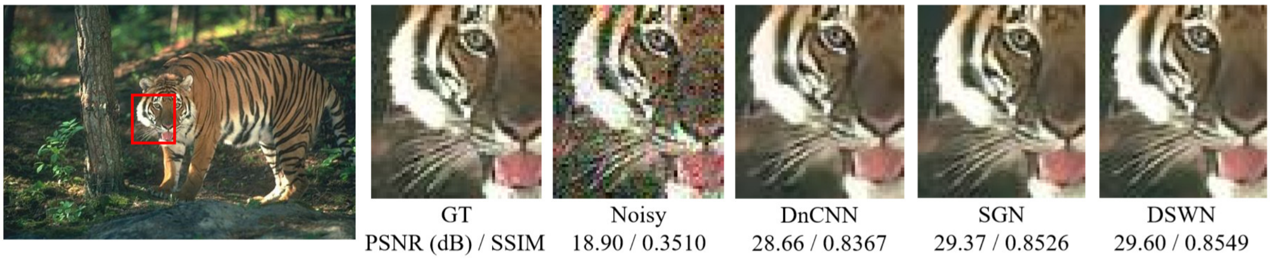
In this section, we first introduce the training details and then provide experimental results on different datasets. We compare DSWN with several state-of-the-art denoising approaches.

4.1. Experimental Setting

DIV2K [25] training and validation datasets provide an adequate amount of high quality images. Most state-of-the-art image denoising solutions select the DIV2K as their training dataset [12, 16]. To train our DSWN model, we

GT PSNR (dB) / SSIM | Noisy 15.13 / 0.2800 | DnCNN 23.57 / 0.6204 | SGN 24.99 / 0.6710 | DSWN 25.67 / 0.6653

Figure 4: Denoising results of the conventional methods and the proposed method on BSD68 dataset (σ = 50).



GT PSNR (dB) / SSIM | Noisy 18.90 / 0.3510 | DnCNN 28.66 / 0.8367 | SGN 29.37 / 0.8526 | DSWN 29.60 / 0.8549

Figure 5: Denoising results of the conventional methods and the proposed method on BSD68 dataset (σ = 30).

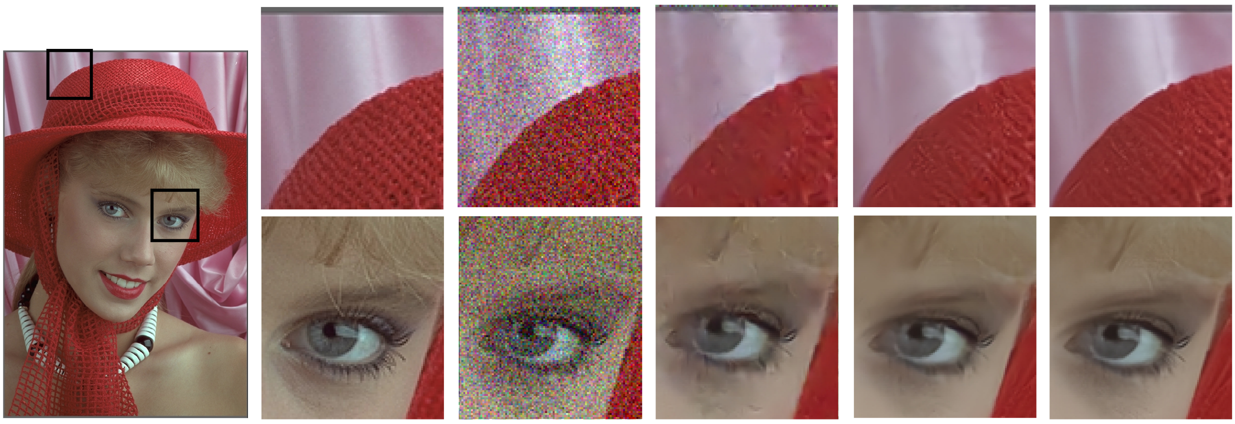
Table 1: RGB image denoising results on BSD68 dataset.

Method	Noisy	CBM3D [26]	DnCNN [8]	FFDNet [19]	IRCNN [27]	SGN [16]	DHDN [13]	DSWN
Noise Level	PSNR (dB)							
σ = 10	28.30	35.89	36.12	36.14	36.06	35.97	36.45	**36.91**
σ = 30	19.03	29.72	30.32	30.31	30.22	30.36	30.41	**30.72**
σ = 50	14.91	27.36	27.92	26.96	27.86	28.01	28.02	**28.29**
Noise Level	SSIM							
σ = 10	0.7114	0.9507	0.9536	0.9540	0.9533	0.9544	0.9572	**0.9578**
σ = 30	0.3363	0.8432	0.8611	0.8603	0.8607	0.8684	0.8639	**0.8751**
σ = 50	0..1993	0.7622	0.7882	0.7881	0.7889	0.8011	0.7961	**0.8081**

also use DIV2K dataset as training and validation dataset for AWGN task. The training dataset consists of 800 high quality images. The resolution of each of these images is 1920 × 1080. The DIV2K validation dataset consists of 100 images; the quality of each image is similar to that of the training dataset. For the testing datasets, we use the BSD68 [28] and dataset Kodak dataset [29] which are used by some recent denoising networks [30]. The Kodak dataset consists of 24 images, each of which has a resolution of 768×512. The BSD68 dataset consists of 68 images, each of which has a resolution of 321 × 481. As for real noise removal task, we conduct experiments on Smartphone Image Denoising Dataset (SIDD) benchmark [31]. The SIDD training dataset consists of 320 images from 10 scenes under different lighting conditions using five representative smartphone cameras. The test dataset is SIDD+ dataset [32] which is generated by NTIRE 2020 Real Image Denoising Challenge with a similar procedure as the SIDD benchmark. Compared with the SIDD validation dataset, the SIDD+ dataset contains more details.

Table 2: RGB image denoising results on Kodak24 dataset.

Method	Noisy	CBM3D [26]	DnCNN [8]	FFDNet [19]	IRCNN [27]	SGN [16]	DHDN [13]	DSWN
Noise Level	PSNR (dB)							
σ = 10	28.24	36.57	36.58	36.80	36.70	36.64	37.33	**37.35**
σ = 30	18.93	30.89	31.28	31.39	31.24	31.70	31.95	**33.08**
σ = 50	14.87	28.62	28.94	29.10	28.92	29.42	29.67	**29.97**
Noise Level	SSIM							
σ = 10	0.6607	0.9432	0.9446	0.9462	0.9448	0.9455	**0.9508**	0.9488
σ = 30	0.2755	0.8459	0.8579	0.8596	0.8581	0.8599	0.8736	**0.8794**
σ = 50	0.1557	0.7772	0.7915	0.7949	0.7939	0.7949	0.8160	**0.8241**



GT PSNR (dB) / SSIM — Noisy 15.13 / 0.2800 — DnCNN 30.72 / 0.8010 — SGN 32.27 / 0.8283 — DSWN 32.74 / 0.8366

Figure 6: Denoising results of the conventional methods and the proposed method on Kodak24 dataset ( ).

When training our model, we randomly crop 256 × 256 patches from the training images. The input patches of the proposed network are randomly flipped and rotated for data augmentation. The parameters of network are Xavier initialized [33]. We train the whole network for 300 epochs overall. The learning rate is initialized as 1e−4 at the first 200 epochs and reduce to 5e−5 in the next 50 epochs. We finetune our model at the last 50 epochs with a 1e−5 learning rate. For optimization, we use Adam optimizer [34] with $\beta 1 = 0.5$, $\beta 2 = 0.999$ and batch size equals to 1. We use L1 loss which is a PSNR-oriented optimization in the training process [35]. The experiments are implemented using a NVIDIA RTX 2080Ti GPU.

4.2. Performance comparison

We compare our proposed network with several state-of-the-art image denoising solutions: CBM3D [9], DnCNN [8], FFDNet [19], IRCNN [27], SGN [16] and DHDN [13], where DHDN is more complex than DSWN. To compare the performance, we determined the peak-signal-noise-ratio (PSNR) [36] and structural similarity (SSIM) [37] as the objective measurements. Table 1 lists the average PSNR of the compared methods and the proposed method for sRGB images in BSD68 dataset. Figure 4 and Figure 5 show some example results in BSD68 dataset with different noise levels. From Table 1, we can see that our DSWN shows the best PSNR and SSIM in all the noise levels. DHDN is the second best method. It adopts twice as many DCR blocks as DSWN and is harder to train. Although DHDN adopts more DCR block in its network architecture, DSWN is still able to achieve a PSNR which is about 1.37 higher than DHDN on average. Figure 4 and Figure 5 show some detail results of DnCNN, SGN and DSWN, where DnCNN is a classical denoising network and SGN has a similar structure to DSWN. We can see all the three denoising networks are able to achieve a obvious improvement compared with noisy images. DSWN is better than DnCNN and SGN in some details such as the texture details of the statue in Figure 4 and the beard of the tiger in Figure 5. Our proposed method is able to handle different noise levels and reserve more details at the same time.

	GT	Noisy	DnCNN	SGN	DSWN
PSNR (dB) / SSIM		14.85 / 0.3284	23.86 / 0.7313	26.26 / 0.7905	26.98 / 0.8161

Figure 7: Denoising results of the conventional methods and the proposed method on Kodak24 dataset ($\sigma = 50$).

GT Noisy DSWN

Figure 8: NTIRE 2020 real image denoising challenge results of DSWN on validation dataset of SIDD+ with unspecified noise.

Table 2 shows DSWN outperforms the compared methods in most cases. From Figure 6 and Figure 7, we can conclude a similar conclusion to the BSD68 dataset. At a higher noise level, denoising networks tend to smooth the noise image too much, because the network is difficult to distinguish true details from noise. DSWN can better preserve details at high noise levels, such as eyelashes and window textures.

4.3. NTIRE 2020 real image denoising challenge

The proposed method is initially proposed to participate in NTIRE 2020 real image denoising challenge [32]. The purpose of the challenge is to remove unspecified noise from images. The training dataset is SIDD benchmark which includes images with various noise levels, dynamic ranges, and brightness. Such a noise model is more complex than the AWGN task. During the challenge, we tried several different solutions on SIDD benchmark dataset and finally fix DSWN as our final solution. Our network achieves a competitive PSNR on the testing dataset and is very efficient among all the solutions for this challenge.

In this section, we first present denoising results generated by DSWN and some compared denoising networks. Then, we conduct some ablation studies on different network structures. In the section, we evaluate the proposed method on SIDD+ validation dataset which provided by the organizers of NTIRE 2020 real image denoising challenge.

4.3.1 Performance comparison

We compare our proposed network with several state-of-the-art image denoising solutions state-of-the-art image denoising solutions: DnCNN [8], RED [10], MemNet [11], ResUnet [38], MWCNN [24] and SGN [16]. Table 3 shows the denoising performance of different methods. For real world denoising task, DSWN is able to improve the PSNR from 26.63 to 36.94 and improve the SSIM by 0.2952. As shown in Figure 8, our proposed network successfully removes unspecified noise from noisy images. DSWN not only perform well under proper exposure, but also shows excellent performance in underexposed scenes. For some scenes with strong noise and many details (such as text on a globe), DSWN is able to denoise without losing the sharp edges of the text.

Table 3: Denoising results on SIDD+ validation dataset.

Method	Noisy	DnCNN [8]	ResUnet [38]	RED [10]	MWCNN [24]	MemNet [11]	SGN [16]	DSWN
PSNR (dB)	26.63	34.03	34.01	27.35	35.87	32.28	35.51	**36.94**
SSIM	0.6622	0.9139	0.9192	0.7331	0.9407	0.8669	0.9411	**0.9574**

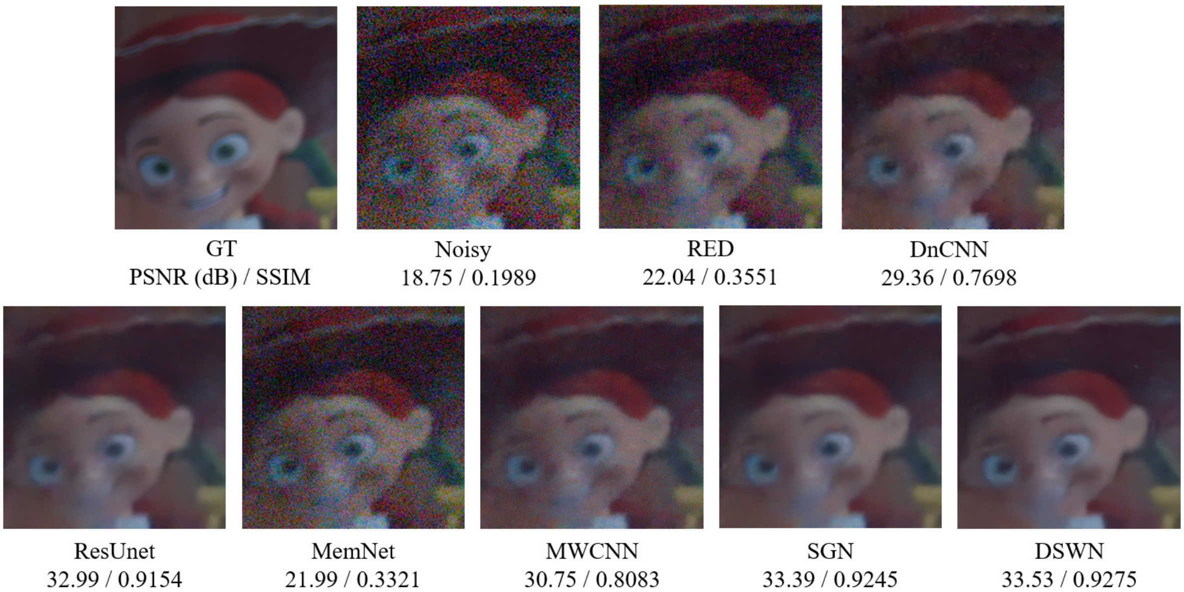


Figure 9: Denoising results of the conventional methods and DSWN on SIDD+ validation dataset.

Figure 9 shows example results of the compared methods and DSWN. The denoised result of DSWN is better than other methods, especially at the edges of the hat, hair and eyebrows. Owing to the real images denoising task is more difficult than AWGN task, some of the compared methods only slightly improved the performance.

In this challenge, we average eight output images of eight input images; these are generated by a combination of a flip and rotation of an input image. Using such an ensemble strategy, the PSNR of our method is able to be further improved by 0.1.

4.3.2 Ablation Study

Table 4: Ablation study on two branches of DSWN on SIDD+ validation dataset.

Method	Residual branch only	End2End branch only	Two Residual branches	DSWN
PSNR (dB)	36.67	36.44	36.86	**36.94**
SSIM	0.9515	0.9466	0.9533	**0.9574**

At the full resolution level of DSWN, we adopt a double branch structure: a residual branch and an end2end branch. At first, we tried such a scheme as an ensemble strategy. However, we observe an interesting phenomenon: the residual branch tends to ignore the dark areas and the end2end branch tends to highlight the dark details. As a result, these two branches collaborate to generate output images with appropriate colors as shown in Figure 10. To prove the effectiveness of the double branch structure, we conduct additional experiments as shown in Table 4. We can see DSWN is able to achieve better PSNR and SSIM than using any branch only. If we replace the end2end branch by another residual branch, our DSWN is still better. This demonstrates that the improvement is not only from the increase of network parameters. In addition, the results from the same two residual branches are similar (Figure 11).

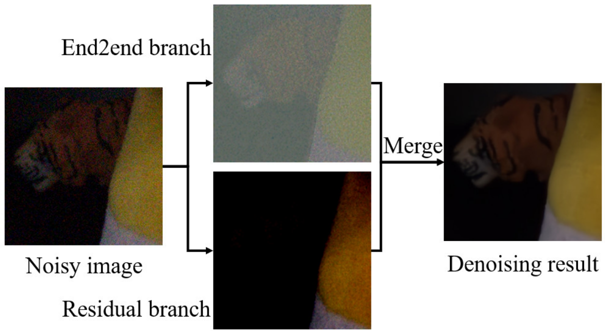


Figure 10: Schematic diagram of double branch learning of at the full resolution level (DSWN).

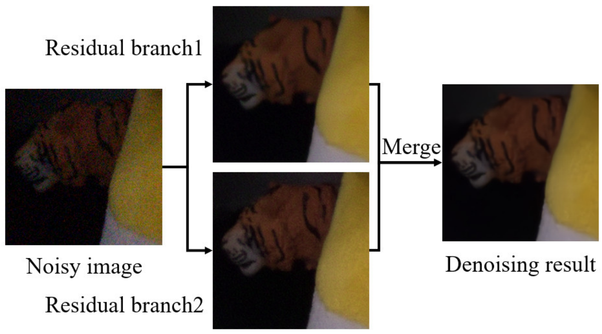


Figure 10: Schematic diagram of double branch learning of at the full resolution level (Two Residual branches).

In addition, our proposed the main structure of DSWN is inspired by SGN. The most obvious structural difference between the two is the incorporation of multi-scale information. DSWN adopt DWT and IDWT to replace PixelShuffle and PixelUnShuffle in SGN. Both of them are no loss of information. In Table 5, we show the ablation study of these two methods. The results suggest that wavelet transform is better than PixelShuffle strategy. This might be because the frequency domain is more suitable for denoising and wavelet transform is able to enlarge receptive field.

Table 5: Ablation study on down / up sampling method of DSWN on SIDD+ validation dataset.

Method	PixelShuffle / PixelUnShuffle	DWT/IDWT
PSNR (dB)	36.59	**36.94**
SSIM	0.9548	**0.9571**

5. Conclusion

In this paper, we proposed a densely self-guided wavelet network (DSWN) for image denoising. DSWN adopts a top-down manner to denoise images. Wavelet transform is adopted to generate input variations with different spatial resolutions before any convolutional operation. Then, we embed a DCR block into three low spatial resolution levels, respectively. At the full resolution level, we employ more DCR blocks and a double branch structure to further improve the quality of output images. The proposed DSWN was validated on AWGN task and real world image denoising benchmark and demonstrated excellent efficiency in NTIRE 2020 real image denoising challenge. DSWN is able to generate higher quality denoising results than the compared state-of-the-art methods.

References

[1] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Transactions on image processing,* vol. 16, no. 8, pp. 2080-2095, 2007.

[2] S. Sardy, P. Tseng, and A. Bruce, "Robust wavelet denoising," *IEEE Transactions on Signal Processing,* vol. 49, no. 6, pp. 1146-1152, 2001.

[3] A. Buades, B. Coll, and J. M. Morel, "A non-local algorithm for image denoising," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, 2005.

[4] X. Lan, S. Roth, D. Huttenlocher, and M. J. Black, "Efficient belief propagation with learned higher-order markov random fields," in *European conference on computer vision*, 2006, pp. 269-282: Springer.

[5] J. Lehtinen *et al.*, "Noise2noise: Learning image restoration without clean data," *arXiv preprint arXiv:1803.04189,* 2018.

[6] C. Tian, Y. Xu, L. Fei, and K. Yan, "Deep Learning for Image Denoising: A Survey."

[7] G. Shi, Z. Yan, Z. Kai, W. Zuo, and Z. Lei, "Toward Convolutional Blind Denoising of Real Photographs."

[8] 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*.

[9] K. Dabov, A. Foi, V. Katkovnik, and K. O. Egiazarian, "Image restoration by sparse 3D transform-domain collaborative filtering," in *Image Processing: Algorithms and Systems VI, 28 January 2008, San Jose, California, USA*, 2008.

[10] X. Mao, C. Shen, and Y.-B. Yang, "Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections," in *Advances in neural information processing systems*, 2016, pp. 2802-2810.

[11] Y. Tai, J. Yang, X. Liu, and C. Xu, "Memnet: A persistent memory network for image restoration," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 4539-4547.

[12] S. Yu, B. Park, and J. Jeong, "Deep iterative down-up cnn for image denoising," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2019, pp. 0-0.

[13] B. Park, S. Yu, and J. Jeong, "Densely connected hierarchical network for image denoising," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2019, pp. 0-0.

[14] A. Abdelhamed, R. Timofte, and M. S. Brown, "Ntire 2019 challenge on real image denoising: Methods and results," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2019, pp. 0-0.

[15] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, 2015, pp. 234-241: Springer.

[16] S. Gu, Y. Li, L. V. Gool, and R. Timofte, "Self-Guided Network for Fast Image Denoising," in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, pp. 2511-2520.

[17] W. Shi *et al.*, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 1874-1883.

[18] X. J. Mao, C. Shen, and Y. B. Yang, "Image Restoration Using Convolutional Auto-encoders with Symmetric Skip Connections."

[19] K. Zhang, W. Zuo, and L. Zhang, "FFDNet: Toward a Fast and Flexible Solution for CNN based Image Denoising," *IEEE Transactions on Image Processing,* pp. 1-1.

[20] J. Chen, J. Chen, H. Chao, and M. Yang, "Image blind denoising with generative adversarial network based noise modeling," in *Proceedings of the IEEE*

Conference on Computer Vision and Pattern Recognition, 2018, pp. 3155-3164.
[21] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual dense network for image restoration," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 2020.
[22] W. Bae, J. Yoo, and J. Chul Ye, "Beyond deep residual learning for image restoration: Persistent homology-guided manifold simplification," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2017, pp. 145-153.
[23] T. Guo, H. Seyed Mousavi, T. Huu Vu, and V. Monga, "Deep wavelet prediction for image super-resolution," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2017, pp. 104-113.
[24] P. Liu, H. Zhang, W. Lian, and W. Zuo, "Multi-level Wavelet Convolutional Neural Networks."
[25] R. Timofte, E. Agustsson, L. V. Gool, M. H. Yang, and Q. Guo, "NTIRE 2017 Challenge on Single Image Super-Resolution: Methods and Results," in *IEEE Conference on Computer Vision & Pattern Recognition Workshops*, 2017.
[26] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Color image denoising via sparse 3D collaborative filtering with grouping constraint in luminance-chrominance space," in *2007 IEEE International Conference on Image Processing*, 2007, vol. 1, pp. I-313-I-316: IEEE.
[27] K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep CNN denoiser prior for image restoration," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 3929-3938.
[28] 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 *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, 2001, vol. 2, pp. 416-423: IEEE.
[29] R. Franzen, "Kodak lossless true color image suite. source: http://r0k.us/graphics/kodak," vol. 4, 1999.
[30] S. Lefkimmiatis, "Universal denoising networks: a novel CNN architecture for image denoising," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 3204-3213.
[31] A. Abdelhamed, S. Lin, and M. S. Brown, "A high-quality denoising dataset for smartphone cameras," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 1692-1700.
[32] A. a. A. Abdelhamed, Mahmoud and Timofte, Radu et. al., "NTIRE 2020 Challenge on Real Image Denoising: Dataset, Methods and Results," *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops,* 2020.
[33] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, 2010, pp. 249-256.
[34] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980,* 2014.
[35] H. Zhao, O. Gallo, I. Frosio, and J. Kautz, "Loss Functions for Image Restoration With Neural Networks," *IEEE Transactions on Computational Imaging,* vol. 3, no. 1, pp. 47-57.
[36] A. Horé and D. Ziou, "Image Quality Metrics: PSNR vs. SSIM," 2010.
[37] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing,* vol. 13, no. 4, pp. 600-612, 2004.
[38] C. Chen, Q. Chen, M. N. Do, and V. Koltun, "Seeing motion in the dark," in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, pp. 3185-3194.