This CVPR 2020 workshop paper is the Open Access version, provided by the Computer Vision Foundation.
Except for this watermark, it is identical to the accepted version;
the final published version of the proceedings is available on IEEE Xplore.

Deep Wavelet Network with Domain Adaptation for Single Image Demoireing

Xiaotong Luo[1], Jiangtao Zhang[1], Ming Hong[1], Yanyun Qu*[1], Yuan Xie*[2], Cuihua Li[1]
[1]School of Informatics, Xiamen University, Xiamen, China
[2]East China Normal University, Shanghai, China
xiaotluo@qq.com, 1328937778@qq.com, mingh@stu.xmu.edu.cn
yyqu@xmu.edu.cn, xieyuan8589@foxmail.com, chli@xmu.edu.cn

Abstract

Convolutional neural networks have made a prominent progress in low-level image restoration tasks. Moire is a kind of high-frequency and irregular interference stripe that appears on the photosensitive element of digital cameras or scanners. It can bring in unpleasant colorful artifacts on images. In this paper, we propose a deep wavelet network with domain adaptation mechanism for single image demoireing, dubbed AWUDN. The feature mapping is mainly performed in the wavelet domain, which can not only cut down computation complexity, but also reduce information loss. Moreover, considering that the images provided by the challenge organizers have strong self-similarity, the global context block is adopted for the learning of feature dependency in different positions. Finally, we introduce the domain adaptation mechanism to fine-tune the pretrained model for reducing the domain gap between training moire dataset and testing moire dataset. Benefiting from these improvements, the proposed method can achieve superior accuracy on the public testing dataset in the NTIRE 2020 Single Image Demoireing Challenge.

1. Introduction

When photographing digital images, the scene is sampled at discrete spatial locations. If the highest frequency of the scene exceeds the sampling rate of the imaging system, it could cause aliasing in frequency domain. Meanwhile, the aliasing can generate sawtooth artifacts or moire stripes on images. Actually, moire is a type of high-frequency and irregular stripe appearing on the photosensitive element of digital devices, which could lead into colorful artifacts on images. In the NTIRE 2020 Single image Demoireing Challenge Track 1, for an input image the aim is to obtain a high-quality image with the best fidelity towards the ground truth moire-free image [18].

*Corresponding author

Figure 1. The visual comparisons on the testing dataset. Left: the moire image "000019_3"; Right: our demoireing result

Low-level image restoration contains a large number of tasks, such as image super-resolution, dehazing, deblurring, denoising and so on. Most of these tasks have the corresponding degradation models to synthetize training datasets. However, moire stripes are irregular in shape. It is complex to formalize the degradation process of a moire image.

Traditional methods for single image demoireing mainly include images filtering and decomposition operations. [8] proposes a sparse matrix factorization algorithm to remove moire stripes on high-frequency textures. With the rise of deep learning, it has made a great breakthrough for image restoration both in objective indexes and subjective visual effects with supervised training. For example, [19] proposes a multi-scale network for image demoireing.

In this paper, we propose a deep wavelet network with domain adaptation for single image demoireing, dubbed AWUDN. The whole network is an U-Net structure, where the downsampling and upsampling of feature maps are replaced with discrete wavelet transform (DWT) and inverse discrete wavelet transform (IDWT) for reducing computation complexity and information loss. Therefore, the feature mapping is performed in wavelet domain, where the basic block adopts the residual-in-residual structure [20] for extracting more residual information effectively. Considering that the dataset provided by the competition has strong self-similarity, i.e., similar texture structure inside image itself, the global context block is introduced in the fron-

t of network structure. It can help establish the relationship between two distant pixels to better use the internal information of the image for restoring texture details. Moreover, there may exist slight domain difference between the source domain training data and the target domain testing data. It means that the distribution of moire images in the training set and the moire images in the test set is inconsistent, which can constrain the performance improvement of the model pretrained on the training dataset. CORAL loss [12] gives us some inspiration, which defines the measurement difference of the second-order statistics between features in the source domain and the target domain. Therefore, the pretrained model WUDN is fine-tuned using coral loss for reducing the domain shift of training dataset and testing dataset in the testing phase. Figure 1 shows the visual comparisons of the moire image and our demoireing result on the test dataset, and it seems that our proposal can remove the moire well.

To summarize, the main contributions are four-fold:

- A deep wavelet network with domain adaptation mechanism, dubbed AWUDN, is proposed for single image demoireing.
- The feature mapping is performed in wavelet domain, where DWT and IDWT operations are used for feature downsampling and upsampling, respectively. It can reduce computation complexity and information loss.
- Global context block is embedded in the front of network for modeling self-similarity dependency. It can help to mine the internal information of the image to restore texture details.
- Domain adaptation is introduced to fine-tune the pretrained model for reducing the domain difference of the training dataset and testing dataset.

2. Related work

Image restoration aims to recover a high-quality image from its degraded counterpart. In general, there exists the corresponding degradation model for most tasks, such as image super-resolution, dehazing and so on. These physical models can be utilized as prior information during the reconstruction process. Moire is a kind of irregular stripes, which may appear in different regions. Therefore, it is difficult to formalize the degradation expression for a moire image. Here, we focus on the works referred to traditional demoireing methods and deep learning based methods.

2.1. Traditional demoireing methods

Traditional moire removal methods are generally based on images filtering and decomposition. Nishioka *et al.* [9] proposed that adding a low-pass filter to the camera can effectively remove moire. K.Pandya *et al.* [1] proposed a method for removing moire in the frequency domain. Wei *et al.* [15] proposed a median-Gaussian filtering method to eliminate ripples in X-ray microscope images. Yang *et al.* [17] proposed a novel textured image demoireing method by signal decomposition and guided filtering. Liu *et al.* [8] proposed a low-level sparse matrix factorization to remove moire on high-frequency textures.

2.2. Deep learning based methods

With the prevalence of deep learning (DL), abundant algorithms for low-level tasks have been proposed, such as image super-resolution and image dehazing. At the same time, deep convolutional neural network has also been applied to the problem of image demoireing. Sun *et al.* [13] first introduced the demoireing problem on camera images and proposed that the moire characteristics of camera images are muti-frequency. Moreover, they also proposed a multi-scale network, which can achieve significant visual effects. However, it fails on large-scale color patches. Subsequently, Gao *et al.* [5] proposed a multi-scale feature enhancement network (MSFE), and applied feature enhancement branch (FEB) with skip connections. Later, Cheng *et al.* [3] also adopted a multi-scale strategy to introduce the attention mechanism to the moire removal, and proposed Channel attention Dynamic feature encoding Residual block (CDR) and Dynamic feature encoding (DFE). However, this method has insufficient information interaction ability between different scales. He *et al.* [6] proposed a multi-scale network based on DenseNet, and added a RGB channel level edge detector to solve the unbalanced intensity between color channels, which can achieve a remarkable visual effect.

In summary, the traditional demoireing methods are more interpretable, but the results are not satisfactory. Existing deep learning methods can achieve preferable performance and fast inference speed, while limiting on the design and improvement of multi-scale network structures.

3. Proposed Method

In this section, we first present the overall network architecture. Then, the loss function is defined to optimize the model. Next, we describe the adaptive fine-tuning and optimization scheme. Finally, we discuss the differences between the proposed network and its related works.

3.1. Network architecture

The proposed WUDN mainly includes three parts: the shallow feature extraction module, the global context block, the wavelet encoder-decoder module as shown in Figure 2. Here, the input moire image and the output demoireing image are denoted as X and Y, respectively.

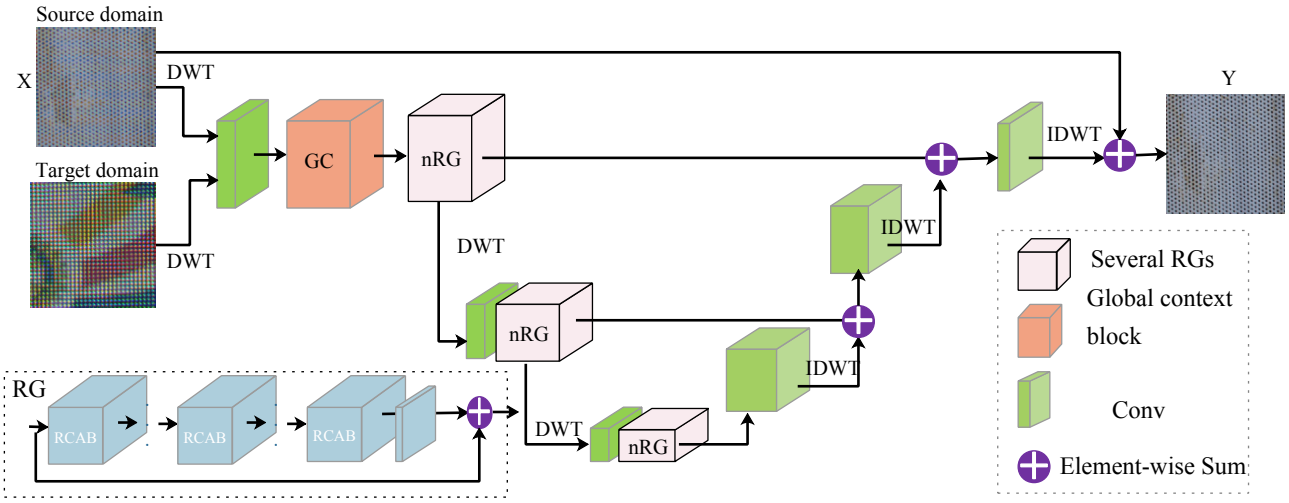

Figure 2. The overall network architecture of the proposed AWUDN

Firstly, discrete wavelet transform (DWT) is performed to the input moire image X, which can turn the RGB image into twelve-channel image X_{dwt} and reduce the image size by half.

$$X_{dwt} = DWT(X). \tag{1}$$

Then, the shallow features F_0 are obtained by the shallow feature extraction module, in which a 3×3 convolutional layer without activation is applied to the moire input in wavelet domain, i.e., X_{dwt}.

$$F_0 = G_0(X_{dwt}), \tag{2}$$

where $G_0(\cdot)$ denotes the shallow filter operation.

Next, global context block is followed to capture the dependence between long-distance pixels.

$$F_{GC} = G_{GC}(F_0), \tag{3}$$

where G_{GC} represents the global context extraction function. F_{GC} is the obtained features.

The wavelet encoder-decoder module is followed for realizing deep feature mapping, which is an U-Net structure. The basic block consists of several residual groups (RGs) that are a residual-in-residual structure as proposed in [20]. The feature maps F_{GC} go through several residual groups at different scales sequentially. These scales are generated by DWT downsampling. Then, the size of features F_{GC} is reduced to one quarter, and feature upsampling is performed by inverse discrete wavelet transform (IDWT). Meanwhile, the feature maps with the same spacial size are element-wise summed. Note that the convolutional layer after IDWT is used to the alignment of channels.

$$X_{ED} = G_{ED}(F_{GC}), \tag{4}$$

where $G_{ED}(\cdot)$ indicates mapping function of the wavelet encoder-decoder module, and X_{ED} represents the reconstructed residual image.

Finally, the reconstructed residual image adds with the input moire image for generating the moire-free image Y:

$$Y = X_{ED} + X. \tag{5}$$

3.2. Global context block

After glancing over the image pairs provided by the challenge organizers, it can be found that the internal image itself has a great degree of self-similarity. Considering that the convolution operation only covers local information of an image, He *et al.* proposed the non-local module [14] drawing on the non-local mean algorithm. It is used to capture the dependence between long-distance pixels, so that the features at the query point position are calculated by weighting and summing the features at all positions on the input features. The non-local module can establish the relationship between two distant pixels on an image, which can be used to better use the internal information of the image for texture details restoration. Cao *et al.* [2] analyzed the characteristics of the non-local module and found that the features of the image patches obtained at different query positions are the same. It means that the input region has no relationship with the calculated features. Therefore, they combined with the characteristics of the global context framework, non-local module and SE block [7], and then proposed the Global Context (GC) block as shown in Figure 3. The GC block can greatly simplify the calculation complexity of the non-local module without decreasing the performance.

GC block can be regarded as a global context content modeling module, which forms the global context characteristics of a specific query at each query position and cal-

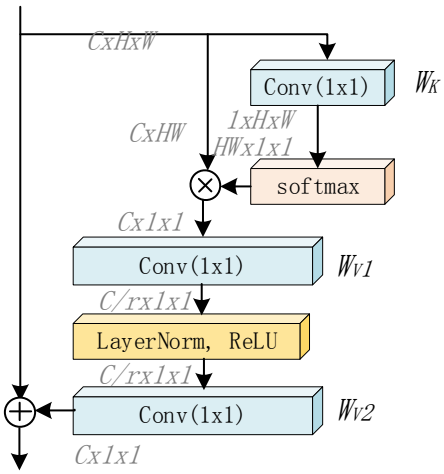

Figure 3. Global context block

culates channel attention for each query point. The computation complexity is the second power of the number of position points. The GC block is obtained by adding the attention weight ω_j of features, the conversion function $\Psi(\cdot)$, and the global context features at each location. The features of the query point z_i can be denoted as

$$z_i = x_i + \Psi(\sum_{j=1}^{N_p} \alpha_j x_j)\Psi(\cdot) = W_{v2}\Big(ReLU\Big(LN\big(W_{v1}(.)\big)\Big)\Big) \tag{6}$$

where ReLU is a nonlinear activation operation. LN refers to Layer Normalization. x_i is a query point. x_j is another point related to the query point. W_{v1}, W_{v2}, and W_k represent the weight values of the three convolutional operations, respectively. $\sum \alpha_j x_j$ is the context modeling operation, which uses the weight average α_j to group the features of all positions together to obtain a global context feature. For more details, please refer to [2].

3.3. Loss function

In learning-based image restoration methods, Mean Absolute Error (MAE) and Mean Squared Error (MSE) on image pixels space are frequently used. In this proposal, we also adopt MAE loss for measuring the differences between the demoireing images and the ground truth.

Meanwhile, inspired by [16], which introduces loss in Fourier domain, we adopt another loss in the discrete cosine transform (DCT) domain. After an image signal is transformed into the DCT domain, the main components of the frequency coefficients are concentrated in a relatively small range. The distribution of frequency coefficients presents a zigzag arrangement. The upper left corner represents the DC component and the low frequency components (i.e., profile) of the image information. While the values in lower right corner are small, they mainly reflects the high frequency (i.e., details) of the image. As shown in Figure 4, the moire images have less low-frequency components but more high-frequency information than the moire-free images. Considering the similarity in the frequency domain can be mutually reinforced with the spatial domain, we combine the two loss functions as follows,

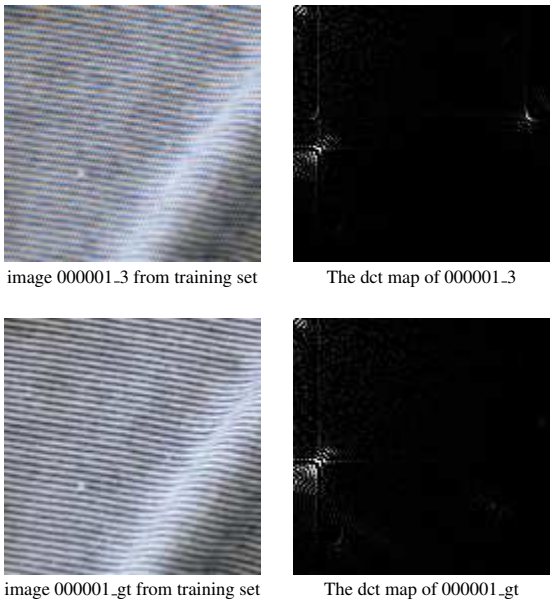

Figure 4. The moire image and the corresponding moire-free image with the respective dct map. **Zoom in for best view**

$$L = L_{pix} + \lambda L_{dct} \tag{7}$$

$$L_{pix} = \frac{1}{N}\sum_{i=1}^{N} \left\| x_i^{gt} - G(x_i^m) \right\|_1 \tag{8}$$

$$L_{dct} = \frac{1}{N}\sum_{i=1}^{N} \left\| D(x_i^{gt}) - D(G(x_i^m)) \right\|_1 \tag{9}$$

where x_i^{gt}, x_i^m denote the ith ground truth sample and moire sample. $D(\cdot)$ represents the discrete cosine transform. N is the total number of training samples. We will discuss the loss function in the Section 4.3.

3.4. Adaptive Fine-tuning

Considering that there may exist slight distribution difference between the moire images in the training dataset and the moire images in the testing dataset, it would limit the performance of the model pretrained on the training dataset. Inspired by Deep CORAL [12], we introduce domain adaptation to fine-tune the pretrained model, dubbed AWUDN, for reducing the domain shift of training dataset and testing dataset in the testing phase. As shown in Figure 2, we add CORAL loss between the source domain features and the target domain features to reduce their second-order

statistical characteristics. Note that the compared features are extracted after the global context block.

The CORAL loss between a single feature layer in two different domains can be denoted as follows. Assuming that the source domain moire and moire-free training sample pairs $D_S = (x, y)$, the target domain test moire samples $D_T = u$. Here, x and u are the same dimension. The expression of CORAL loss is shown as

$$L_{Coral} = \frac{\|C_S - C_T\|_F^2}{4d^2}, \tag{10}$$

where $\|\cdot\|_F^2$ represents the square matrix $Frobenius$ norm. d is the number of feature channels. C_S, C_T are the feature covariance matrixes of the source and target data, and can be expressed as

$$C_S = \frac{1}{n_S - 1}(F_S^T F_S - \frac{(1^T F_S)^T (1^T F_S)}{n_S}), \tag{11}$$

$$C_T = \frac{1}{n_T - 1}(F_T^T F_T - \frac{(1^T F_T)^T (1^T F_T)}{n_T}), \tag{12}$$

where $\mathbf{1}$ represents the column vector whose elements are all 1. F_S and F_T denote the features of the moire image in the source domain and the moire image in the target domain. The number of samples of source data and target data are n_S and n_T, respectively. Therefore, we introduce CORAL loss in network fine-tuning to map the features of the moire images in the training set and the moire images in the testing set to a common feature subspace for reducing the domain differences between them.

$$L_{ft} = L + \gamma L_{Coral}, \tag{13}$$

where L represents the loss defined in Section 3.3, and γ is the weight parameter.

3.5. Optimization scheme

The whole optimization process is shown in Algorithm 1. We first optimize the network with paired dataset by Eq. (7). Next, the pretrained model WUDN is loaded for fine-tuning with paired training dataset and moire test dataset without the ground truth. The paired samples in the source domain and the moire samples in the target domain are passed through the global context block to obtain features for the measurement of the second-order statistical characteristics. The total optimization loss of the network during fine-tuning is shown by Eq. (13).

3.6. Discussions

In this subsection, we mainly discuss the differences between the proposed method and its related methods.

Algorithm 1 Adaptive fine-tuning for single image demoireing.

1: **Input**: Paired training samples D_S in source domain, only test moire samples D_T, batch size m, total iteration number T, the current iteration number$i$
2: Load the pretrained model parameters of WUDN optimized by the loss defined in Eq. (7).
3: **while** $i <= T$ **do**
4: Sample paired data $(x_i, y_i)_{i=1}^m$ from D_S and moire data $(x_i)_{i=1}^m$ from D_T;
5: Input x_i, u_i for extracting the features for optimizing CORAL loss;
6: Update AWUDN by minimizing the objective: $L_{ft} = L + \lambda L_{Coral}$;
7: $i \leftarrow i + 1$;
8: **end while**

WUDN vs. RCAN. In WUDN, we also adopt the residual-in-residual structure with channel attention proposed in RCAN [20] as basic feature mapping block. For further reducing computation complexity, we synthesize DWT and IDWT for decreasing the size of feature maps rather than the same spatial size in the network pipeline.

WUDN vs. U-Net. WUDN adopts the U-Net structure. The difference lies in: DWT as a kind of downsampling way in WUDN to reduce the size of feature maps instead of a convolution with a stride of 2 in U-Net. Meanwhile, IDWT is utilized for feature upsampling rather than sub-pixel convolution [11], deconvolution [4], or nearest-neighbor upsampling + convolution [10].

Besides, we also add the global context block considering the self-similarity characteristic of given moire images. In addition, we fine-tune the pretrained model using coral loss to compensate for the performance caused by the difference in the data domain.

4. Experiments

4.1. Datasets

The dataset provided by the NTIRE 2020 Single Image Demoireing Challenge contains 11000 diverse image pairs totally, where 10000 image pairs are used for training, 500 for validation, 500 for testing. However, the ground truth for the validation dataset and the testing dataset has not been released, so that we can only get online feedback from the validation server. During training, we use 9950 image pairs for training and the last 50 image pairs for validation. The size of image pairs is all 128×128.

4.2. Implementation details

During training, the whole 128×128 RGB images are as input to WUDN. Random flipping and rotation are used to

augment the training data. Adam optimizer with $\beta_1 = 0.9$, $\beta_2 = 0.999$ is utilized to optimize the proposed network. The mini-batch size is set to 16. The initial learning rate is set to $2e-4$ and decreased to half per 200 epochs for a total of 1000 epochs. The kernel size of all convolutional layers is set to 3×3 and the activation functions are all ReLU. We use PyTorch to realize the network. The hardware configuration is GTX Titan XP. We train the proposed method for approximately 3 days.

In WUDN, the initial convolutional layer contains 32 kernels. After the second DWT, the kernels are changed to 128, and then changed to 512 after the third DWT. The residual groups in each nRG is set to 5 and residual blocks are also set to 5. γ is set to 100.

Model performance is evaluated by objective index, *i.e.*, peak signal-to-noise ratio (PSNR) as Eq. (14).

$$PSNR = 10log_{10}(\frac{(2^n-1)^2}{MSE}). \tag{14}$$

4.3. Ablation analysis

In this subsection, we discuss the effects of wavelet domain U-Net, global context block, DCT loss function and adaptive fine-tuning. The baseline is a U-Net structure without global context block. Meanwhile, the convolution with a stride of 2 is used for feature downsampling and sub-pixel convolution is used for feature upsampling. Note that the PSNR values are obtained by the online server.

Table 1. Ablation studies of the effects of wavelet domain U-Net, Global context block (GC) in PSNR (dB) on the validation dataset.

Model	U-Net	WU-Net	WU-Net + GC
PSNR	40.73	40.91	**41.68**

Wavelet domain U-Net (WU-Net). To validate the effect of feature mapping in wavelet domain, i.e., DWT and IDWT operations for changing the size of feature maps, we compare it with convolution with stride of 2 and sub-pixel convolution upsampling. The DWT and IDWT operations are orthogonal transform, which can reduce information loss. As shown in Table 1, the PNSR is improved from 40.73 dB to 40.91 dB.

Global context block (GC). To validate the effect of global context block, we add it upon the wavelet domain U-Net. As shown in Table 1, the PNSR is improved from 40.91 dB to 41.68 dB. It indicates that global context block can help to mine the similarity of internal image and restore more rich textures.

DCT loss. To fully employ the information of an image, we introduce the loss in DCT domain. As shown in Figure 5, after adding the DCT loss with L_1, the validation curve is superior to only use L_1 loss. The loss of image domain

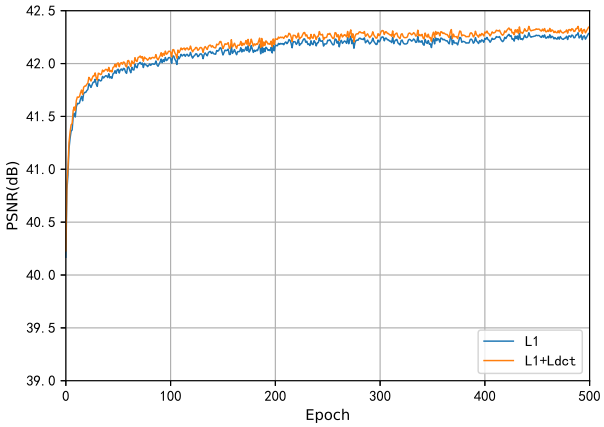

Figure 5. The convergence curves with L_1 loss and DCT loss on the last 50 image pairs of the given training dataset

is a global measure which is hard to restrict each pixel on the image. The high-frequency details distribute differently in the image and frequency domain. The loss of the DCT domain can help further constrain the similarity of the high-frequency detail. Note that the models in Table 1 are all trained by L_1 loss.

Fine-tuning. To reduce the domain difference, we fine-tune the pretrained model WUDN with CORAL loss for 50 epochs. As shown in Table 2, the PSNR is increased from 41.85 dB to 41.94 dB. It shows that domain adaptation can reduce the domain shift of the training dataset and testing dataset and further improve performance.

Table 2. The quantitative comparisons between WUDN and AWUDN in PSNR(dB) on the validation dataset

Model	WUDN	AWUDN
PSNR	41.85	**41.94**

4.4. Comparisons with the state-of-the-arts

In this section, we compare our proposal with the state-of-the-art methods RCAN[20], which is a very deep residual network for image super-resolution with the residual-in-residual structure embedded with the channel-wise attention mechanism. We retrain RCAN for image demoireing and remove its upsampling module. The quantitative results are shown in Table 3. AWUDN+ represents the results of using self-ensemble strategy, which can further improve performance as a post-processing methods. AWUDN can outperform RCAN by 0.23 dB. Moreover, AWUDN+ can achieve the best performance for 42.22 dB. Although the number of parameters of AWUDN is larger than RCAN, the amount of calculation is smaller than that of RCAN.

The quantitative results on the last 50 image pairs in the training dataset are shown in Figure 6, which can recover moire-free image similar to the ground truth. Moreover, we

Table 3. The quantitative comparisons between RCAN and AWUDN in PSNR(dB) on the validation dataset.

Model	RCAN	AWUDN	AWUDN+
PSNR	41.71	41.94	**42.22**
Params(M)	15.29	144.33	144.33
GFlops	248.93	45.17	45.17

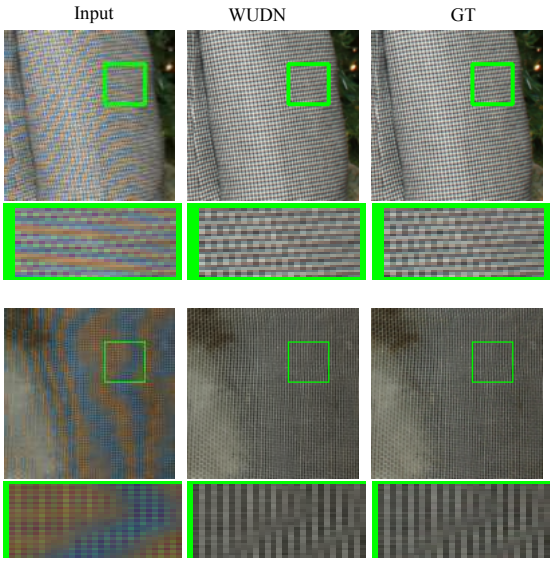

Figure 6. Visual comparisons of our proposal on the last 50 images of the training dataset. **Zoom in for best view**

also give several example on the validation dataset in Figure 7. The proposed method can obtain significant quantitative and qualitative results.

4.5. Future work

In the proposed network, we perform feature mapping mainly in wavelet domain. However, the information of different frequencies is treated equally. Considering moire is a type of high-frequency and irregular interference stripe, the information of different frequencies after discrete wavelet transform should be treated differently. We will consider improving it in future work.

5. Conclusions

In this paper, we propose a deep wavelet network with domain adaptation for single image demoireing, named AWUDN. The whole network is an U-Net structure, where the downsampling and upsampling of features are replaced with DWT and IDWT for reducing information loss and computation complexity. The basic block of the proposed method adopts residual-in-residual structure for extracting more texture information. Besides, considering the self-similarity inside the given moire image, we add the global context block in the network structure for the learning of

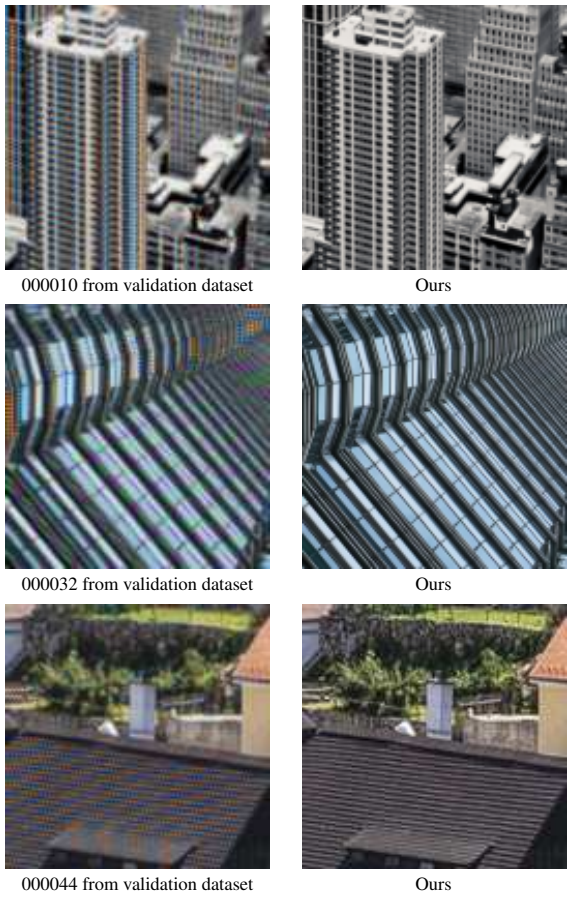

Figure 7. Visual comparisons of our proposal on validation dataset. **Zoom in for best view**

dependence between long-distance pixels. Finally, we finetune the pretrained model with domain adaptation mechanism for reducing the domain shift of training dataset and testing dataset in the testing phase.

6. Acknowledgments

This work is supported by the National Natural Science Foundation of China under Grant 61876161, Grant 61772524, Grant U1065252 and partly by the Beijing Municipal Natural Science Foundation under Grant 4182067, and partly by the Fundamental Research Funds for the Central Universities associated with Shanghai Key Laboratory. of Trustworthy Computing

References

[1] Igor N Aizenberg and Constantine Butakoff. Frequency domain medianlike filter for periodic and quasi-periodic noise removal. In *Image Processing: Algorithms and Systems*, 2002.

[2] Yue Cao, Jiarui Xu, Stephen Lin, Fangyun Wei, and Han Hu. Gcnet: Non-local networks meet squeeze-excitation net-

works and beyond. *arXiv:1904.11492*, 2019.

[3] Xi Cheng, Zhenyong Fu, and Jian Yang. Multi-scale dynamic feature encoding network for image demoiréing. *arXiv:1909.11947*, 2019.

[4] Chao Dong, Chen Change Loy, and Xiaoou Tang. Accelerating the super-resolution convolutional neural network. In *ECCV*, 2016.

[5] Tianyu Gao, Yanqing Guo, Xin Zheng, Qianyu Wang, and Xiangyang Luo. Moiré pattern removal with multi-scale feature enhancing network. In *ICMEW*, 2019.

[6] Bin He, Ce Wang, Boxin Shi, and Ling-Yu Duan. Mop moire patterns using mopnet. In *ICCV*, 2019.

[7] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *CVPR*, 2018.

[8] Fanglei Liu, Jingyu Yang, and Huanjing Yue. Moiré pattern removal from texture images via low-rank and sparse matrix decomposition. In *Visual Communications and Image Processing (VCIP)*, 2015.

[9] Kimihiko Nishioka, Naoki Hasegawa, Katsuya Ono, and Yutaka Tatsuno. Endoscope system provided with low-pass filter for moire removal, 2000.

[10] Mehdi SM Sajjadi, Bernhard Schölkopf, and Michael Hirsch. Enhancenet: Single image super-resolution through automated texture synthesis. In *ICCV*, 2017.

[11] Wenzhe Shi, Jose Caballero, Ferenc Huszár, Johannes Totz, Andrew P Aitken, Rob Bishop, Daniel Rueckert, and Zehan Wang. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In *CVPR*, 2016.

[12] Baochen Sun and Kate Saenko. Deep coral: Correlation alignment for deep domain adaptation. In *ECCV*, 2016.

[13] Yujing Sun, Yizhou Yu, and Wenping Wang. Moiré photo restoration using multiresolution convolutional neural networks. *IEEE Transactions on Image Processing*, 2018.

[14] Xiaolong Wang, Ross Girshick, Abhinav Gupta, and Kaiming He. Non-local neural networks. In *CVPR*, 2018.

[15] 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*, 2012.

[16] Guang Yang. Dagan: deep de-aliasing generative adversarial networks for fast compressed sensing mri reconstruction. *IEEE Transactions on Medical Imaging*, 2017.

[17] Jingyu Yang, Fanglei Liu, Huanjing Yue, Xiaomei Fu, Chunping Hou, and Feng Wu. Textured image demoiréing via signal decomposition and guided filtering. *IEEE Transactions on Image Processing*, 2017.

[18] Shanxin Yuan, Radu Timofte, Ales Leonardis, Gregory Slabaugh, et al. Ntire 2020 challenge on image demoireing: Methods and results. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2020.

[19] Shanxin Yuan, Radu Timofte, Gregory Slabaugh, Ales Leonardis, Bolun Zheng, Xin Ye, Xiang Tian, Yaowu Chen, Xi Cheng, Zhenyong Fu, et al. Aim 2019 challenge on image demoireing: Methods and results. In *ICCVW*, 2019.

[20] Yulun Zhang, Kunpeng Li, Kai Li, Lichen Wang, Bineng Zhong, and Yun Fu. Image super-resolution using very deep residual channel attention networks. In *ECCV*, 2018.