

Deep Learning for Image Denoising: A Survey

Chunwei Tian^{1,2}, Yong Xu^{1,2,*}, Lunke Fei³, and Ke Yan^{1,2}

1. Bio-Computing Research Center, Harbin Institute of Technology, Shenzhen, Shenzhen. 518055 Guangdong, China
2. Shenzhen Medical Biometrics Perception and Analysis Engineering Laboratory, Shenzhen, Shenzhen. 518055 Guangdong, China
3. School of Computers, Guangdong University of Technology, Guangzhou 510006, China

{chunweitian@163.com, yongxu@ymail.com, flksxm@126.com, yanke401@163.com}

Abstract. Since the proposal of big data analysis and Graphic Processing Unit (GPU), the deep learning technology has received a great deal of attention and has been widely applied in the field of imaging processing. In this paper, we have an aim to completely review and summarize the deep learning technologies for image denoising proposed in recent years. Moreover, we systematically analyze the conventional machine learning methods for image denoising. Finally, we point out some research directions for the deep learning technologies in image denoising. *abstract environment.*

Keywords: Deep Learning, Convolution neural networks, GPU; Image Denoising.

1 Introduction

Image processing has numerous applications including image segmentation [28], image classification [25,38,32,12], object detection [13], video tracking [36], image restoration [48] and action recognition [35]. Especially, the image denoising technology is one of the most important branches of image processing technologies and is used as an ex-ample to show the development of the image processing technologies in last 20 years [42]. Buades et al. [5] proposed a non-local algorithm method to deal with image denoising. Lan et al. [19] fused the belief propagation inference method and Markov Random Fields (MRFs) to address image denoising. Dabov et al. [9] proposed to transform grouping similar two-dimensional image fragments into three-dimensional data arrays to improve sparsity for image denoising. These selection and extraction methods have amazing performance for image denoising. However, the conventional methods have two challenges [45]. First, these methods are non-convex, which need to manually set parameters. Second, these methods refer a complex optimization problem for the test stage, resulting in high computational cost.

In recent years, researches have shown that deep learning technologies can reply to deeper architecture to automatically learn and find more suitable image

features rather than manual setting parameters, which effectively address drawbacks of traditional methods mentioned above [18]. Big data and GPU are also essential for deep learning technologies to improve the learning ability [16]. The learning ability of deep learning is finished by model (also referred to as network) and the model consists of many layers, including the convolutional layer, pooling layer, batch normalization layer and full connection layer. In other words, deep learning technologies can convert input data (e.g. images, speech and video) into outputs (e.g. object category, password unlocking and traffic information) by the model [24]. Especially, convolutional neural network (CNN) is one of the most typical and successful deep learning network for image processing [20]. CNN was originated LeNet from 1998 and it was successfully used in hand-written digit recognition, achieving excellent performance [21]. However, convolutional neural networks (CNNs) haven't been widely used in other real applications before the arise of GPU and big data. In other words, the real success of CNNs attributed to ImageNet Large Scale Visual Recognition Challenge 2012 (ILSVRC 2012) where new CNN was proposed, named AlexNet and became a world champion in this ILSVRC 2012 [18,43].

In subsequent years, deeper neural networks have becoming popular and obtain promising performance for image processing [29]. Karen Simonyan et al. [29] in-creased the depth of neural networks to 16-19 weighted layers and convolution filter size of each layer was 3×3 for image recognition. Christian Szegedy et al. [30] provided a mechanism by using sparsely connected layer [2] instead of fully connected layers to increase the width and depth of the neural networks for image classification, named as Inception V1. Inception V1 effectively prevented to overfitting from enlarged size (width) of network and reduced the computing resource from increased depth of network. Previous researches show that the deep networks essentially use an end-to-end multilayer fashion to fuse different level fashion [17] and classifiers and the extracted features can be more robust by increasing the number of depth in networks. Despite deep networks have obtained successfully applications for image processing [27], they can generate vanishing gradient or exploding gradient [4] with increased network depth. That makes network hamper convergence. This problem can be solved by normalized initialization [39]. However, when deeper neural networks get to converge, networks are saturated and degrade quickly with increasing depth of networks. The appearance of residual network effectively dealt with problems above for image recognition [15]. ResNeXt method is tested to be very effectively for image classification [40]. Spatial-temporal Attention (SPA) method is very competitive for visual tracking [50]. Residual Dense Network (RDN) is also an effective tool for image super-resolution [49]. Furthermore, DiracNets [44], IndRNN [23] and variational U-Net [11] also provide us with many competitive technologies for image processing. These deep networks are also widely applied in image denoising, which is the branch of image processing technologies. For example, the combination of kernel-prediction net and CNN is used to obtain denoised image [3]. BMCNN utilizes NSS and CNN to deal with image denoising [1]. GAN is used to remove noise from noisy image [33].

Although the researches above expose that deep learning technologies have obtained enormous success in the applications of image denoising, to own knowledge, there is no comparative study of deep learning technologies for image denoising. Deep learning technologies refer to properties of image denoising to propose wise solution methods, which are embedded in multiple hidden layers with end-end connection to better deal with them. Therefore, a survey is important and necessary to review the principles, performance, difference, merits, shortcomings and technical potential for image processing. Deeper CNNs (e.g. AlexNet, GoogLeNet, VGG and ResNet), which can show the ideas of deep learning technologies and successful reasons for image denoising. To better show the robustness of deep learning denoising, the performance of deep learning for image denoising is shown. The potential challenges and directions of deep learning technologies for image denoising in the future are also offered in this paper.

The remainder of this paper is organized as follows. Section 2 overviews of typical deep learning methods. Section 3 provides deep learning technologies for image denoising. Section 4 points out some potential research directions. Section 6 presents the conclusions of this paper.

2 Typical deep network

Nowadays, the most widely used model is trained with end-to-end in a supervised fashion, which is easy and simple to implement to train models. The popular network architecture is CNNs (ResNet). This network is widely used to deal with applications of image processing and obtain enormous success. The following sections will show the popular deep learning technology, discuss the merits and differences of the method in Section 2.

2.1 ResNet

Deep CNNs have result in a lot of breakthroughs for image recognition. Especially, deep network plays an important role on image classification [30]. Many other visual recognition applications are beneficial from deep networks. However, deeper network can have vanishing/exploding gradients [30]. This problem has been effectively solved by normalized initialization [33], which makes the network converge. When this network starts converging, performance of the network gets degraded. For example, the depth of this network are increased, the errors in the training model are increasing. The problem is effectively addressed by ResNet [15]. The ideas of ResNet are that outputs of each two layers and their inputs are added as the new input. ResNet include many blocks and the block is shown in Fig.1, where x and f , respectively, denote input and activation function. A residual block is obtained by $f(x) + x$. The ResNet is popular based on the following reasons. First, ResNet is deep rather than width, which effectively controls the number of parameters and overcomes the overfitting problem. Second, it uses less pooling layers and more downsampling operations to improve transmission efficiency. Third, it uses BN and average pooling for regularization,

which accelerate the speed of training model. Finally, it uses 3×3 filters of each convolutional layer to train model, which is faster than using the combination of 3×3 and 1×1 filters. As a result, ResNest takes the first place in ILSVRC 2015 and reduces 3.57% error on the ImageNet test set.

In addition, deformation networks of Residual network are popular and have been widely used in image classification, image denoising [41] and image resolution [31].

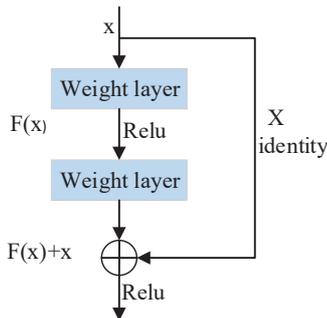


Fig. 1. Residual network: a building block

3 Image Denoising

Image denoising is topic applications for image processing. We take image denoising as an example to show the performance and principle for deep learning technologies in image processing applications.

The aim of image denoising is to obtain clean image x from a noisy image y which is explained by $y = x + n$. n denotes the additive white Gaussian noise (AWGN) with variance σ^2 . From the machine learning knowledge, we know that the image prior is an important for image denoising. In the past ten years, a lot of methods are proposed for model with image priors, such as Markov random field (MRF) method [19], BM3D [9], NCSR [10] and NSS [6]. Although these methods perform well for image denoising, they have two drawbacks. First, these methods need to optimize, which results in increasing computational cost. Second, these methods are non-convex, which need manual settings to improve performance. To address the problems, some discriminative learning schemes were proposed. A trainable nonlinear reaction diffusion method was proposed and used to learn image prior [26]. A cascade of shrinkage fields fuse the random field-based model and half-quadratic algorithm into a single architecture [46]. Despite methods improve the performances for image denoising, they are limited to the specified forms of prior. Another shortcoming is that these methods cant use a model to deal with blind image denoising.

Deep learning technologies can effectively deal with problems above. And deep learning technologies are chosen for image denoising based on the following three-fold. First, they have deep architecture, which can learn more extractions. Second, BN and ReLU are added into deep architectures, which can accelerate the training speed. Third, networks of deep learning methods can run on GPU, which can train more samples and improve the efficiency. The proposed DnCNN [45] uses BN and ResNet to perform image denoising. This network not only deals with blind image denoising, but also addresses image super-resolution task and JPEG image deblocking. Its architecture is as shown in Fig. 2. Specifically, it obtains the residual image from the model and it needs to use $y = x + n$ to obtain clean image when it is in the test phase. It obtained PSNR of 29.13, which is higher than the state-of-the-art BM3D method of 28.57 for BSD68 dataset with $\sigma = 25$.

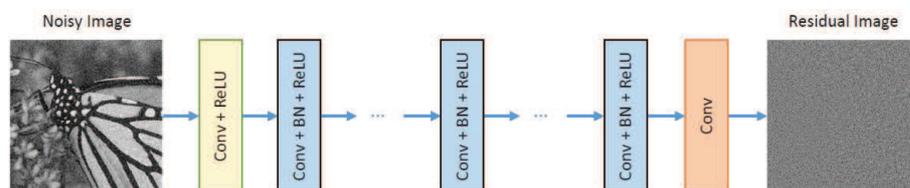


Fig. 2. The architecture of DnCNN

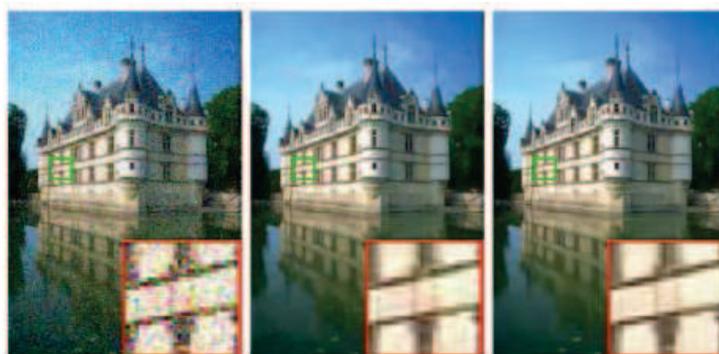


Fig. 3. Results of CBM3D and FFDNet for color image denoising (a)Noisy($\sigma=35$) (b)CBM3D(29.90dB) (c)FFDNet(30.51dB)

FFDNet [46] uses noise level map and noisy image as input to deal with different noise levels. This method exploits a single model to deal with multiple noise levels. It is also faster than BM3D on GPU and CPU. As shown in

Fig.3, performance of FFDNet outperforms the CBM3D [19] method in image denoising. IRCNN [47] fuses the model-based optimization method and CNN to address image denoising problem, which can deal with different inverse problems and multiple tasks with one single mode. In addition, it adds dilated convolution into network, which improves the performance for denoising. Its architecture is shown as Fig. 4.

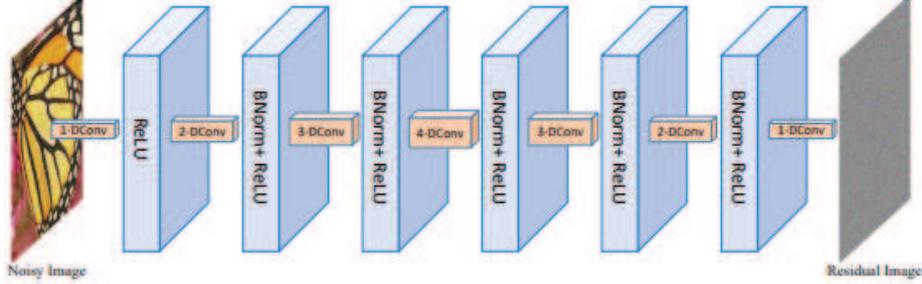


Fig. 4. The architecture of IRCNN

In addition, many other methods also obtain well performance for image denoising. For example, fusion of the dilated convolution and ResNet is used for image denoising [37]. It is a good choice for combing disparate sources of experts for image denoising [8]. Universal denoising networks [22] for image denoising and deep CNN denoiser prior to eliminate multiplicative noise [34] are also effective for image denoising. As shown in Table 1, deep learning methods are superior to the conventional methods. And the DnCNN method obtains excellent performance for image denoising.

Table 1. Comparisons of different methods with $\sigma = 25$ for image denoising.

Methods	PSNR	Dataset
BM3D [9]	28.57	BSD68
WNNM [14]	28.83	BSD68
TNRD [7]	28.92	BSD68
DnCNN [45]	29.23	BSD68
FFDNet [46]	29.19	BSD68
IRCNN [47]	29.15	BSD68
DDRN [37]	29.18	BSD68

4 Research directions

4.1 The challenges of deep learning technologies in image denoising

According to existing researches, deep learning technologies achieve promising results in image denoising. However, these technologies also suffer from some challenges as follows. (1) Current deep learning denoising methods only deal with AWGN, which are not effective for real noisy images, such as low light images. (2) They can't use a model to deal with all the low level vision tasks, such as image denoising, image super-resolution, image blurring and image deblocking. (3) They can't use a model to address the blind Gaussian noise.

4.2 Some potential directions of deep learning technologies for image denoising

According to the previous researches, deep learning technologies have the following changes for image denoising application above. First, deep learning technologies design different network architectures to deal with tasks above. Second, they can fuse the optimization and discrimination methods. Third, they can use multiple tasks to design the network. Fourth, they can change the input of the neural networks.

5 Conclusion

This paper first comprehensively introduces the development of deep learning technologies on image processing applications. And then shows the implementations of typical CNNs. After that, image denoising is illustrated in detail, which concludes the differences and ideas of different methods for image denoising in real world. Finally, this paper shows the challenges of deep learning methods for image processing applications and offers solutions. This review offers important cues on deep learning technologies for image processing applications. We believe that this paper could provide researchers with a useful guideline working in the related fields, especially for the beginners worked in deep-learning.

6 Acknowledgment

This paper was supported in part by Shenzhen Municipal Science and Technology Innovation Council under Grant no. JCYJ20170811155725434, in part by the National Natural Science Foundation under Grant no. 61876051.

References

1. Byeongyong Ahn and Nam Ik Cho. Block-matching convolutional neural network for image denoising. *arXiv preprint arXiv:1704.00524*, 2017.

2. Sanjeev Arora, Aditya Bhaskara, Rong Ge, and Tengyu Ma. Provable bounds for learning some deep representations. In *International Conference on Machine Learning*, pages 584–592, 2014.
3. Steve Bako, Thijs Vogels, Brian McWilliams, Mark Meyer, Jan Novák, Alex Harvill, Pradeep Sen, Tony Derose, and Fabrice Rousselle. Kernel-predicting convolutional networks for denoising monte carlo renderings. *ACM Trans. Graph*, 36(4):97, 2017.
4. Yoshua Bengio, Patrice Simard, and Paolo Frasconi. Learning long-term dependencies with gradient descent is difficult. *IEEE transactions on neural networks*, 5(2):157–166, 1994.
5. Antoni Buades, Bartomeu 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*, volume 2, pages 60–65. IEEE, 2005.
6. Antoni Buades, Bartomeu Coll, and Jean-Michel Morel. Nonlocal image and movie denoising. *International journal of computer vision*, 76(2):123–139, 2008.
7. Yunjin Chen and Thomas Pock. Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration. *IEEE transactions on pattern analysis and machine intelligence*, 39(6):1256–1272, 2017.
8. Joon Hee Choi, Omar Elgendy, and Stanley H Chan. Integrating disparate sources of experts for robust image denoising. *arXiv preprint arXiv:1711.06712*, 2017.
9. Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on image processing*, 16(8):2080–2095, 2007.
10. Weisheng Dong, Lei Zhang, Guangming Shi, and Xin Li. Nonlocally centralized sparse representation for image restoration. *IEEE Transactions on Image Processing*, 22(4):1620–1630, 2013.
11. Patrick Esser, Ekaterina Sutter, and Björn Ommer. A variational u-net for conditional appearance and shape generation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8857–8866, 2018.
12. Lunke Fei, Guangming Lu, Wei Jia, Shaohua Teng, and David Zhang. Feature extraction methods for palmprint recognition: A survey and evaluation. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2018.
13. Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 580–587, 2014.
14. Shuhang Gu, Lei Zhang, Wangmeng Zuo, and Xiangchu Feng. Weighted nuclear norm minimization with application to image denoising. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2862–2869, 2014.
15. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
16. Qibin Hou, Ming-Ming Cheng, Xiaowei Hu, Ali Borji, Zhuowen Tu, and Philip Torr. Deeply supervised salient object detection with short connections. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5300–5309. IEEE, 2017.
17. Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*, 2015.

18. Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.
19. Xiangyang Lan, Stefan Roth, Daniel Huttenlocher, and Michael J Black. Efficient belief propagation with learned higher-order markov random fields. In *European conference on computer vision*, pages 269–282. Springer, 2006.
20. Steve Lawrence, C Lee Giles, Ah Chung Tsoi, and Andrew D Back. Face recognition: A convolutional neural-network approach. *IEEE transactions on neural networks*, 8(1):98–113, 1997.
21. Yann LeCun, Bernhard Boser, John S Denker, Donnie Henderson, Richard E Howard, Wayne Hubbard, and Lawrence D Jackel. Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4):541–551, 1989.
22. Stamatios Lefkimiatis. Universal denoising networks: A novel cnn architecture for image denoising. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3204–3213, 2018.
23. Shuai Li, Wanqing Li, Chris Cook, Ce Zhu, and Yanbo Gao. Independently recurrent neural network (indrnn): Building a longer and deeper rnn. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5457–5466, 2018.
24. Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghahfoorian, Jeroen AWM van der Laak, Bram Van Ginneken, and Clara I Sánchez. A survey on deep learning in medical image analysis. *Medical image analysis*, 42:60–88, 2017.
25. Yongbin Qin and Chunwei Tian. Weighted feature space representation with kernel for image classification. *Arabian Journal for Science and Engineering*, pages 1–13, 2017.
26. Uwe Schmidt and Stefan Roth. Shrinkage fields for effective image restoration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2774–2781, 2014.
27. Pierre Sermanet, David Eigen, Xiang Zhang, Michaël Mathieu, Rob Fergus, and Yann LeCun. Overfeat: Integrated recognition, localization and detection using convolutional networks. *arXiv preprint arXiv:1312.6229*, 2013.
28. Jianbo Shi and Jitendra Malik. Normalized cuts and image segmentation. *IEEE Transactions on pattern analysis and machine intelligence*, 22(8):888–905, 2000.
29. Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
30. Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9, 2015.
31. Ying Tai, Jian Yang, Xiaoming Liu, and Chunyan Xu. Memnet: A persistent memory network for image restoration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4539–4547, 2017.
32. Chunwei Tian, Qi Zhang, Guanglu Sun, Zhichao Song, and Siyan Li. Fft consolidated sparse and collaborative representation for image classification. *Arabian Journal for Science and Engineering*, 43(2):741–758, 2018.
33. Subarna Tripathi, Zachary C Lipton, and Truong Q Nguyen. Correction by projection: Denoising images with generative adversarial networks. *arXiv preprint arXiv:1803.04477*, 2018.

34. Guodong Wang, GuoTao Wang, Zhenkuan Pan, and Zhimei Zhang. Multiplicative noise removal using deep cnn denoiser prior. In *Intelligent Signal Processing and Communication Systems (ISPACS), 2017 International Symposium on*, pages 1–6. IEEE, 2017.
35. Heng Wang, Alexander Kläser, Cordelia Schmid, and Cheng-Lin Liu. Action recognition by dense trajectories. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 3169–3176. IEEE, 2011.
36. Li Wang, Ting Liu, Gang Wang, Kap Luk Chan, and Qingxiong Yang. Video tracking using learned hierarchical features. *IEEE Transactions on Image Processing*, 24(4):1424–1435, 2015.
37. Tianyang Wang, Mingxuan Sun, and Kaoning Hu. Dilated residual network for image denoising. *arXiv preprint arXiv:1708.05473*, 2017.
38. Jie Wen, Xiaozhao Fang, Yong Xu, Chunwei Tian, and Lunke Fei. Low-rank representation with adaptive graph regularization. *Neural Networks*, 108:83–96, 2018.
39. Yuxin Wu and Kaiming He. Group normalization. *arXiv preprint arXiv:1803.08494*, 2018.
40. Saining Xie, Ross Girshick, Piotr Dollár, Zhuowen Tu, and Kaiming He. Aggregated residual transformations for deep neural networks. In *Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on*, pages 5987–5995. IEEE, 2017.
41. Weiyang Xie, Yunsong Li, and Xiuping Jia. Deep convolutional networks with residual learning for accurate spectral-spatial denoising. *Neurocomputing*, 2018.
42. Jun Xu, Lei Zhang, Wangmeng Zuo, David Zhang, and Xiangchu Feng. Patch group based nonlocal self-similarity prior learning for image denoising. In *Proceedings of the IEEE international conference on computer vision*, pages 244–252, 2015.
43. Yang You, Zhao Zhang, Cho-Jui Hsieh, James Demmel, and Kurt Keutzer. 100-epoch imagenet training with alexnet in 24 minutes. *ArXiv e-prints*, 2017.
44. Sergey Zagoruyko and Nikos Komodakis. Diracnets: training very deep neural networks without skip-connections. *arXiv preprint arXiv:1706.00388*, 2017.
45. Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE Transactions on Image Processing*, 26(7):3142–3155, 2017.
46. Kai Zhang, Wangmeng Zuo, and Lei Zhang. Ffdnet: Toward a fast and flexible solution for cnn based image denoising. *IEEE Transactions on Image Processing*, 2018.
47. Kai Zhang, Wangmeng Zuo, and Lei Zhang. Learning a single convolutional super-resolution network for multiple degradations. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 6, 2018.
48. Lei Zhang and Wangmeng Zuo. Image restoration: From sparse and low-rank priors to deep priors. *IEEE Signal Processing Magazine*, 34(5):172–179, 2017.
49. Yulun Zhang, Yapeng Tian, Yu Kong, Bineng Zhong, and Yun Fu. Residual dense network for image super-resolution. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.
50. Zheng Zhu, Wei Wu, Wei Zou, and Junjie Yan. End-to-end flow correlation tracking with spatial-temporal attention. *illumination*, 42:20, 2017.