

# Enhanced CNN for image denoising

ISSN 2468-2322

Received on 8th November 2018

Revised on 27th December 2018

Accepted on 7th January 2019

doi: 10.1049/trit.2018.1054

www.ietdl.org

Chunwei Tian<sup>1,2</sup>, Yong Xu<sup>1,2</sup> ✉, Lunke Fei<sup>3</sup>, Junqian Wang<sup>1,2</sup>, Jie Wen<sup>1,2</sup>, Nan Luo<sup>4</sup><sup>1</sup>Bio-Computing Research Center, Harbin Institute of Technology, Shenzhen, Shenzhen 518055, People's Republic of China<sup>2</sup>Shenzhen Medical Biometrics Perception and Analysis Engineering Laboratory, Harbin Institute of Technology, Shenzhen, Shenzhen 518055, People's Republic of China<sup>3</sup>School of Computer Science and Technology, Guangdong University of Technology, Guangzhou 510006, People's Republic of China<sup>4</sup>Institute of Automation Heilongjiang Academy of Sciences, Harbin 150090, People's Republic of China

✉ E-mail: yongxu@yml.com

**Abstract:** Owing to the flexible architectures of deep convolutional neural networks (CNNs) are successfully used for image denoising. However, they suffer from the following drawbacks: (i) deep network architecture is very difficult to train. (ii) Deeper networks face the challenge of performance saturation. In this study, the authors propose a novel method called enhanced convolutional neural denoising network (ECNDNet). Specifically, they use residual learning and batch normalisation techniques to address the problem of training difficulties and accelerate the convergence of the network. In addition, dilated convolutions are used in the proposed network to enlarge the context information and reduce the computational cost. Extensive experiments demonstrate that the ECNDNet outperforms the state-of-the-art methods for image denoising.

## 1 Introduction

Image denoising is a classical technique of image restoration and has been successful in many fields such as pathological analysis and human entertainment [1, 2]. The degradation model is widely used in denoising problem to recover clear image, which is expressed as  $y = x + \mu$ , where  $x$  is a clean image,  $y$  is a noisy image and  $\mu$  is the additive Gaussian noise with standard deviation  $\sigma$ . According to the Bayesian theory, it is known that the prior is very important for image denoising [3]. For example, wavelet transformation with a prior of Markov random field is used to suppress noise [4]. Combing the self-similarities and sparse representation can improve the performance and reduce the storage for image denoising [5]. Block-matching and 3D filtering (BM3D) converts 2D image data into 3D data arrays and uses the sparse method to deal with the obtained 3D data arrays to remove noise [6]. Enforcing the gradient histogram of the noisy image is approximate to the theoretical gradient histogram of the clean image for image denoising [7]. In addition, Nonlocally centralised sparse representation (NCSR) [8], gradient methods [9, 10], total-variation methods [11, 12] and weight nuclear norm minimisation (WNNM) [13] are also very effective for image denoising.

Although the above methods have obtained great performance for denoising task, they still face the following problems [3]: (i) they need to set manually the parameters to obtain the optimal results. (ii) They use complex optimisation to improve the performance, which increases the computational cost.

Owing to the flexible connection fashion of the deep network architecture and strong learning ability, deep learning techniques have become the most effective methods to address the above problems for image denoising. Specifically, deep convolutional neural networks (CNNs) have attracted more attention in image denoising [14]. For example, CNN uses residual learning method to improve the performance in image denoising [3]. It first uses a model to deal with multiple restoration tasks such as image denoising, image super-resolution and image deblocking. The fusion of CNN and characteristics of denoising task is useful to remove unknown [15]. Combining CNN and nature of images is very effective to obtain a clean image. For example, CNN utilises non-local similarity to deal with colour noisy images [16].

Discriminative learning methods embedded into optimisation method obtain great performance for real noisy images [17]. CNN consolidated unsupervised learning is a good choice for image restoration [18]. Using the principle of enhanced signal-to-design novel network architecture is also very popular to recover image [19]. Integrating spatial domain into CNN can better filter noise [20]. The combination of traditional denoising methods and CNN such as BMCNN is very competent to separate noise from noisy image [21]. The fusion of multiple features is very beneficial for image denoising [22]. Deep CNN has good visual effects on multiplicative noises [23]. Deep CNN is a good tool for medical image denoising [24, 25]. The recently proposed deep cascade convolutional residual denoising network (DCCRDN) repeatedly uses concatenate operations to train the models for image denoising [26]. Although the above deep network methods have obtained great performance for denoising tasks, most of these methods suffer from the drawbacks of vanishing or exploding gradients when the network architecture is very deep. In addition, the above methods sacrifice the computational cost to improve the performance. For example, they apply multiple concatenation operations to train the denoising model.

In this paper, we propose a novel network referred to as enhanced convolutional neural denoising network (ECNDNet). ECNDNet utilises residual learning technique [27] to prevent vanishing and exploding gradient problems. Moreover, batch normalisation (BN) [28] is used to accelerate the convergence of the trained model and make the network easy to train. To decrease the computational burden, we use dilated convolution [29] to capture more context information. Extensive experiments demonstrate that our proposed ECNDNet method outperforms the popular image denoising methods such as fast and flexible denoising net (FFDNet) [15], image restoration CNN (IRCNN) [17] and BM3D [6].

The main contributions of this work are summarised as follows:

- (i) The depth of the proposed ECNDNet is only set to 17 layers, which can effectively reduce the computational cost.
- (ii) ECNDNet uses residual learning mechanism to prevent vanishing and exploding gradient problems. Besides, it utilises BN technique to normalise data and improve the efficiency of the training model.

(iii) ECNDNet uses dilated convolutions to enlarge the receptive field and improve the performance.

The remaining of this paper is organised as follows. Section 2 presents related work of the proposed method. Section 3 provides the proposed method. Section 4 shows the extensive experimental results of this paper. Section 5 offers the conclusion.

## 2 Related work

### 2.1 BN and residual learning

One of the reasons for CNN’s success is its end-to-end connection. The end-to-end connection architecture of CNN generally includes initial parameter [30], gradient optimisation methods [31, 32] and rectified linear unit (ReLU) [33]. Although the general network architecture has obtained good performance, they face vanishing/exploding gradient problems and have difficulty in training deep networks. In this paper, we use BN and residual learning to address the above problems. The detailed information about BN [27] and residual learning [28] are explained as follows: the distribution of sample data is changed after it passes the convolution layer. This phenomenon is called internal covariate shift problem. This problem can be addressed by BN technique. That is, first, BN normalises the training data in every batch. Then, it uses scale and shift operations to recover the distribution of training data. The above two important parameters of BN are updated when the trained network is back propagation. BN is set before the activation function of each layer. BN enjoys the following merits: (i) it can accelerate the convergence of the training model and makes the network easier to train. (ii) It makes the different batches of training data keep uniform distribution and

improves the performance of the network. (iii) It has low sensitivity for initialisation.

To the best of our knowledge, although increasing the depth of network can improve the performance for image denoising, deeper network may lead to the vanishing or exploding gradient problems. Residual learning is a good tool to solve this problem. It mainly adds the input (original images) and residual block (the output of several feature layers) as the input of the current layer to guarantee the performance. As shown in Fig. 1, we assume that  $x$  and  $f(x)$  represent the input and the output of stack several layers, respectively. The input of the next layer of the stack several layers is  $f(x) + x$ .

### 2.2 Dilated convolution

As we know, more features can improve the performance for image processing [34–36]. Enlarging the receptive field in the CNN is very effective for extracting more features for image denoising [37]. There are two popular ways to enlarge the receptive field: (i) enlarging the width of the network (also referred to as increasing the filter size). (ii) Increasing the depth of the network. However, the first way may produce more parameters, which results in over fitting of the network. It also increases the computational cost. The second way may lead to vanishing/exploding gradients when the depth of the network is big. As a consequence, dilated convolution is a good choice to balance the above ways. Dilated convolution uses a dilated filter with dilation factor  $f$  to increase the obtained information. That is, a dilated filter can be expressed as a filter with size  $(2f + 1)(2f + 1)$ . For example, when is 1, the receptive field of the first layer is 3. The receptive fields of the other layers are 5, 7, 9, ..., respectively. In addition, combining the dilated filter and the convolutional kernel of  $3 \times 3$  is very popular for image processing [29]. For more details on dilated convolution, please refer to [29].

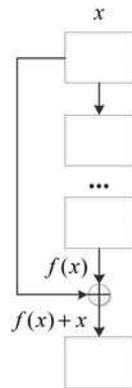


Fig. 1 Idea of residual learning mechanism

## 3 Proposed denoising method

### 3.1 Network architecture

According to the previous research, we know that the denoising method can be expressed as  $y = x + \mu$ . In this paper, the objective function of learning  $f(y)$  is as follows:

$$l(p) = \frac{1}{N} \sum_{j=1}^N \|f(y_j; p) - (y_j - x_j)\|^2 \quad (1)$$

Formula (1) is the objective function to train the denoising model, where  $p$  represents the parameters,  $y_j$  represents the  $j$ th noisy image patch and  $x_j$  represents the  $j$ th label image patch. Specifically, the image patches can reduce the computational cost

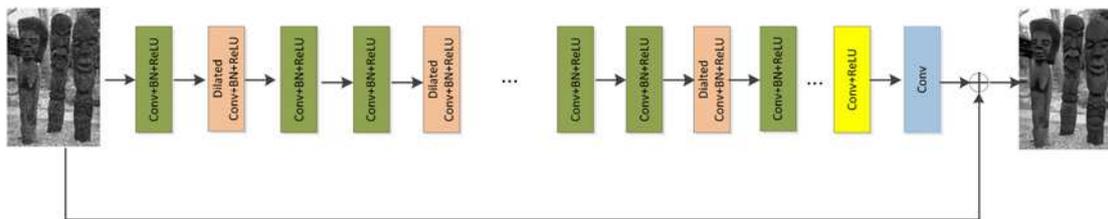


Fig. 2 Architecture of ECNDNet

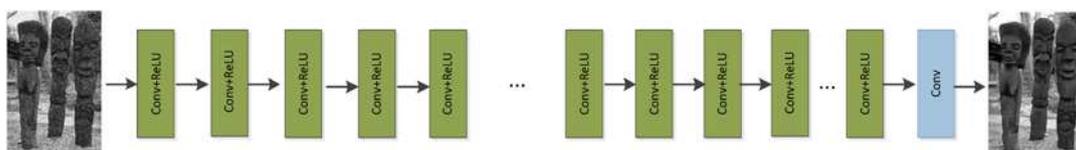


Fig. 3 Architecture of CRNet

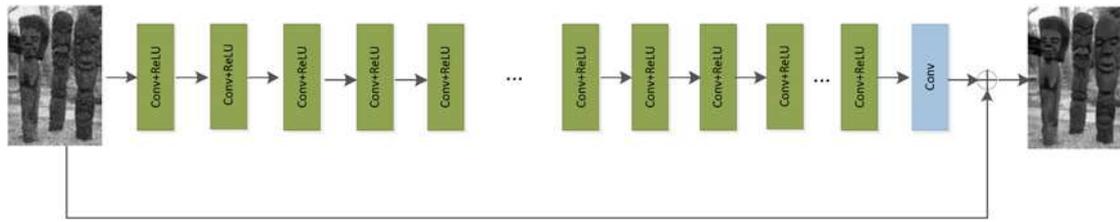


Fig. 4 Architecture of CRRNet

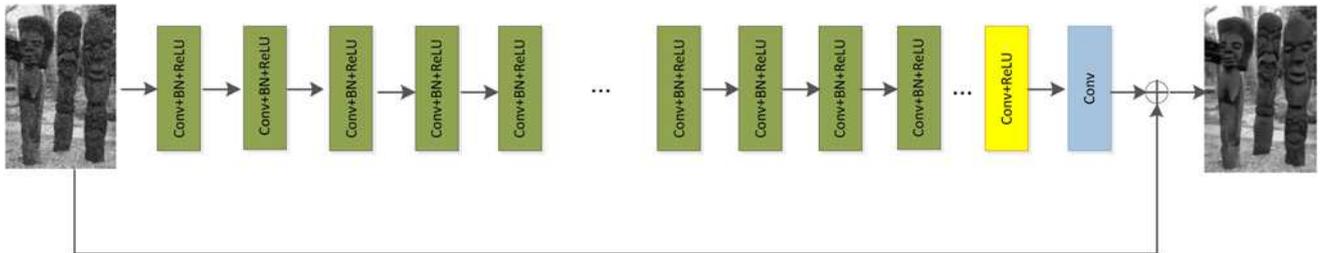


Fig. 5 Architecture of CRRBNet

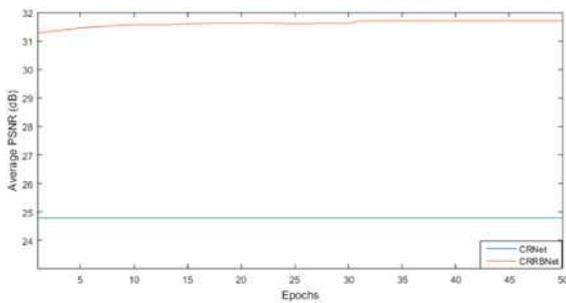


Fig. 6 Gaussian denoising results of CRNet and CRRBNet on BSD68 is shown. CRNet only has convolution and ReLU. CRRBNet includes BN and ReLU and residual learning. They are trained with  $\sigma = 15$

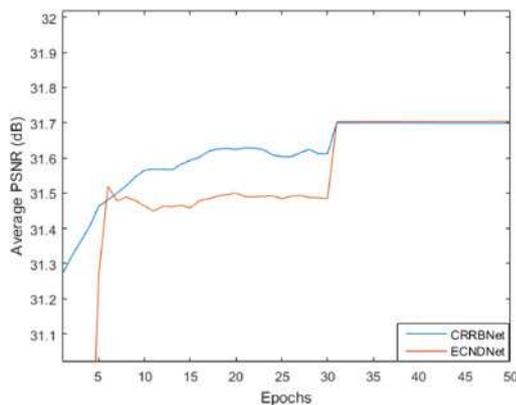


Fig. 7 Gaussian denoising results of CRRBNet and ECNDNet on BSD68 is shown. CRRBNet has BN and ReLU and residual learning. ECNDNet includes BN, ReLU, residual learning and dilated convolution. They are trained with  $\sigma = 15$

and learn more features [38]. Thus, we divided the image into patches is reasonable for image denoising. In addition, very deep architecture is another non-ignorable factor which can result in vanishing or exploding gradient problems. As a result of these concerns, we proposed a novel network called ECNDNet. ECNDNet consists of dilated convolution, residual learning, BN, convolution (Conv) and ReLU. We empirically find that sets the dilated convolution to the 2nd, 5th, 9th and 12th layers can not only increase the captured information, but also reduces the computational cost than that of each layer with dilated convolution. Moreover, the use of BN and residual learning makes this network more effective for image denoising. The architecture of the designed network is shown in Fig. 2. Also the depth of the proposed network is 17. It has four types in this network: Conv, ReLU, BN and dilated Conv. Specifically, they are convolution, rectified linear units, BN and dilated convolution, respectively. The 1st and 16th layers are Conv+ReLU. The 2nd, 5th, 9th and 12th layers are dilated Conv+BN+ReLU. Specifically, the dilated factor is important to enlarge the receptive field for dilated convolution. Here we use dilated factor of 2 and the receptive fields of all 17 are 3, 7, 9, 11, 15, 17, 19, 21, 25, 27, 29, 33, 35, 37, 39, 41 and 43, respectively. It can map the context features from  $3 \times 3$  to  $43 \times 43$ . The final layer is Conv. The other layers are Conv+BN+ReLU. The size of the convolutional kernels is  $128 \times 1 \times 40 \times 40$  for the first and the last layers, respectively. The size of other convolutional kernels is  $128 \times 64 \times 40 \times 40$ .

The merits of the proposed method have three-fold: (i) it uses 17 layers network and residual learning to prevent the problems of vanishing or exploding gradients. (ii) It uses BN technique to accelerate convergence and make the network easier to train. (iii) It uses dilated convolutions to enhance the performance of the designed network and reduce the computational cost.

### 3.2 Discussion

The proposed method relies on residual learning, BN and dilated convolution, they are complementary for image denoising. In this part, we will prove the effectiveness of these methods for

Table 1 Average PSNR (dB) results from different methods on BSD 68

Methods	BM3D	WNNM	EPLL	MLP	CSF	TNRD	IRCNN	ECNDN
$\sigma = 15$	31.07	31.37	31.21	—	31.24	31.42	31.63	31.71
$\sigma = 25$	28.57	28.83	28.68	28.96	28.74	28.92	29.15	29.22
$\sigma = 50$	25.62	25.87	25.67	26.03	—	25.97	26.19	26.23



image-denoising. Here CRNet, CRRNet and CRRBNet have the same the number of network layers, convolutional kernel size and initial parameters. Specifically, CRNet consists of Conv and ReLU as shown in Fig. 3, where Conv and ReLU denote the convolution and rectified linear units, respectively. CRRNet consists of Conv, ReLU and residual learning technique as shown in Fig. 4. Here Figs. 1–4 are the schematic diagrams, in this paper CRRBNet consists of Conv, ReLU, residual learning and BN as shown in Fig. 5. First, we illustrate the peak signal-to-noise ratio (PSNR) of every training epoch for CRNet and CRRBNet. From Fig. 6, we know that the combination of BN and residual learning is effective for image denoising. Then, we prove that the dilated convolution is useful for image denoising as shown in Fig. 7.



Fig. 9 Widely used 12 images

## 4 Experimental results

### 4.1 Experimental setting

We design a 17-layer network called ECNDNet. Its depth is the same as denoising CNN (DnCNN). Its loss function (also referred to as objective function) is shown as in (1). We choose Adam [39] to optimise the converge model. The initial parameters are set as follows: (i) learning rate, beta\_1, beta\_2 and epsilon are  $1 \times 10^{-3}$ , 0.9, 0.999 and  $1 \times 10^{-8}$ , respectively. (ii) The initial weights are set as shown in [40]. (iii) The number of batches is 128. (iv) The number of epochs is 180 for the trained model. In addition, the learning rates of the 180 epochs are  $1 \times 10^{-3}$  to  $1 \times 10^{-8}$ .

We choose PyTorch tool [41] to train the denoising model in this paper. All the experiments are implemented in the environment of Ubuntu 16.04 and python 2.7 and run on PC with Intel Core i7 7800X CPU, RAM 16G and a Nvidia GeForce GTX 1080 Ti GPU. The types of Nvidia CUDA and cuDNN are 9.0 and 7.5, respectively.

### 4.2 ECNDNet for grey image denoising

We choose 400 images [42] with size of  $180 \times 180$  for Gaussian denoising. The format of training images is '.png'. According to the IRCNN [17] and fast and flexible denoising network (FFDNet) [15], we use BSD68 [43] and Set 12 to test the denoising model. In addition, we use popular methods such as BM3D [6], WNNM [13], expected patch log likelihood (EPLL) [38], cascade of shrinkage fields (CSF) [44], trainable nonlinear reaction diffusion (TNRD) [42], IRCNN [17] and multi-layer perceptron [45] to verify the performance of gray image denoising. To test the robustness of our proposed method for low-level and high-level noise, we choose  $\sigma = 15$ ,  $\sigma = 25$  and  $\sigma = 50$  to conduct comparative experiments. For example, the PSNR of our proposed method is 31.71 dB higher than that of the state-of-the-art method such as IRCNN as shown in Table 1 ( $\sigma = 15$ ). Besides, the

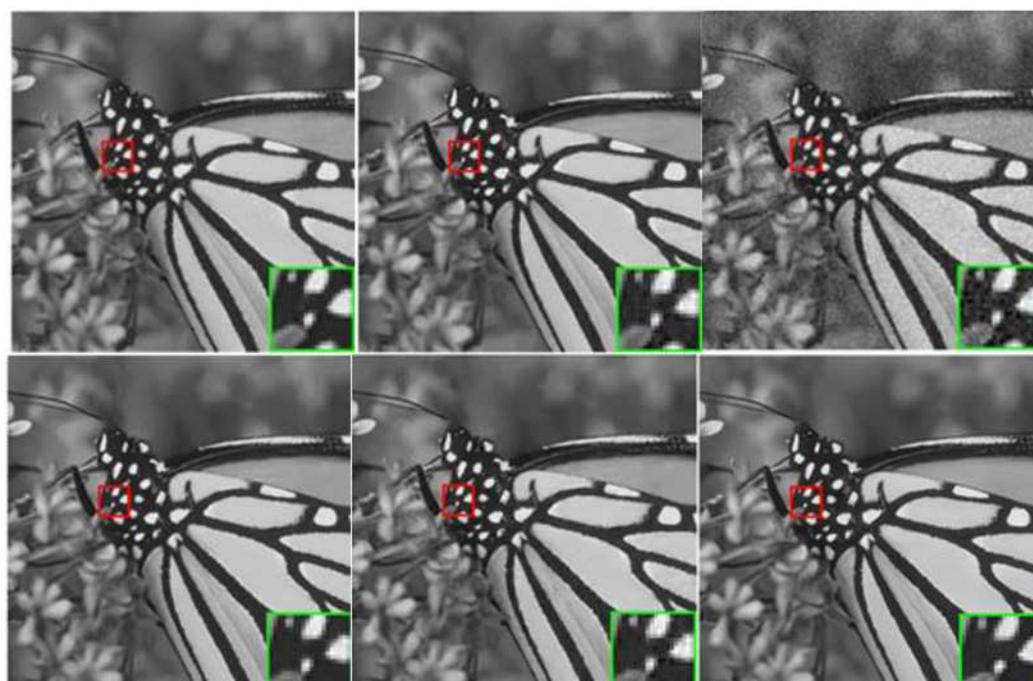


Fig. 10 Denoising results of one grey image with  $\sigma = 15$

- a Original image
- b BM3D/31.85 dB
- c Noisy/24.59 dB
- d WNNM/32.71 dB
- e EPLL/32.10 dB
- f ECNDNet/33.11 dB

**Table 3** Run time of different methods on image of size  $256 \times 256$ ,  $512 \times 512$  and  $1024 \times 1024$  with noise level 25

Methods	Device	$256 \times 256$	$512 \times 512$	$1024 \times 1024$
		Gray	Gray	Gray
BM3D	CPU	0.65	2.85	11.89
WNNM	CPU	203.1	773.2	2536.4
EPLL	CPU	25.4	45.5	422.1
MLP	CPU	1.42	5.51	19.4
CSF	CPU/GPU	2.11/—	5.67/0.92	40.8/1.72
TNRD	CPU/GPU	0.45/0.010	1.33/0.032	4.61/0.116
DnCNN-s	GPU	0.008	0.068	0.154
ECNDNet	GPU	0.012	0.079	0.205

best and second best performance are shown in italic and bold, respectively.

We use Fig. 8 to vividly show the performance of our method and other comparative methods with  $\sigma = 50$  on BSD68 dataset. To show the performance of our proposed method for the images of different categories, we validate it using the Set12 dataset.

From Table 2, it is known that our proposed method has good performance for each category image (Fig. 9). For example, the average PSNR of our method is 30.39 dB higher than that of BM3D when noise level is 25. Specifically, the best PSNR is marked in *italic* and the second PSNR is marked in **bold** as shown in Table 2. The detailed results of the comparative experiments are shown in [3, 15, 17]. Fig. 10 shows the denoising performance of different methods of an image.

## 5 Run time

PSNR and run time of processing an image are two important factors of image denoising. The performance of the proposed method has been proved in Section 4.2. The run time of processing an image is tested for gray image denoising as follows. We utilise noisy image sizes of  $256 \times 256$ ,  $512 \times 512$  and  $1024 \times 1024$  with  $\sigma = 50$  to test the speed of different methods for an image. Specifically, we use PyTorch to test run time of DnCNN-s and ECNDNet. From Table 3, we know that our ECNDNet is competitive with popular methods such as BM3D, WNNM, EPLL, CSF, TNRD and DnCNN-s in run time. In summary, our proposed method is robust for image denoising.

## 6 Conclusion

In this paper, a deep CNN called ECNDNet is proposed to solve the image denoising problem.

Specifically, BN, residual learning and dilated convolution are used to enhance network performance. BN can deal with internal covariate shift problem and makes the network easier to train. Residual learning technique can address the problem of vanishing or exploding gradients. It is used to obtain clean images from noisy images and residual images. Dilated convolution can extract more context information and reduce the computational cost. In addition, BN, residual learning and dilated convolution are complement for image denoising. Extensive experiments show that ECNDNet is more effective than the popular denoising methods such as IRCNN. In the future, we will combine model base-optimisation and discriminative learning methods to remove the noise from real noisy images.

## 7 Acknowledgments

This paper was supported in part by the Guangdong Province high-level personnel of special support program under grant no. 2016TX03X164, in part by the Shenzhen Municipal Science

and Technology Innovation Council under grant no. JCYJ20170811155725434.

## 8 References

- [1] Li, S., Yin, H., Fang, L.: ‘Group-sparse representation with dictionary learning for medical image denoising and fusion’, *IEEE Trans. Biomed. Eng.*, 2006, **59**, (12), pp. 3450–3459
- [2] Zhang, L., Zuo, W.: ‘Image restoration: from sparse and low-rank priors to deep priors’, *IEEE Signal Process. Mag.*, 2017, **34**, (5), pp. 172–179
- [3] Zhang, K., Zuo, W., Chen, Y., *et al.*: ‘Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising’, *IEEE Trans. Image Process.*, 2017, **26**, (7), pp. 3142–3155
- [4] Mallat, M., Roose, D.: ‘Wavelet-based image denoising using a Markov random field a priori model’, *IEEE Trans. Image Process.*, 1996, **6**, (4), pp. 549–565
- [5] Mairal, J., Bach, F., Ponce, J., *et al.*: ‘Non-local sparse models for image restoration’. Proc. IEEE Int. Conf. Computer Vision, September/October 2009, pp. 2272–2279
- [6] Dabov, K., Foi, A., Katkovnik, V., *et al.*: ‘Image denoising by sparse 3-D transform-domain collaborative filtering’, *IEEE Trans. Image Process.*, 2007, **16**, (8), pp. 2080–2095
- [7] Zuo, W., Zhang, L., Song, C., *et al.*: ‘Gradient histogram estimation and preservation for texture enhanced image denoising’, *IEEE Trans. Image Process.*, 2014, **23**, (6), pp. 2459–2472
- [8] Dong, W., Zhang, L., Shi, G., *et al.*: ‘Nonlocally centralized sparse representation for image restoration’, *IEEE Trans. Image Process.*, 2013, **22**, (4), pp. 1620–1630
- [9] Beck, A., Teboulle, M.: ‘Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems’, *IEEE Trans. Image Process.*, 2009, **18**, (11), pp. 2419–2434
- [10] Zhu, M., Chan, T.: ‘An efficient primal-dual hybrid gradient algorithm for total variation image restoration’, UCLA CAM Report, <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.474.898&rep=rep1&type=pdf>, 2008
- [11] Chan, T.F., Chen, K.: ‘An optimization-based multilevel algorithm for total variation image denoising’, *Multiscale. Model. Simul.*, 2006, **5**, (2), pp. 615–645
- [12] Frohn, C., Henn, S., Witsch, K.: ‘Nonlinear multigrid methods for total variation image denoising’, *Comput. Vis. Sci.*, 2004, **7**, (3–4), pp. 199–206
- [13] Gu, S., Zhang, L., Zuo, W., *et al.*: ‘Weighted nuclear norm minimization with application to image denoising’. Proc. IEEE Conf. Computer Vision Pattern Recognition, June 2014, pp. 2862–2869
- [14] Lefkimiatis, S.: ‘Universal denoising networks: a novel CNN architecture for image denoising’. Proc. IEEE Conf. Computer Vision Pattern Recognition, June 2018, pp. 3204–3213
- [15] Zhang, K., Zuo, W., Zhang, L.: ‘FFDNet: toward a fast and flexible solution for CNN based image denoising’, *IEEE Trans. Image Process.*, 2018, pp. 4608–4622
- [16] Lefkimiatis, S.: ‘Non-local color image denoising with convolutional neural networks’. Proc. IEEE Conf. Computer Vision Pattern Recognition, July 2017, pp. 3587–3596
- [17] Zhang, K., Zuo, W., Gu, S., *et al.*: ‘Learning deep CNN denoiser prior for image restoration’. Proc. IEEE Conf. Computer Vision Pattern Recognition, July 2017, pp. 3587–3596
- [18] Du, B., Xiong, W., Wu, J., *et al.*: ‘Stacked convolutional denoising auto-encoders for feature representation’, *IEEE Trans. Cybern.*, 2017, **47**, (4), pp. 1017–1027
- [19] Wu, D., Kim, K., Fakhri, G.E., *et al.*: ‘A cascaded convolutional neural network for X-ray low-dose CT image denoising’, arXiv preprint arXiv:1705.04267, 2017
- [20] Bako, S., Vogels, T., McWilliams, B., *et al.*: ‘Kernel-predicting convolutional networks for denoising Monte Carlo renderings’, *ACM Trans. Graph.*, 2017, **36**, (4), pp. 1–14
- [21] Ahn, B., Cho, N.I.: ‘Block-matching convolutional neural network for image denoising’, arXiv preprint arXiv:1704.00524, 2017
- [22] Liu, P., Zhang, H., Zhang, K., *et al.*: ‘Multi-level wavelet-CNN for image restoration’, arXiv preprint arXiv:1805.07071, 2018
- [23] Tian, C., Xu, Y., Fei, L., *et al.*: ‘Deep learning for image denoising: a survey’, arXiv preprint arXiv:1810.05052, 2018
- [24] Kang, E., Chang, W., Yoo, J., *et al.*: ‘Deep convolutional framelet denoising for low-dose ct via wavelet residual network’, *IEEE Trans. Med. Imaging*, 2018, **37**, (6), pp. 1358–1369
- [25] Gondara, L.: ‘Medical image denoising using convolutional denoising auto encoders’. Proc. IEEE Int. Conf. Data Mining Workshops (ICDMW), 2016, pp. 241–246
- [26] Kokkinos, F., Lefkimiatis, S.: ‘Deep image demosaicking using a cascade of residual denoising networks’, arXiv preprint arXiv:1803.05215, 2018
- [27] He, K., Zhang, X., Ren, S., *et al.*: ‘Deep residual learning for image recognition’. IEEE Int. Conf. Computer Vision, June 2016, pp. 770–778
- [28] Ioffe, S., Szegedy, C.: ‘Batch normalization: accelerating deep network training by reducing internal covariate shift’, arXiv preprint arXiv:1502.03167, 2015
- [29] Yu, F., Koltun, V.: ‘Multi-scale context aggregation by dilated convolutions’, arXiv preprint arXiv:1511.07122, 2015
- [30] He, K., Zhang, X., Ren, S., *et al.*: ‘Delving deep into rectifiers: surpassing human-level performance on imagenet classification’. IEEE Int. Conf. Computer Vision, June 2015, pp. 1026–1034
- [31] Duchi, J., Hazan, E., Singer, Y.: ‘Adaptive subgradient methods for online learning and stochastic optimization’, *J. Mach. Learn. Res.*, 2011, **12**, pp. 2121–2159

- [32] Kingma, D., Ba, J.: 'Adam: a method for stochastic optimization'. *Int. Conf. for Learning Representations*, 2015
- [33] Krizhevsky, A., Sutskever, I., Hinton, G.E.: 'Imagenet classification with deep convolutional neural networks', *Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105
- [34] Tian, C., Zhang, Q., Sun, G., *et al.*: 'FFT consolidated sparse and collaborative representation for image classification', *Arab. J. Sci. Eng.*, 2018, **42**, (2), pp. 741–758
- [35] Fei, L., Lu, G., Jia, W., *et al.*: 'Feature extraction methods for palmprint recognition: a survey and evaluation', *IEEE Trans. Syst., Man, Cybern. Syst.*, 2018
- [36] Guo, K., Wu, S., Xu, Y.: 'Face recognition using both visible light image and near-infrared image and a deep network', *CAAI Trans. Intell. Technol.*, 2017, **2**, (1), pp. 39–47
- [37] Wang, T., Sun, M., Hu, K.: 'Dilated residual network for image denoising', arXiv preprint arXiv:1708.05473, 2017
- [38] Zoran, D., Weiss, Y.: 'From learning models of natural image patches to whole image restoration'. *IEEE Conf. Computer Vision*, June 2011, pp. 479–486
- [39] Kingma, D., Adam, J.B.: 'A method for stochastic optimization', *Int. Conf. Learn. Representations (ICLR)*, 2015, **5**
- [40] He, K., Zhang, X., Ren, S., *et al.*: 'Delving deep into rectifiers: surpassing human-level performance on imagenet classification'. *IEEE Conf. Computer Vision*, June 2015, pp. 1026–1034
- [41] Paszke, A., Gross, S., Chintala, S.: 'Pytorch', 2017
- [42] Chen, Y., Pock, T.: 'Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2016, **PP**, (99), pp. 1–1
- [43] Roth, S., Black, M.J.: 'Fields of experts', *Int. J. Comput. Vis.*, 2009, **82**, (2), pp. 205–229
- [44] Schmidt, U., Roth, S.: 'Shrinkage fields for effective image restoration'. *Proc. IEEE Conf. Computer Vision Pattern Recognition*, June 2014, pp. 2774–2781
- [45] Burger, H.C., Schuler, C.J., Harmeling, S.: 'Image denoising: Can plain neural networks compete with BM3D?'. *Proc. IEEE Conf. Computer Vision Pattern Recognition*, June 2012, pp. 2392–2399