



# Fast, nonlocal and neural: a lightweight high quality solution to image denoising

Yu Guo, Axel Davy, Gabriele Facciolo, Jean-Michel Morel, Qiyu Jin

## ► To cite this version:

Yu Guo, Axel Davy, Gabriele Facciolo, Jean-Michel Morel, Qiyu Jin. Fast, nonlocal and neural: a lightweight high quality solution to image denoising. IEEE Signal Processing Letters, 2021, 28, pp.1515-1519. 10.1109/LSP.2021.3099963 . hal-03521188

**HAL Id: hal-03521188**

**<https://hal.science/hal-03521188>**

Submitted on 11 Jan 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Fast, nonlocal and neural: a lightweight high quality solution to image denoising

Yu Guo, *Graduate Student Member, IEEE*, Axel Davy, Gabriele Facciolo,  
Jean-Michel Morel and Qiyu Jin, *Member, IEEE*

**Abstract**—With the widespread application of convolutional neural networks (CNNs), the traditional model based denoising algorithms are now outperformed. However, CNNs face two problems. First, they are computationally demanding, which makes their deployment especially difficult for mobile terminals. Second, experimental evidence shows that CNNs often over-smooth regular textures present in images, in contrast to traditional non-local models. In this letter, we propose a solution to both issues by combining a nonlocal algorithm with a lightweight residual CNN. This solution gives full latitude to the advantages of both models. We apply this framework to two GPU implementations of classic nonlocal algorithms (NLM and BM3D) and observe a substantial gain in both cases, performing better than the state-of-the-art with low computational requirements. Our solution is between 10 and 20 times faster than CNNs with equivalent performance and attains higher PSNR. In addition the final method shows a notable gain on images containing complex textures like the ones of the *MIT Moiré* dataset.

**Index Terms**—BM3D, convolutional neural network, image denoising, nonlocal methods

## I. INTRODUCTION

Denoising is one of the most critical issues in image processing. Indeed, noise lowers the final image quality and impacts all downstream computer vision tasks. After a variance stabilizing transform (VST) the raw observed image can be expressed as

$$y = x + n, \quad (1)$$

where  $y$  is the observed image,  $x$  is the underlying clean image, and  $n$  is an additive approximately white Gaussian noise (AWGN). In the past few decades and before the emergence of CNNs, several image prior models have been proposed, particularly the non-local models (NLM [1], NLBayes [2], BM3D [3], WNNM [4]) and sparse models (LSSC [5]). These algorithms restore well textures and repeated structures, but they also leave behind residual noise and cause blur.

This work was supported by National Natural Science Foundation of China (No. 12061052), Natural Science Fund of Inner Mongolia Autonomous Region (No. 2020MS01002), Prof. Guoqing Chen’s “111 project” of higher education talent training in Inner Mongolia Autonomous Region and the network information center of Inner Mongolia University. Work partly financed by Office of Naval research grants N00014-17-1-2552 and N00014-20-S-B001, DGA Defals challenge n° ANR-16-DEFA-0004-01, MENRT and Fondation Mathématique Jacques Hadamard. (*Corresponding author: Qiyu Jin.*)

Y. Guo and Q. Jin are with School of Mathematical Science, Inner Mongolia University, Hohhot 010020, China (e-mail: yuguomath@aliyun.com; qyjin2015@aliyun.com).

A. Davy, G. Facciolo and J-M. Morel are with Centre Borelli, ENS Paris-Saclay, CNRS, 4, avenue des Sciences 91190 Gif-sur-Yvette, France (e-mail: axel.davy@normalesup.org; gabriele.facciolo@ens-paris-saclay.fr; morel@ens-paris-saclay.fr).

The non-local models are now clearly outperformed by data-driven CNNs such as DnCNN [6], FFDNet [7] and others [8]–[11]. Yet these data-driven algorithms are energy greedy and require high-performance GPUs. This hinders their deployment on a large scale, especially for mobile terminals such as mobile phones. Furthermore, although CNNs show excellent PSNR performance, it has been observed that they over-smooth regular textures and self-similar structure [12].

Aware of the scale and speed of CNNs, Gu et al. [13] proposed a self-guided fast network using multi-resolution input. Ma et al. [14] combined a pyramid neural network with the Two-Pathway Unscented Kalman Filter and proposed a fast algorithm for real-world noise removal. Aiming at raw image denoising of mobile devices, Wang et al. [15] proposed a lightweight and effective neural network and k-Sigma noise transformation method.

The second problem has led many researchers to incorporate non-local routines into their CNN design. Wang [16] proposed a non-convolutional denoising lightweight network trained on texture patches. The texture patches are detected, denoised by the network and then aggregated to the BM3D result. Yang et al. [17] referred to the BM3D algorithm and proposed “extraction” and “aggregation” layers to model the block matching in BM3D, and used CNNs to replace the collaborative filtering stage. Cruz et al. [12] placed plug-and-play non-local filters after CNNs, thus adding a nonlocal regularity term. In [18] the authors, inspired by non-local neural networks [19], introduced non-local CNNs into image restoration. Lefkimmiatis [20] performed block matching and weighted non-local sum on the results of 2D Convolution, thereby effectively integrating non-locality into CNNs. Taking advantage of the relaxation of the k-nearest neighbors matching selection rule, Plötz and Roth [21] proposed a novel non-local layer. All of the above mentioned papers gave evidence that a non-local routine indeed improves the receptive field of CNNs and enhances their performance. Unfortunately, given the high dimensionality of their feature space, non-local operations on CNNs further increase their power consumption.

Both nonlocal algorithms and deep learning based ones have their own advantages and disadvantages, and they complement each other’s shortcomings. In this letter, a network is proposed to combine a nonlocal algorithm and a deep learning based method to benefit from the advantages of both techniques. In the proposed network, the output of a nonlocal method is used as a pre-processing algorithm for it protects the image details while reducing noise, then a network is used to learn the residual of the restored image of the nonlocal method.

TABLE I  
CPSNR(DB) COMPARISON BETWEEN OUR PROPOSED METHOD AND THE  
STATE-OF-THE-ART METHODS WITH THE FIXED PARAMETER.

Dataset	$\sigma$	BM3D [22]	BM3D -G [23]	DnCNN [6]	ADNet [8]	DnCNN $K = 10$	Ours $K = 10$	Ours $K = 16$
<i>Kodak</i> [24]	25	31.88	31.72	32.34	32.30	32.08	32.38	<b>32.40</b>
	35	30.29	30.14	30.75	30.75	30.47	30.82	<b>30.86</b>
	50	28.68	28.55	29.15	29.18	28.85	29.26	<b>29.32</b>
	75	26.98	26.88	—	27.51	27.08	27.56	<b>27.64</b>
<i>McM</i> [25]	25	31.72	31.65	32.49	32.47	32.23	32.52	<b>32.55</b>
	35	30.18	30.11	30.93	30.93	30.60	30.96	<b>31.02</b>
	50	28.53	28.48	29.23	29.29	28.92	29.33	<b>29.43</b>
	75	26.76	26.72	—	27.47	27.06	27.48	<b>27.59</b>
<i>CBSD68</i> [26]	25	30.76	30.62	<b>31.25</b>	31.19	31.06	31.23	<b>31.25</b>
	35	29.09	28.94	29.59	29.56	29.37	29.59	<b>29.62</b>
	50	27.48	27.35	28.01	27.95	27.73	27.97	<b>28.02</b>
	75	25.69	25.60	—	26.24	25.95	26.23	<b>26.30</b>
<i>MIT</i> [27]	25	30.41	30.39	29.92	29.97	29.63	30.59	<b>30.64</b>
	35	28.82	28.80	28.34	28.41	28.04	28.99	<b>29.07</b>
	50	27.07	27.06	26.76	26.86	26.47	27.38	<b>27.46</b>
	75	25.34	25.36	—	25.22	24.80	25.61	<b>25.69</b>
<i>Urban</i> [28]	25	30.97	30.75	31.17	31.23	30.86	31.64	<b>31.68</b>
	35	29.25	29.05	29.37	29.42	28.00	29.89	<b>29.96</b>
	50	27.43	27.28	27.48	27.57	27.12	28.09	<b>28.17</b>
	75	25.32	25.26	—	25.50	25.02	25.98	<b>26.07</b>

Here, the network corrects the restored image given by the nonlocal method, that is, removes artifacts, color distortions and residual noise from the image, rather than full denoising. As a result, a very lightweight network can produce excellent results.

The rest of the paper is organized as follows. In Section II, we verify the effectiveness of the proposed scheme. We find that preprocessing by BM3D leads to a drastic reduction of the size of the CNN yielding an optimal solution. These experiments made with fixed noise CNNs lead us to propose a lightweight and flexible color image denoising method in Section III. This algorithm obtains high-quality results at a lower computational cost for images with variable noise. Experiments evaluating image quality and computational cost are presented in Section IV, where our algorithm is compared with state-of-the-art denoising algorithms.

## II. COMBINATION OF BM3D AND CNN

We take inspiration in residual learning, which has been used in other areas of image restoration such as demosaicing [29] and super-resolution [30]. We test a processing pipeline where a fast, nonlocal and neural filter first removes noise by a GPU-accelerated BM3D algorithm (BM3D-G) [23]. Then we feed both the restored image and noisy image to the CNN to learn a residual, finally obtaining the restored image by adding the residual to the BM3D preprocessed image. This specific model is shown in Fig. 1.

Using the BM3D-G preprocessing step reduces floating point operations and lightens the network. With reference to the simple and effective design of DnCNN [6],  $3 \times 3$  convolutions are used to build the network. To test how far BM3D could allow for a reduction of the scale of the network, we test a network with 10 convolutional layers ( $K = 10$ , 44.6% of the DnCNN scale) and a network with 16 convolutional layers ( $K = 16$ , 77.1% of the DnCNN scale).

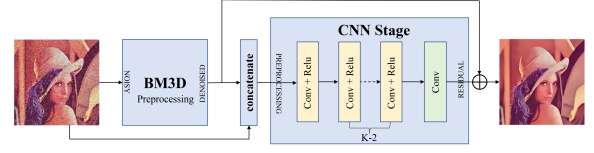


Fig. 1. Algorithm structure. Using residual learning to combine BM3D and CNNs.

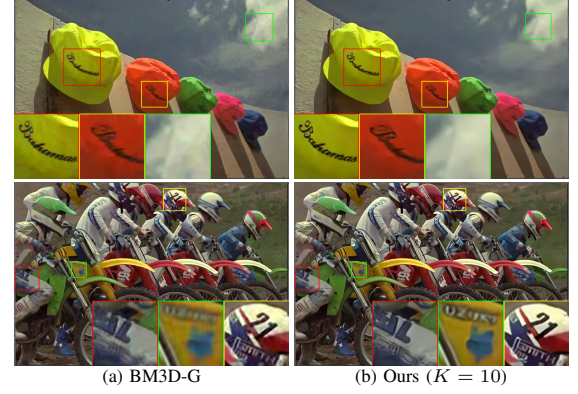


Fig. 2. Result of BM3D before and after being post-processed by a lightweight CNN.

The training hyperparameters for our network are the same as [6]. The training dataset is the Waterloo Exploration Database (WED) [31], which contains 4744 natural images. We crop the images into 200,000  $128 \times 128$  patches and used an  $L_1$  loss defined by

$$\mathcal{L}(\Theta) = \frac{1}{N} \sum_{i=1}^N |x_i - [F(z_i; \Theta) + BM3D(y_i)]|, \quad (2)$$

$$z_i = \text{Concatenate}(BM3D(y_i), y_i), \quad (3)$$

where  $y_i$  is a noisy image and  $x_i$  is the corresponding clean image. We considered five datasets to comprehensively test the performance of the algorithm. There are three natural image sets (Kodak [24], MacMaster [25], CBSD68 [26]) and two texture image sets (MIT moiré [27], Urban100 [28]). Our comparison criterion is the composite PSNR (CPSNR) [32].

In Table I, we can see that with only  $K = 10$  convolutional layers, our proposed method reaches the performance of DnCNN while having half of its size. With  $K = 16$  convolutional layers, our proposed scheme outperforms DnCNN. On the texture image sets, our method shows a significant improvement. On the MIT moiré dataset, the average gain of our proposed method is as high as 0.60dB, and on Urban100, the CPSNR improves by nearly 0.54dB.

These experiments show that using BM3D to preprocess noisy images has two advantages:

- After the BM3D preprocessing, the difficulty of solving the inverse problem is reduced. This explains why the network can be shallower without performance loss.
- The use of BM3D preprocessing improves the non-local performance of the network, with a just small increase in cost. The combination greatly improves network effectiveness in processing texture images or self-similar structures.

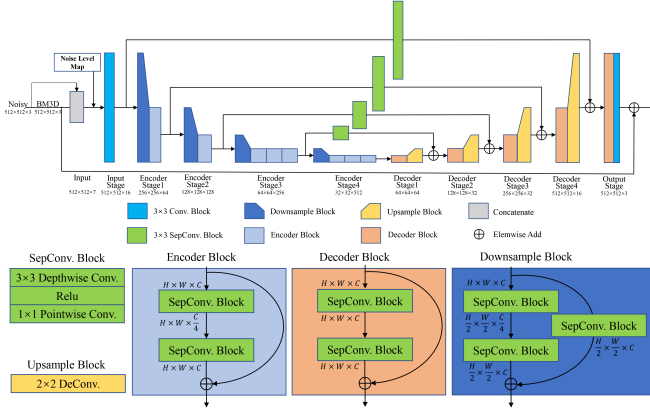


Fig. 3. Flexible and lightweight U-Net-like structure of denoising.

Although BM3D has remarkable performance, it introduces unpleasant artifacts, that become conspicuous at higher noise levels. Fig. 2 (a) illustrates that there are ripples, color distortion and blur in the BM3D restored images. The CNN effectively corrects these artifacts and provides pleasant restored images (see Fig. 2 (b)).

In summary, the combination of BM3D and CNN not only overcomes each other's shortcomings but also improves the quality of denoised images. Yet these experiments were conducted with a fixed noise model, requiring training for each noise level. This is not practical for mobile devices. In the next section we address this problem.

### III. A FLEXIBLE AND LIGHTWEIGHT DENOISING ALGORITHM

CNNs trained for a fixed noise level perform well but cannot be practically applied in real scenes due to huge memory requirements. In order to be widely used in mobile terminals (such as mobile phones), an algorithm has to satisfy two conditions:

- **Flexibility:** The algorithm must be able to deal with various levels of noise flexibly, instead of training for each level of noise like most CNNs, as done in [7].
- **Lightweight:** Most devices do not have high-performance GPUs. So the algorithm must be lightweight to ensure high-speed operation in a low-power environment.

Considering the above two characteristics and the conclusion of Sec. II, we designed a flexible and lightweight U-Net-like denoising algorithm. The overall network structure is shown in Fig. 3. We refer to the design of FFDNet [7] to introduce a noise level map to deal flexibly with noise. We use a U-Net architecture [33] with 4 encoders and 4 decoders. In order to preserve the low-dimensional characteristics of the network, skip connections are added. The entire network uses  $3 \times 3$  depthwise separable convolutions [34] to reduce computational consumption. The stride-2 depthwise separate convolution is used for downsampling, and the corresponding upsampling uses a  $2 \times 2$  deconvolution. The details of the network block are shown in Fig. 3. See Sec. II for the relevant settings of network training and testing.

TABLE II  
CPSNR(DB) VALUE COMPARISON BETWEEN OUR PROPOSED METHOD AND THE STATE-OF-THE-ART METHODS WITH THE FLEXIBLE PARAMETER.

Dataset	$\sigma$	NLM -G [23]	BM3D -G [23]	FDnCNN [6]	FFDNet [7]	PMRID [15]	Ours-FL NLM	Ours-FL BM3D
Kodak	25	30.07	31.72	32.24	32.25	32.01	32.11	<b>32.26</b>
	35	28.72	30.14	30.68	30.69	30.53	30.65	<b>30.77</b>
	50	27.20	28.55	29.07	29.10	29.04	29.16	<b>29.25</b>
	75	25.47	26.88	27.39	27.43	27.48	27.56	<b>27.65</b>
McM	25	29.79	31.65	<b>32.39</b>	32.36	31.74	31.91	32.22
	35	28.42	30.11	<b>30.84</b>	30.83	30.36	30.53	30.79
	50	26.86	28.48	29.18	27.29	28.90	29.05	<b>29.24</b>
	75	25.04	26.72	27.36	27.37	27.28	27.38	<b>27.50</b>
CBSD68	25	28.83	30.62	31.13	<b>31.14</b>	30.98	31.01	31.13
	35	27.56	28.94	29.51	29.51	29.40	29.45	<b>29.54</b>
	50	26.06	27.35	27.95	27.95	27.85	27.91	<b>27.98</b>
	75	24.36	25.59	26.16	26.17	26.21	26.26	<b>26.32</b>
MIT	25	28.49	30.39	29.74	29.73	29.16	29.91	<b>30.53</b>
	35	27.16	28.80	28.21	28.23	27.80	28.44	<b>29.02</b>
	50	25.58	27.06	26.69	26.72	26.44	26.94	<b>27.46</b>
	75	23.73	25.36	25.04	25.08	25.01	25.30	<b>25.76</b>
Urban	25	28.75	30.75	30.99	30.87	29.94	30.71	<b>31.22</b>
	35	27.41	29.05	29.22	29.14	28.31	29.12	<b>29.60</b>
	50	25.71	27.28	27.41	27.36	26.64	27.37	<b>27.88</b>
	75	23.57	25.26	25.27	25.26	24.80	25.35	<b>25.89</b>
SIDD	25	36.09	37.70	38.15	38.34	38.51	38.55	<b>38.66</b>
	35	34.49	36.20	36.76	36.99	37.35	37.37	<b>37.48</b>
	50	32.65	34.44	35.19	35.44	36.04	36.03	<b>36.11</b>
	75	30.40	32.30	33.30	33.63	34.41	34.40	<b>34.43</b>

## IV. EXPERIMENTS AND ANALYSIS

### A. Quantitative and Qualitative Comparison

We conducted comprehensive testing on five image sets in order to fully evaluate the performance of the algorithm. Three flexible deep learning algorithms (flexible DnCNN (FDnCNN) [6], FFDNet [7], PMRID [15]) were used for comparison. PMRID is a lightweight raw image denoising algorithm for mobile devices, that we retrained in order to compare it with our algorithm.

Table II shows that our proposed method outperforms other flexible CNNs algorithms, especially for MIT moiré and Urban100, which are rich in details. On the MIT moiré image set our proposed method beats other CNNs methods by 0.75dB. Nowadays, users and manufacturers prefer 4K resolution images. Therefore, we selected 25 normal light images taken by Google Pixel ( $4048 \times 3044$ ) from the SIDD dataset [35] for testing. For these high resolution images, our proposed schemes maintain their advance.

A comparison of Table I and Table II shows that the performance of FDnCNN has decreased compared with DnCNN. Similarly, Ours-FL (BM3D) has a slight decrease in CPSNR compared to Ours ( $K = 16$ ). However, in the following evaluation of the cost of the algorithm, we find that these drops are worthwhile. Indeed, it is more difficult to train with flexible parameters than fixed parameters, and the search for local optimal solutions consumes more resources. The comparison with [15] also shows that while direct application of lightweight networks can reduce the computational cost, its performance is lower in textured areas. Therefore, using BM3D for preprocessing can effectively reduce the difficulty and improve performance.



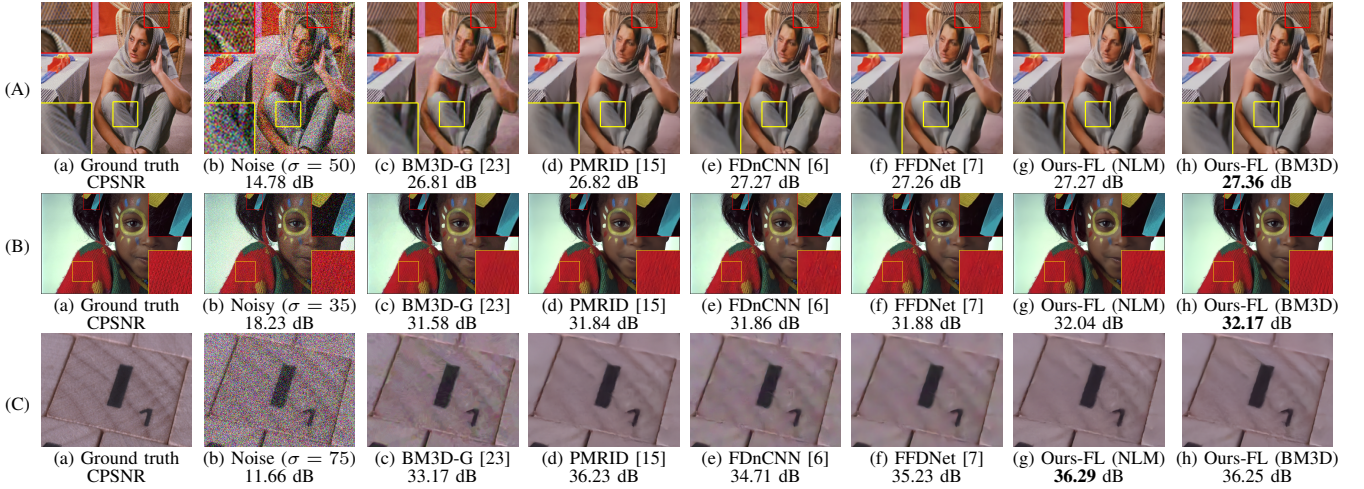


Fig. 4. Barbara, image 15 of *Kodak* ( $768 \times 512$ ) and image 156 of *SIDD* ( $4048 \times 3044$ ) processed by different methods.

TABLE III  
COMPARISON OF NETWORK COMPLEXITY.

Image size		NLM -G [23]	BM3D -G [23]	DnCNN [6]	FDnCNN [6]	DnCNN $K = 10$	FFDNet [7]	PMRID [15]	ADNet [8]	Ours $K = 10$	Ours-FL NLM/BM3D
$512 \times 512$	GFLOPs	—	—	175.78	175.32	78.48	55.97	4.71	136.71	78.82	4.52
	Memory (MB)	65	71	904	841	773	747	785	905	839	802
	Time-GPU (s)	0.03	0.09	0.17	0.15	0.07	0.05	0.03	0.13	0.07(+0.09)	0.03(+0.03/0.09)
	Time-CPU (s)	—	—	42.58	41.31	18.55	3.43	1.86	27.65	18.57(+0.09)	1.82(+0.03/0.09)
$1600 \times 1200$	GFLOPs	—	—	1287.42	1284.10	574.83	409.93	34.47	1001.27	577.30	33.12
	Memory (MB)	137	184	3015	2579	2069	1421	1823	3011	2605	1867
	Time-GPU (s)	0.16	0.66	1.21	1.05	0.50	0.35	0.21	1.13	0.50(+0.66)	0.21(+0.16/0.66)
	Time-CPU (s)	—	—	87.02	82.98	39.20	21.40	12.44	73.03	39.29(+0.66)	9.19(+0.16/0.66)
Model size (MB)		0.13	0.15	2.67	2.64	1.20	3.38	4.11	2.09	1.19(+0.15)	3.87(+0.13/0.15)

Fig. 4 compare the visual quality of the considered algorithms. In Fig. 4 (A), the texture of the chair is messy and the shadows on the pants are over-smoothed in the image restored by FDNcNN and FFDNet, while the texture of the chair is perfectly preserved and the shadows are preserved by our method. The same conclusion can be drawn from Fig. 4 (B) and (C).

From these results we conclude that our proposed method is good at preserving details and texture compared to CNNs algorithms. It also creates fewer artifacts than non-local methods. In fact, the non-local part of our proposed framework can be replaced with any non-local denoising algorithm, as long as its processing speed can be guaranteed. Based on this, we also tested the use of NLM-G [23] as preprocessing algorithm. Compared to BM3D, NLM is computationally less expensive and therefore widely used in cameras. Ours-FL (NLM) yields significantly better results than NLM. On high resolution images, the NLM scheme is far less complex than the BM3D scheme (see Table III). Yet, its PSNR and visual quality is only slightly inferior to that of the BM3D scheme (see Table II and Fig. 4 (C)).

### B. Network Complexity and Running Time

The network complexity and running time were compared on a PC with Intel Core i7-9750H 2.60GHz, 16GB memory, and Nvidia GTX-1650 GPU. We tested images of two sizes:

$512 \times 512$  (the most commonly used benchmark) and  $1600 \times 1200$  (two megapixels, the maximum size that DnCNN can run on a 4G video memory device). Four indicators: Floating point operations (FLOPs), Memory (MB), Time-GPU (s) and Time-CPU (s) are used for comparison.

The experimental results are shown in Table III. In terms of computation, our model is only 8.08% of FFDNet and 2.58% of FDNcNN. The significant reduction in the amount of calculation makes it possible to deploy on mobile terminals. In terms of run time, our algorithm retains the advantages on both CPU and GPU. As the resolution increases, the time of BM3D begins to increase, but NLM+lightweight CNN provides a valuable alternative, with an excellent trade-off between computational cost and visual quality.

For a higher cost solution with optimal image quality, BM3D remains the best choice. Since BM3D has already been successfully implemented in several smart phones, a BM3D+lightweight CNN solution is feasible.

## V. CONCLUSION

In this letter, we explored a way to combine the advantages of non-local denoising with those of denoising CNNs. We saw that non-locality recovers image texture, and that combining a nonlocal algorithm with a lightweight CNN corrects the image artifacts. The syncretic algorithm requires much less computational power than current CNNs and can therefore be used in mobile terminals.

## REFERENCES

- [1] A. Buades, B. Coll, and J.-M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 490–530, 2005.
- [2] M. Lebrun, A. Buades, and J.-M. Morel, "Implementation of the "non-local bayes" (nl-bayes) image denoising algorithm," *Image Processing On Line*, vol. 3, pp. 1–42, 2013.
- [3] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, 2007.
- [4] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm minimization with application to image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 2862–2869.
- [5] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2009, pp. 2272–2279.
- [6] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, 2017.
- [7] K. Zhang, W. Zuo, and L. Zhang, "Ffdnet: Toward a fast and flexible solution for cnn-based image denoising," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4608–4622, 2018.
- [8] C. Tian, Y. Xu, Z. Li, W. Zuo, L. Fei, and H. Liu, "Attention-guided cnn for image denoising," *Neural Netw.*, vol. 124, pp. 117–129, 2020.
- [9] Y. I. Jang, Y. Kim, and N. I. Cho, "Dual path denoising network for real photographic noise," *IEEE Signal Process. Lett.*, vol. 27, pp. 860–864, 2020.
- [10] Y. Song, Y. Zhu, and X. Du, "Grouped multi-scale network for real-world image denoising," *IEEE Signal Process. Lett.*, vol. 27, pp. 2124–2128, 2020.
- [11] Y. Wang, X. Song, and K. Chen, "Channel and space attention neural network for image denoising," *IEEE Signal Process. Lett.*, vol. 28, pp. 424–428, 2021.
- [12] C. Cruz, A. Foi, V. Katkovnik, and K. Egiazarian, "Nonlocality-reinforced convolutional neural networks for image denoising," *IEEE Signal Process. Lett.*, vol. 25, no. 8, pp. 1216–1220, 2018.
- [13] S. Gu, Y. Li, L. Van Gool, and R. Timofte, "Self-guided network for fast image denoising," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2019, pp. 2511–2520.
- [14] R. Ma, H. Hu, S. Xing, and Z. Li, "Efficient and fast real-world noisy image denoising by combining pyramid neural network and two-pathway unscented kalman filter," *IEEE Trans. Image Process.*, vol. 29, pp. 3927–3940, 2020.
- [15] Y. Wang, H. Huang, Q. Xu, J. Liu, Y. Liu, and J. Wang, "Practical deep raw image denoising on mobile devices," in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 1–16.
- [16] Y.-Q. Wang, "Small neural networks can denoise image textures well: a useful complement to bm3d," *Image Processing On Line*, vol. 6, pp. 1–7, 2016.
- [17] D. Yang and J. Sun, "Bm3d-net: A convolutional neural network for transform-domain collaborative filtering," *IEEE Signal Process. Lett.*, vol. 25, no. 1, pp. 55–59, 2018.
- [18] D. Liu, B. Wen, Y. Fan, C. C. Loy, and T. S. Huang, "Non-local recurrent network for image restoration," in *Proc. Neural Inf. Process. Syst.*, 2018, p. 1680–1689.
- [19] X. Wang, R. Girshick, A. Gupta, and K. He, "Non-local neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 7794–7803.
- [20] S. Lefkimmatis, "Non-local color image denoising with convolutional neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 5882–5891.
- [21] T. Plötz and S. Roth, "Neural nearest neighbors networks," in *Proc. Neural Inf. Process. Syst.*, 2018, p. 1095–1106.
- [22] 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 *Proc. IEEE Int. Conf. Image Process.*, vol. 1, 2007, pp. I – 313–I – 316.
- [23] A. Davy and T. Ehret, "Gpu acceleration of nl-means, bm3d and vbm3d," *J. Real-Time Image Process.*, vol. 18, p. 57–74, 02 2021.
- [24] L. Zhang, X. Wu, A. Buades, and X. Li, "Color demosaicking by local directional interpolation and nonlocal adaptive thresholding," *J. Electron. Imaging*, vol. 20, no. 2, p. 023016, 2011.
- [25] E. Dubois, "Frequency-domain methods for demosaicking of bayer-sampled color images," *IEEE Signal Process. Lett.*, vol. 12, no. 12, pp. 847–850, 2005.
- [26] S. Roth and M. J. Black, "Fields of experts," *Int. J. Comput. Vision*, vol. 82, no. 2, pp. 205–229, 2009.
- [27] M. Gharbi, G. Chaurasia, S. Paris, and F. Durand, "Deep joint demosaicking and denoising," *ACM Trans. Graph.*, vol. 35, no. 6, p. 191, 2016.
- [28] J. Huang, A. Singh, and N. Ahuja, "Single image super-resolution from transformed self-exemplars," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 5197–5206.
- [29] R. Tan, K. Zhang, W. Zuo, and L. Zhang, "Color image demosaicking via deep residual learning," in *Proc. IEEE Int. Conf. Multimedia Expo*, 2017, pp. 793–798.
- [30] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 184–199.
- [31] K. Ma, Z. Duanmu, Q. Wu, Z. Wang, H. Yong, H. Li, and L. Zhang, "Waterloo exploration database: New challenges for image quality assessment models," *IEEE Trans. Image Process.*, vol. 26, no. 2, pp. 1004–1016, 2017.
- [32] D. Alleysson, S. Susstrunk, and J. Herault, "Linear demosaicing inspired by the human visual system," *IEEE Trans. Image Process.*, vol. 14, no. 4, pp. 439–449, 2005.
- [33] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2015, pp. 234–241.
- [34] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv:1704.04861*, 2017.
- [35] A. Abdelhamed, S. Lin, and M. S. Brown, "A high-quality denoising dataset for smartphone cameras," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 1692–1700.