DEEP MULTI-SCALE GABOR WAVELET NETWORK FOR IMAGE RESTORATION

Hang Dong[1*], *Xinyi Zhang*[2*] *Yu Guo*[2], *Fei Wang*[1]

[1]College of Artificial Intelligence, Xi'an Jiaotong University
[2]School of Software Engineering, Xi'an Jiaotong University

ABSTRACT

Due to the limitations of the imaging processors and complex weather conditions, image degradation is often inevitable. Existing deep learning-based image restoration methods often rely on the powerful feature representation capacity of deep networks and pay less attention to the inherent properties of the degradation signal, e.g. variations in spatial scale and orientations across the image, which makes them ineffective for the image restoration tasks. In this paper, we propose a Multi-scale Gabor Wavelet Network (MsGWN) for image restoration. We apply the multi-scale architecture to extract the contaminated feature from input at different spatial scales, and thus the contaminated feature can be effectively restored in a corse-to-fine manner. However, using multi-scale architecture alone cannot remove the degradations with different orientations. To overcome this problem, we introduce a Gabor Wavelet Module (GWM) to further extract the contaminated features from four orientations. By decomposing the features into four multi-orientation components, the restoration process can be facilitated by avoiding learning the mixed degradations all-in-one. We evaluate the proposed method on image demoiréing, image deraining, and image dehazing. Experiments on these applications demonstrate that the proposed method can achieve favorable results against the state-of-the-art approaches.

Index Terms— Image restoration, Multi-scale, Multi-orientation, Gabor Wavelet, Deep learning

1. INTRODUCTION

Due to the limitation of imaging hardware and complex environment, image degradation is inevitable. Image restoration aims to estimate a clean image from the degraded one. Although traditional methods based on some image priors can generate favorable results, they are less effective when the priors do not hold for certain scenes. Recently, convolutional neural networks (CNNs) have made great achievements in some image restoration tasks, for instance, image demoiréing, image deraining, and image dehazing.

Image demoiréing. Moiré, shaped like dynamic stripes, frequently appears in photography, which deteriorates the image quality in the circumstance of shooting a digital screen. Early demoiréing methods used filtering [1] or decomposition [2] to solve this problem. However, the complex and dynamic moiré can't be accurately modeled by these traditional methods. Recent CNN-based approaches have taken an exploration of image demoiréing [3, 4]. To address the challenge of moiré spreading, Gao et al. [3] introduced a feature enhancing branch to fuse low and high-frequency features.

Image Deraining. Image deraining methods are developed to deal with rainy images, where the rain-streaks are fickle. Recent approaches adopted deep CNNs to remove rain-streaks [5, 6, 7, 8, 9, 10]. Xia et al. [7] utilized the recurrent neural network(RNNs) to decompose the rain removal into multiple stages. Zhang et al. [6] propose a multi-streaming network for joint rain event detection and deraining.

Image dehazing. Image dehazing aims to restore the haze-free scene from a hazy image. Although the traditional dehazing method [11] is capable of getting fine results, they are limited under complex circumstances. To solve this problem, some deep convolutional neural networks [12, 13, 10, 14] try to directly estimate clean images. Other CNN-based methods [15, 16] are built on the hazy model and use sub-networks to estimate transmission map, airlight, and hazy-free scene simultaneously.

To deal with image restoration, existing CNN-based methods focus on the effectiveness of features exploitation [7, 10]. On the other hand, some algorithms aim to enlarge the receptive field [13]. However, the spatial scales and orientations of the degradations may vary across the image [17], which cannot be effectively solved by existing methods. In general, decomposing the degradations into multiple components with different properties is an efficient way to remove the degradations with various patterns. In [18], the authors propose a Haar Wavelet Module to decompose the contaminated features into different downsampled components while keeping all the information. However, the Haar wavelet adopted in their method can only capture high-frequency information in the horizontal and vertical directions, without considering the information of other orientations.

To tackle with this problem, we design a Multi-scale Gabor Wavelet Network to restore the degradations at different spatial scales and orientations. Specifically, to dispose of the disadvantage of Haar Wavelet Transform, we propose a Gabor

[*] Equally contributed

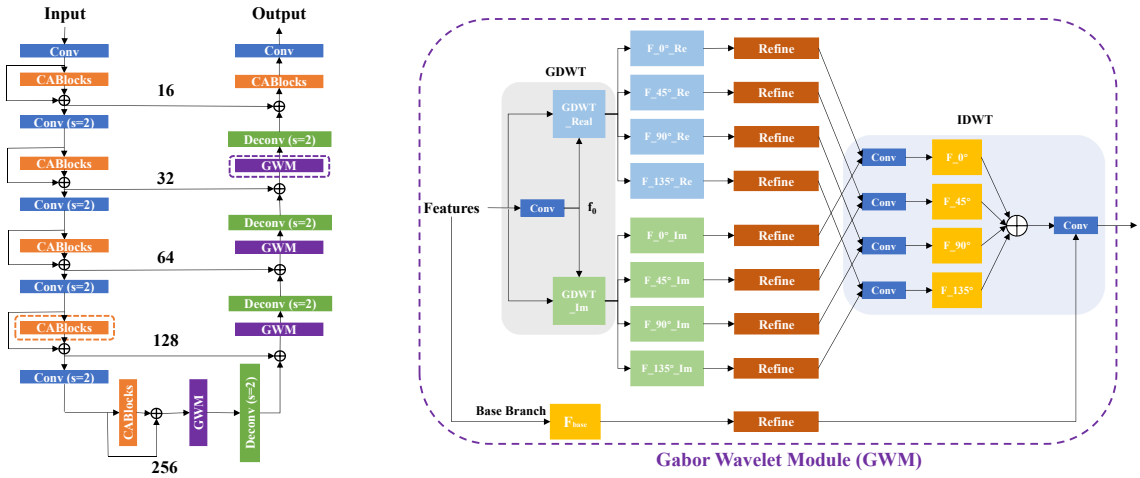

Fig. 1. **Architecture of the proposed MsGWN model.** Our multi-scale model is built upon the U-Net architecture and uses the Gabor Wavelet Module (GWM) to remove the degradations in the wavelet domain.

Wavelet Module to process the contaminated features at different orientations. In this module, we use a Discrete Gabor Wavelet Transform [19] unit to decompose the contaminated features into multi-orientation components. Then, a set of independent refinement units are used to remove the degradations in the contaminated features at different orientations. Finally, a trainable Inverse Discrete Wavelet Transform unit is deployed to reconstruct the restored features from all the components.

The contributions of this work are summarized as follows:

- We propose an effective Multi-scale Gabor Wavelet Network for image restoration. By adopting the multi-scale architecture and Gabor Wavelet Module, our method can effectively remove the degradations with different spatial scales and orientations.
- To preserve the information that cannot be captured by the Discrete Gabor Wavelet Transform, we introduce a base branch paralleled to the branches of Gabor Wavelet components to exploit the remaining information. This base branch can also guarantee that the features can be fully reconstructed in the Inverse Discrete Wavelet Transform.
- We analyze the contribution of the Gabor Wavelet Module and show that the proposed algorithm can generate competitive performance against the state-of-the-art image restoration methods.

2. METHOD

2.1. Network Architecture

To enlarge the receptive field and restore contaminated features at different spatial scales, our model is built upon the U-Net architecture. At the encoder part, we downsample the features of input degraded image to $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$ and $\frac{1}{16}$ of the original resolution. At each scale level of the encoder, channel-attention blocks (CABlocks) are deployed to extract features at different spatial scales. The decoder part is designed to upsample and refine the contaminated features from the encoder. At each scale level of the decoder, a novel Gabor Wavelet Module (GWM) is proposed to process the contaminated features at different orientations. More details of the GWM will be elaborated in Section 2.2.

2.2. Gabor Wavelet Module

To restore the contaminated features at different orientations, we propose a Gabor Wavelet Module (GWM) for each level of the decoder part. The Gabor Wavelet Module consists of four parts: a Discrete Gabor Wavelet Transform (DGWT) unit to decompose the contaminated features, eight independent refinement units to remove the degradations, a trainable Inverse Discrete Wavelet Transform (IDWT) unit to reconstruct the features, and a base branch to reserve the remaining information and guarantee the completion of the reconstructed features.

Gabor Wavelet Transform [20] applies a set of complex functions to decompose a signal into components corresponding to different frequencies and orientations. Therefore, Gabor Wavelets Transform is widely used in image processing because of its strong ability in texture analysis. The 2-D discrete Gabor function can be defined as [21]

$$\Phi_{u,v}(z) = \frac{\|k_{u,v}\|^2}{\sigma^2} e^{(-\|k_{u,v}\|^2 \|z\|^2 2\sigma^2)} [e^{(ik_{u,v}z)} - e^{-\sigma^2/2}], \tag{1}$$

where $\sigma = 2\pi$, $k_{u,v} = \begin{pmatrix} k_v \, cos \, \phi_u \\ k_v \, sin \, \phi_u \end{pmatrix}$, $k_v = f_0 \, / \, 2^{\frac{v}{2}}$ $(v = 0, 1, ..., V_{max})$ is the frequency of Gabor function, f_0 is the reference frequency, and $\phi_u = \frac{u\pi}{U_{max}}$ $(u = 0, 1, ..., U_{max})$ is the orientation of Gabor function.

The network scheme of the GWM is shown in the right part of Figure 1. At the beginning, we use the DGWT in (1) to decompose the contaminated features into multi-orientation components. The proposed DGWT unit consists of

a $GDWT\_Real$ branch to acquire the real part components and a $GDWT\_Im$ branch to acquire the imaginary part components. The $GDWT\_Real$ and $GDWT\_Im$ are implemented by convolutional operations, whose parameters are calculated from (1) and stationary during the training process. For better trade-offs between the performance and efficiency, our DGWT unit has an adaptive frequency ($V_{max} = 0$) and 4 orientations ($U_{max} = 3$). We set up a filter every 45 degrees to accurately extract frequency information with different orientations. Since the frequency of the degradation is various in different degraded images, we apply a convolutional layer to acquire the optimal reference frequency f_0 in (1). Then, each component is fed into a refinement unit to remove the degradation at different orientations. Since limited amounts of discrete Gabor functions cannot preserve all the information during the DGWT process [19], a base branch (the bottom branch in the GWM) is proposed to exploit the remaining information from the contaminated features. Finally, a trainable IDWT unit combines all the wavelet branches and the base branch to reconstruct the recovered features.

2.3. Implementation and Datasets

As shown in Figure 1, we use four strided convolutional layers and four strided deconvolutional layers to downsample and upsample the features. We choose three CABlocks as the refinement unit in the GWM. All the CABlocks have the same configuration as [22]. The filter size is set as 7×7 pixels in the first convolutional layer in the encoder module, 4×4 pixels in all the deconvolutional layers, convolutional layers in DGWT, and the convolutional layers in IDWT, and 3×3 in all other convolutional layers (except the convolutional layers in the CABlock). We train our model with the mean square error (MSE) as the loss function to focus on the fidelity of the network output. The entire training process contains 40 epochs via the ADAM solver [23]. The initial learning rate is set as 5×10^{-5} with a decay rate of 0.5 every 8 epochs.

Datasets. We choose the newly released LCDMoire dataset [24] in the AIM contest as the training and validation sets for the image demoiréing task. For image deraining and image dehazing problems, we choose the Rain1200 [6] and RESIDE [25] datasets as the training and validation sets. We train our model on all the three training set with the same augmentation in [26]. During the training phase, we crop 256×256 patches from the image as inputs and the batch size is set to 16.

3. EXPERIMENT

In this section, we evaluate the proposed MsGWN model on demoiréing, dehazing and deraining tasks. To compare with the Haar Wavelet Transform, we also build a deep Haar Wavelet Network, named **Haar** in this section, by replacing the GWM in the proposed network with the clique Haar Wavelet Module in [27]. We present quantitative and qualitative comparisons with state-of-the-art approaches. In addition, we carry out ablation studies to analyze several design choices of the proposed model.

Table 1. **Demoiréing results on the LCDMoire dataset.** **Red texts** indicate the best detection precision.

Methods		DnCNN [28]	MSFE [3]	SUN [4]	Haar	Ours
LCDMoire	PSNR	29.08	36.66	37.41	39.13	**40.02**
	SSIM	0.906	0.981	0.982	0.989	**0.992**

3.1. Performance Evaluation

We use LCDMoire dataset [24] to evaluate the performance of image demoiréing. The qualitative results of different methods are presented in Figure 2. It can be observed that all the comparison methods failed to remove the Moiré pattern in the selected region. Compared with them, our method can successfully remove the Moiré pattern with multi-scale and multi-orientation. Quantitative evaluation in terms of PSNR, SSIM in Table 1 also shows that the proposed MsGWN can obtain favorable performance.

In deraining task, we compare our model with DDN [5], JORDER [9], DID-MDN [6], RESCAN [7], NLEDN [8], DuRN-S [10] and Haar methods. As shown in Table 2, the proposed MsGWN can obtain better scores than other evaluated methods in terms of PSNR and SSIM. According to Figure 3, one can find that the DID-MDN [6] fails to reconstruct the image without the rain-streaks. Although DuRN [10] and Haar methods can remove most rain-streaks, they cannot preserve the detailed information of the complex scene. In contrast, our method can recover much better details while eliminating the effect of rain-streaks.

We take image dehazing as the third task for evaluating our model. We compare our method with DCP [11], AODNet [16], GFN [12], GCANet [13], GDN [14], DuRN [10] and Haar. As shown in Table 3, our method outperforms all the other methods in the SOTS dataset (the test subset of RESIDE dataset). We also present the visual results of dehazed image in Figure 4. It is clearly observed that the result of Haar method contains significant haze residual, while the results of DCP [11] and DuRN [10] contain color distortions in some regions. Meanwhile, our MsGWN model can generate clearer result without color distortion. It is noted that, although the distribution of haze is more uniform than Moiré pattern and rain-streaks, our method can still acquire state-of-the-art performance by effectively exploiting the contaminated features. The promising results in image dehazing demonstrate that our methods can be applied to most image restoration problems.

3.2. Ablation Study and Analysis

To demonstrate the effectiveness of the proposed Gabor Wavelet Module, we first replace the Gabor Wavelet Module with three channel-attention blocks as the baseline model of the multi-scale architecture. Then we train the Haar method mentioned in Section 3.1 to evaluate the effectiveness of feature decomposing. Moreover, we also remove the base branch in the MsGWN to illustrate the importance of the base

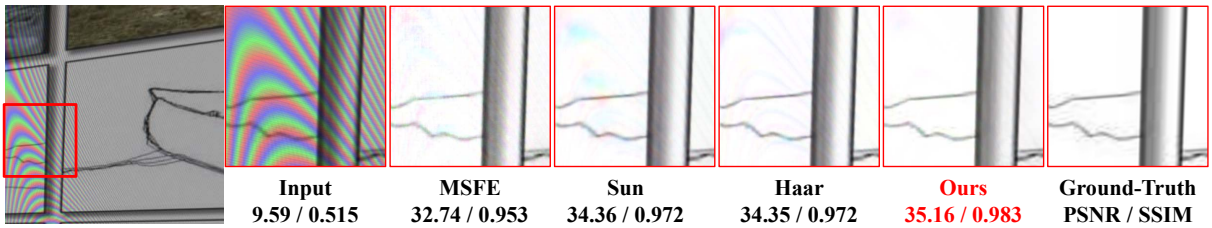

Fig. 2. Visual comparisons on the LCDMoire validation set. Red texts indicate the best detection precision.

Table 2. Deraining results on the Rain1200 dataset. Red texts indicate the best detection precision.

Methods		DDN [5]	JORDER [9]	DID-MDN [6]	RESCAN [7]	NLEDN [8]	DuRN-S [10]	Haar	Ours
Rain1200	PSNR	23.53	30.35	28.30	32.48	33.16	33.21	32.62	**33.28**
	SSIM	0.706	0.876	0.871	0.910	0.919	0.925	0.913	**0.927**

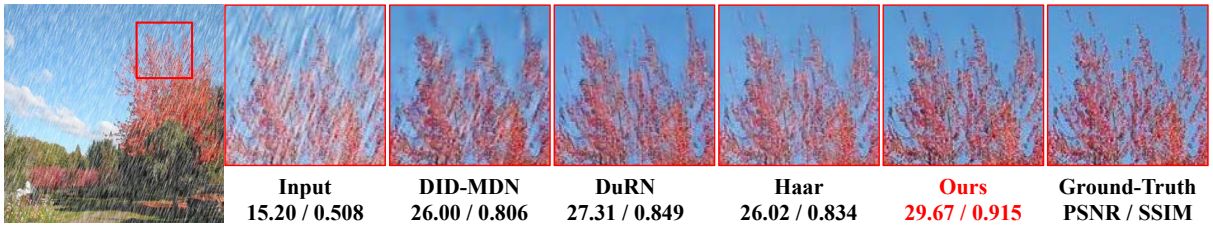

Fig. 3. Visual comparisons on the Rain1200 dataset. Red texts indicate the best detection precision.

Table 3. Dehazing results on the SOTS dataset. Red texts indicate the best detection precision.

Methods		DCP [11]	AODNet [16]	GFN [12]	GCANet [13]	GDN [14]	DuRN [10]	Haar	Ours
SOTS	PSNR	18.75	18.80	23.47	26.32	27.01	29.31	30.27	**30.64**
	SSIM	0.859	0.834	0.894	0.941	0.946	0.952	0.962	**0.966**

Fig. 4. Visual comparisons on the SOTS dataset. Red texts indicate the best detection precision.

branch. As shown in Table 4, the Haar model performs better than the baseline by decomposing the features into different components, which can facilitate the training process. However, the proposed Gabor Wavelet Module also shows superior performance than Haar model by removing the degradations at four orientations. It is also noted that removing the base branch will bring significant performance decrease in all the three applications.

Table 4. Analysis on the effect of each component in MsGWN. Red texts indicate the best detection precision.

Methods		baseline	Haar	MsGWN(w/o base)	MsGWN
LCDMoire	PSNR	37.21	39.13	37.64	**40.02**
	SSIM	0.982	0.989	0.983	**0.992**
Rain1200	PSNR	32.3	32.65	31.85	**33.28**
	SSIM	0.910	0.913	0.884	**0.927**
SOTS	PSNR	29.91	30.27	29.46	**30.64**
	SSIM	0.958	0.962	0.957	**0.966**

4. CONCLUSION

In this paper, we propose a Multi-scale Gabor Wavelet Network (MsGWN) for image restoration. By adopting the multi-scale architecture and Gabor Wavelet Module, the proposed network can decompose the contaminated features into multi-scale and multi-orientation components. Therefore, the contaminated features can be restored at different spatial scales and orientations. Extensive evaluations of different restoration tasks demonstrate that the proposed method is effective for image restoration.

Acknowledgement: This work is partially supported by National Major Science and Technology Projects of China grant under number 2019ZX01008103, National Natural Science Foundation of China (61603291), Natural Science Basic Research Plan in Shaanxi Province of China (Program No.2018JM6057), and the Fundamental Research Funds for the Central Universities.

5. REFERENCES

[1] Zhouping Wei, Jian Wang, Helen Nichol, Sheldon Wiebe, and Dean Chapman, "A median-gaussian filtering framework for moiré pattern noise removal from x-ray microscopy image," *Micron*, vol. 43, no. 2, pp. 170–176, 2012. 1

[2] Jingyu Yang, Xue Zhang, Changrui Cai, and Kun Li, "Demoiréing for screen-shot images with multi-channel layer decomposition," in *2017 IEEE Visual Communications and Image Processing (VCIP)*, 2017. 1

[3] Gao T, Guo Y, Zheng X, Wang Q, and Luo X, "Moiré pattern removal with multi-scale feature enhancing network," in *2019 IEEE International Conference on Multimedia Expo Workshops (ICMEW)*, 2019. 1, 3

[4] Yujing Sun, Yizhou Yu, and Wenping Wang, "Moiré photo restoration using multiresolution convolutional neural networks," *IEEE Transactions on Image Processing*, vol. PP, no. 99, pp. 1–1, 2018. 1, 3

[5] Xueyang Fu, Jiabin Huang, Delu Zeng, Huang Yue, Xinghao Ding, and John Paisley, "Removing rain from single images via a deep detail network," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017. 1, 3, 4

[6] He Zhang and Vishal M Patel, "Density-aware single image de-raining using a multi-stream dense network," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2018. 1, 3, 4

[7] Xia Li, Jianlong Wu, Zhouchen Lin, Hong Liu, and Hongbin Zha, "Recurrent squeeze-and-excitation context aggregation net for single image deraining," in *European Conference on Computer Vision*, 2018. 1, 3, 4

[8] Guanbin Li, Xiang He, Wei Zhang, Huiyou Chang, Le Dong, and Liang Lin, "Non-locally enhanced encoder-decoder network for single image de-raining," *arXiv*, 2018. 1, 3, 4

[9] Wenhan Yang, Robby T Tan, Jiashi Feng, Jiaying Liu, Zongming Guo, and Shuicheng Yan, "Deep joint rain detection and removal from a single image," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017. 1, 3, 4

[10] Xing Liu, Masanori Suganuma, Zhun Sun, and Takayuki Okatani, "Dual residual networks leveraging the potential of paired operations for image restoration," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 7007–7016. 1, 3, 4

[11] Kaiming He, Jian Sun, and Xiaoou Tang, "Single image haze removal using dark channel prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pp. 2341–2353, 2011. 1, 3, 4

[12] Wenqi Ren, Lin Ma, Jiawei Zhang, Jinshan Pan, Xiaochun Cao, Wei Liu, and Ming-Hsuan Yang, "Gated fusion network for single image dehazing," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2018. 1, 3, 4

[13] Dongdong Chen, Mingming He, Qingnan Fan, Jing Liao, Liheng Zhang, Dongdong Hou, Lu Yuan, and Gang Hua, "Gated context aggregation network for image dehazing and deraining," in *IEEE Winter Conference on Applications of Computer Vision*. IEEE, 2019, pp. 1375–1383. 1, 3, 4

[14] Xiaohong Liu, Yongrui Ma, Zhihao Shi, and Jun Chen, "Griddehazenet: Attention-based multi-scale network for image dehazing," *IEEE International Conference on Computer Vision*, 2019. 1, 3, 4

[15] Bolun Cai, Xiangmin Xu, Kui Jia, Chunmei Qing, and Dacheng Tao, "Dehazenet: An end-to-end system for single image haze removal," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5187–5198, 2016. 1

[16] Boyi Li, Xiulian Peng, Zhangyang Wang, Jizheng Xu, and Dan Feng, "Aod-net: All-in-one dehazing network," in *IEEE International Conference on Computer Vision*, 2017, pp. 4770–4778. 1, 3, 4

[17] Siyuan Li, Iago Breno Araujo, Wenqi Ren, Zhangyang Wang, and Xiaochun Cao, "Single image deraining: A comprehensive benchmark analysis," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2019. 1

[18] Pengju Liu, Hongzhi Zhang, Zhang Kai, Lin Liang, and Wangmeng Zuo, "Multi-level wavelet-cnn for image restoration," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018. 1

[19] Sing Lee Tai, "Image representation using 2d gabor wavelets," in *TPAMI*, 1996. 2, 3

[20] D. Gabor, "Theory of communication," *Iee Proc London*, vol. 93, no. 73, pp. 58, 1946. 2

[21] Zhang Baochang, Shan Shiguang, Chen Xilin, and Gao Wen, "Histogram of gabor phase patterns (hgpp): a novel object representation approach for face recognition," *IEEE Transactions on Image Processing*, vol. 16, no. 1, pp. 57–68, 2007. 2

[22] Yulun Zhang, Kunpeng Li, Li Kai, Lichen Wang, Bineng Zhong, and Fu Yun, "Image super-resolution using very deep residual channel attention networks," in *European Conference on Computer Vision*, 2018. 3

[23] Diederik P Kingma and Jimmy Ba, "Adam: A method for stochastic optimization," in *International Conference on Learning Representations*, 2015. 3

[24] Shanxin Yuan, Radu Timofte, Gregory Slabaugh, and Ales Leonardis, "Image super-resolution using very deep residual channel attention networks," 2019. 3

[25] Boyi Li, Wenqi Ren, Dengpan Fu, Dacheng Tao, Dan Feng, Wenjun Zeng, and Zhangyang Wang, "Reside: A benchmark for single image dehazing," *arXiv*, 2017. 3

[26] Xinyi Zhang, Hang Dong, Zhe Hu, Wei-Sheng Lai, Fei Wang, and Ming-Hsuan Yang, "Gated fusion network for joint image deblurring and super-resolution," in *British Machine Vision Conference*, 2018. 3

[27] Zhisheng Zhong, Tiancheng Shen, Yibo Yang, Chao Zhang, and Zhouchen Lin, "Joint sub-bands learning with clique structures for wavelet domain super-resolution," in *Neural Information Processing Systems*, 2018. 3

[28] 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*, vol. 26, no. 7, pp. 3142–3155, 2016. 3