Dual-domain Deep Convolutional Neural Networks for Image Demoireing

An Gia Vien, Hyunkook Park, and Chul Lee
Department of Multimedia Engineering
Dongguk University, Seoul, Korea
viengiaan@mme.dongguk.edu, hyunkook@mme.dongguk.edu, chullee@dongguk.edu

Abstract

We develop deep convolutional neural networks (CNNs) for moiré artifacts removal by exploiting the complex properties of moiré patterns in multiple complementary domains, i.e., the pixel and frequency domains. In the pixel domain, we employ multi-scale features to remove the moiré artifacts associated with specific frequency bands using multi-resolution feature maps. In the frequency domain, we design a network that processes discrete cosine transform (DCT) coefficients to remove moiré artifacts. Next, we develop a dynamic filter generation network that learns dynamic blending filters. Finally, the results from the pixel and frequency domains are combined using the blending filters to yield moiré-free images. In addition, we extend the proposed approach to arbitrary-length burst image demoireing. Specifically, we develop a new attention network to effectively extract useful information from each image in the burst and align them with the reference image. We demonstrate the effectiveness of the proposed demoireing algorithm by evaluating on the test set in the NTIRE 2020 Demoireing Challenge: Track 1 (Single image) and Track 2 (Burst).

1. Introduction

Moiré patterns occur in images captured by digital cameras when the subject contains repetitive details that exceed the resolution of the camera sensor [32]. The captured images contain strange-looking patterns called moiré artifacts. Moiré patterns have various and complex shapes; they can appear as stripes, curves, and ripples. Furthermore, moiré patterns overlap different color variations superposed onto the images. Moiré artifacts cause significant degradation to the visual quality of the images and the performance of subsequent image processing and computer vision applications. Thus, it is crucial to develop effective demoireing algorithms to remove moiré artifacts, and many of these have recently been proposed.

The most common classical approach to removing moiré patterns is to add an optical low-pass filter to the camera lens for anti-aliasing [25]. However, this approach may cause over-smoothing due to the loss of high-frequency components. Another approach is employing color filter array subsampling [24] based on the gradients of color difference interpolation. However, this approach has a high computational complexity which renders it unsuitable for practical applications, and its output quality relies heavily on the green channel. Recently, a signal processing-based approach [32] that explores low-rank and sparsity constraints for moiré pattern removal in the frequency domain was developed. However, it may fail in regions with complex moiré patterns.

Recent works have shown that deep learning-based approaches are more effective than model-based algorithms. For example, Sun *et al*. [28] proposed a method for modeling moiré patterns by learning from a huge dataset. Even though they provided better results than model-based approaches, their network may yield poor results when test images are taken with different camera settings from their training data. He *et al*. [13] developed a neural network for moiré pattern removal by investigating multiple properties of the moiré patterns in the pixel domain. More recently, in the AIM 2019 Demoireing Challenge [34], several deep learning-based approaches to remove moiré artifacts for images captured the monitors have been proposed [6]. However, these approaches still have difficulties in removing severe moiré artifacts or those with strong color textures due to the lack of accurate models of moire patterns.

In this work, we develop deep convolutional neural networks (CNNs) to remove moiré artifacts in images in multiple complementary domains, specifically in the pixel and frequency domains. The proposed network is composed of three subnetworks: the *pixel network*, *frequency network*, and *fusion network*. First, the pixel network converts images with moiré patterns into feature maps and processes these features at different resolutions, as a moiré pattern

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. NRF-2019R1A2C4069806).

spans a wide range of frequencies. Second, inspired by recent observations on moiré patterns [32, 13], moiré artifacts are removed in the frequency domain using the discrete cosine transform (DCT). The frequency network processes DCT coefficients to remove moiré artifacts in the frequency domain. Subsequently, the dynamic filter generation network learns the dynamic blending filters. Finally, the outputs of the pixel network and the frequency network are combined by the dynamic blending filters to generate a moiré-free image. Additionally, we extend the proposed network to burst-image demoireing, which removes moiré artifacts in multiple images with different geometric transformations of the same scene. We demonstrate the effectiveness of the proposed demoireing algorithm through the NTIRE 2020 Demoireing Challenge [33]. We achieved an average PSNR of 38.28 dB for Track 1 (Single image) and 38.50 dB for Track 2 (Burst).

The remainder of this paper is organized as follows. Section 2 provides an overview of related works. Section 3 describes the proposed algorithm. Section 4 discusses the experimental results. Finally, Section 5 concludes the paper.

2. Related Works

Moiré pattern removal: Moiré patterns are a common degradation that occurs in images captured by conventional cameras because of the interference between the frequency of textures in images or display screens and camera sensors. Several algorithms have been proposed to remove different types of moiré patterns. For example, Yang *et al.* [32] proposed low-rank constraint and sparse matrix decomposition to remove moiré patterns in high-frequency textures by observing and analyzing the moiré patterns of textures in the frequency domain. Recently, several deep learning-based demoireing algorithms have been developed and have shown to be more effective than conventional model-based algorithms. Sun *et al*. [28] exploited intrinsic correlations between moiré patterns and image components at different levels in a multi-resolution pyramid network. He *et al.* [13] proposed a framework to remove moiré patterns by considering three components: a multi-scale feature aggregation in the pixel domain, an edge predictor to estimate the edge map of moiré-free images, and an appearance classification to classify moiré patterns using multiple appearance labels. Recently, in the AIM 2019 Demoireing Challenge [34], several deep network architectures were proposed [6]. These networks employed state-of-the-art blocks and modules that have been applied to image restoration to remove moiré patterns.

Image restorations: Image restoration generally focuses on noise removal, contrast enhancement, or high-frequency detail reconstruction. Deep learning models have been successfully applied to image restoration tasks, including super-resolution [10, 16, 36, 18], denoising [37, 19], deblurring [2], dehazing [4, 35], and compression artifact reduction [9, 5]. These learning-based algorithms have achieved state-of-the-art performance in image quality improvement. It was shown that a block or module developed for a certain image restoration task also performs well in other restoration tasks [38].

Demoireing can be considered as image restoration, as it attempts to reconstruct a clean image by removing moiré artifacts. Among the state-of-the-art modules and blocks in image restoration, the dense block (DB) [14, 29, 35] and residual dense block (RDB) [38] are most closely related to demoireing. DB shows effectiveness in super-resolution by preserving low-level information to reconstruct high-frequency details, while RDB is an extension of DB that extracts abundant local features via densely connected convolutional layers and avoids gradient vanishing in a deep network. Because our goal is not only to remove moiré patterns but also to reconstruct missing information, DB and RDB are important and relevant modules necessary for this work. However, because moiré patterns are complex and difficult to distinguish from texture and color in images, the straightforward adoption of DB and RDB modules may not provide a high performance.

Attention mechanisms in deep learning: Attention mechanisms facilitate deep neural networks to determine where to focus and improve the representation of interest. Recently, attention mechanisms have been shown to be a critical component in deep learning and have been extensively employed in computer vision [39, 21, 1, 11]. Among the many variations of the attention module, the convolutional block attention module (CBAM) [30] showed efficacy in image denoising [3, 27] and super-resolution [8, 7] because it directs the network to focus on essential features and suppresses unnecessary ones. Therefore, we employ CBAM in the proposed network.

Additionally, there is an approach to incorporate attention processing to allow the models to identify misaligned regions before merging the features [31]. By determining misaligned image regions at an early stage of the network, the algorithm yields high-quality results with less artifacts than conventional algorithms. Thus, we employ an attention module in [31] for burst-image demoireing to avoid misaligned features before moiré pattern removal.

3. Proposed Algorithms

3.1. Dual-domain Network

We design a dual-domain network to effectively remove moiré artifacts and generate a high-quality clean image. Figure 1 shows the architecture of the proposed network, which takes the moiré image as input and then reconstructs a clean image. Specifically, the proposed dual-domain net-

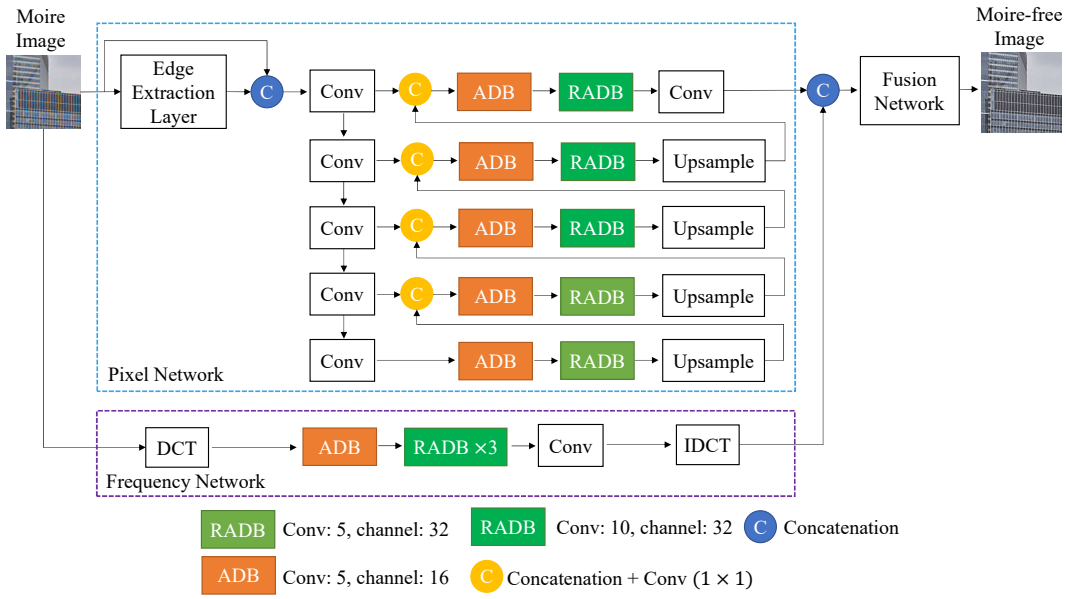

Figure 1: Architecture of the proposed dual-domain network.

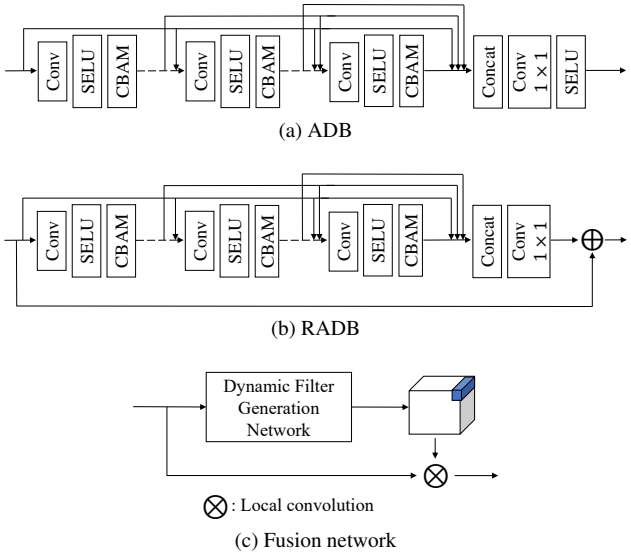

Figure 2: Architectures of the ADB, RADB, and fusion network.

work is composed of three subnetworks: the pixel network, frequency network, and fusion network. First, we estimate the individual results produced by the pixel network and the frequency network. Subsequently, the fusion network uses these results to generate moiré-free images by learning the dynamic filters.

Pixel network: The pixel network processes an input image in the pixel value domain. Based on the recent observation that multi-scale contextual information is effective in image restoration [6, 12], we use multi-scale features in the pixel value domain. In addition, assuming that the moiré artifacts have global structures in images, we first extract edge maps using the edge extraction layer and concatenate them with the inputs. The pixel network is composed of multiple branches of different resolutions. The branch at the top-level processes feature maps of the original resolution of the input image, while the other branches process coarser feature maps. The first convolutional layer (Conv) with a kernel size of 2×2 and a stride of 2 in each branch is responsible for downsampling the feature map from the higher-level branch by a factor of 2. By converting the input image into multiple feature maps at different resolutions, we exploit different levels of details from the input images. At each branch, the output feature maps from the first layer are fed into a sequence of attention dense block (ADB) and residual attention dense blocks (RADB) as shown in Figures 2(a) and (b), respectively, which are composed of CBAM [30], DB [14, 29], and RDB [38]. We increase the resolution of the feature map at each branch using the upsampling module, which is a combination of a single convolutional layer and pixel shuffle [26] and then concatenate a feature map with that of the finer branch. Finally, at the end of the top branch, a convolutional layer is used to generate the final output image.

Frequency network: Moiré patterns are complex in terms of distributions in the frequency domain [32]. According to an observation in [13], images with moiré patterns are indistinguishable from clean images, as moiré patterns are spread across a wide range of frequency bands. Therefore, exploring the properties of moiré patterns in the frequency

domain is necessary for their efficient removal. Thus, we develop an additional subnetwork to process the DCT coefficients of the input images to remove moiré artifacts in the frequency domain. The network is built by cascading a single ADB and three RADBs, shown in Figures 2(a) and (b), respectively, and a convolutional layer. The final output image is obtained by applying the inverse discrete cosine transform (IDCT).

Fusion network: We obtain the moiré-free image by combining the two images obtained from the pixel network and the frequency network, as shown in Figure 1. Since these two images have different characteristics, they are used as complementary candidates of the moiré-free image. A common approach to yield the final result from the two candidates is to use a convolutional layer with a kernel size of 1×1 for pixel-wise blending. However, this approach may cause the final image to retain the artifacts if either network fails to accurately remove them. To alleviate this issue, we instead employ a dynamic filter generation network [15] that takes the aforementioned pair of candidates as input and outputs local blending filters. These filters are then used to yield the moiré-free image. Figure 2(c) shows the proposed fusion network using dynamic local blending filters. The coefficients of the filters are learned through a dynamic blending filter network [15].

3.2. Extension to Burst-Image Demoireing

We extend the dual-domain network in Section 3.1 to burst-image demoireing. In burst-image demoireing, a set of images captured of the same target, where each image has a different geometric transformation, is considered. While each image contains different moiré patterns, they still retain pieces of useful information about the underlying clean image. This additional information in the sequence should be exploited for the effective removal of moiré artifacts. To this end, we add an additional subnetwork in the pixel network, *i.e.*, the attention network, for feature extraction and alignment, as shown in Figure 3. The architecture of the proposed attention network is illustrated in Figure 4.

Attention network: The advantage of burst images over a single image is the redundant information across the images. To fully exploit this advantage, we develop an attention network composed of global feature extraction, feature alignment, and feature merging. Global feature extraction is the first global module [22] in the attention network. The key design goal is to capture the variations between each image and generate global feature maps that contain additional information from the burst that can be combined effectively. Additionally, to ensure that the additional information can directly influence the moiré pattern removal, we fuse the global features with its input features using a convolutional layer. Second, feature alignment considers the misalignment in the images due to the different geometric transformations and selects the center image as the reference. We employ the attention module in [31] to predict the attention maps against the reference. These attention maps can suppress the different geometric distortions in the non-reference images, which prevents the undesirable features from reaching the merging process whose results are used as inputs for the dual-domain network. Finally, feature merging processes the stack of aligned features with respect to the reference and the global module to obtain the final global features containing additional information for the moiré artifact removal network.

Dual-domain network: We use the same frequency network in Section 3.1 for burst-image demoireing. Specifically, the frequency network takes only the center image in the sequence as input and removes moiré artifacts therein.

3.3. Implementation Details

The proposed algorithm includes three neural networks: the pixel network, frequency network, and fusion network. In our implementation, each convolution is followed by a scaled exponential linear unit (SELU) activation function [17]. In each training batch, we apply geometric transformations of $90°$, $-90°$, and $180°$ rotations and horizontal and vertical flipping, thereby producing seven additional augmented versions of each image. We trained the networks using the AdamW optimizer [20] with $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The learning rate was fixed to 10^{-3}, and the batch size was set to 16. We trained the network using the L_2 loss. We implemented our model using PyTorch [23].

We experimentally found that training the networks separately is more efficient than end-to-end training in terms of both training time and memory space. Thus, first, we trained the pixel network and the frequency network separately. Then, we trained the fusion network. In the challenge, we also employed ensemble strategies by applying geometric transformations of $90°$, $-90°$, and $180°$ rotations and horizontal and vertical flipping, thereby producing seven additional augmented versions of each image. Therefore, we yield eight different output images using the proposed network. Inverse geometric transformations are applied to the output images generated from the augmented images. Finally, the final output image is generated by averaging the eight output images.

We used the training dataset provided by the NTIRE 2020 Demoireing Challenge [33].[1] Because the challenge does not provide ground-truth images for validation, we randomly selected 500 images from the training dataset as the test set for the experiments. Thus, the new training dataset contains 9,500 images out of 10,000. The training took approximately two days for single image and three days for burst images using a computer with Intel® Core™

[1] https://competitions.codalab.org/competitions/22223

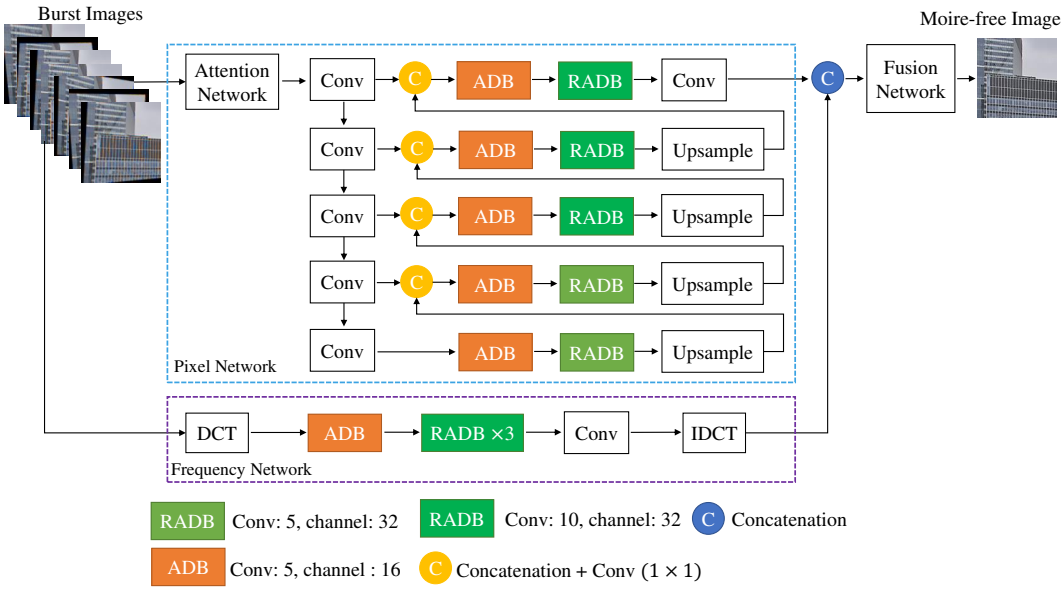

Figure 3: Architecture of the proposed network for burst-image demoireing.

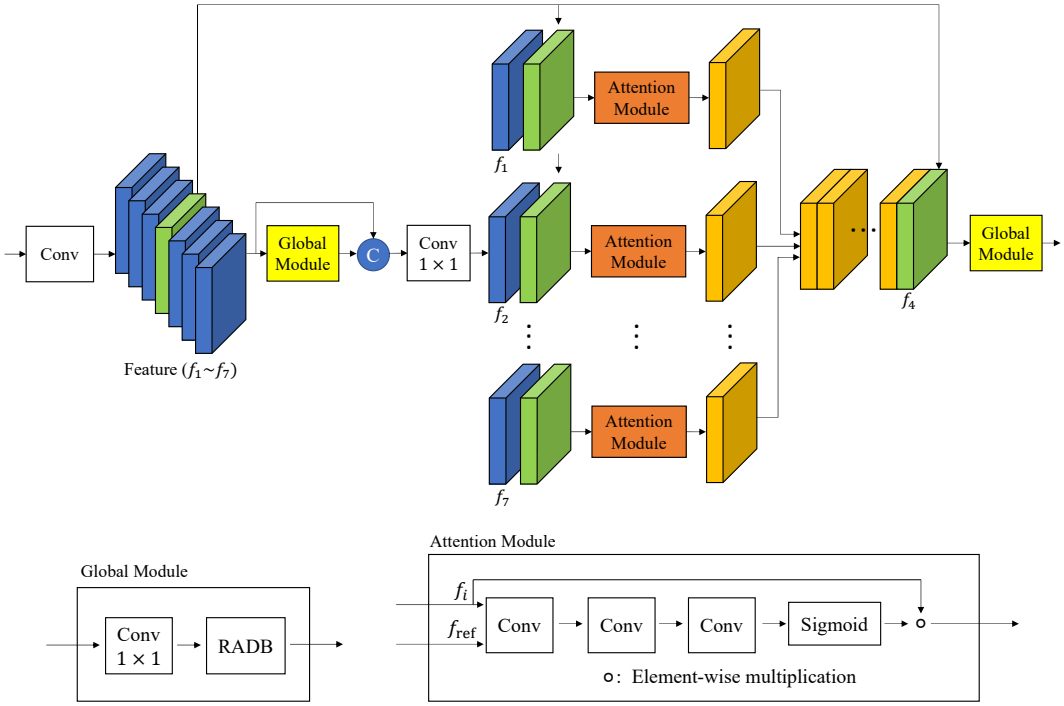

Figure 4: Architecture of the proposed attention network.

i9-9900X @4.4GHz CPU, 64GB RAM, and Nvidia RTX™ 2080 Ti GPU.

4. Experimental Results

4.1. Quantitative and Qualitative Evaluation

In the testing phase, the dual-domain network is fully end-to-end, in that it takes a moiré image as input and produces a moiré-free image. Table 1 shows the quantitative comparisons on the test set for both tracks in the challenge,

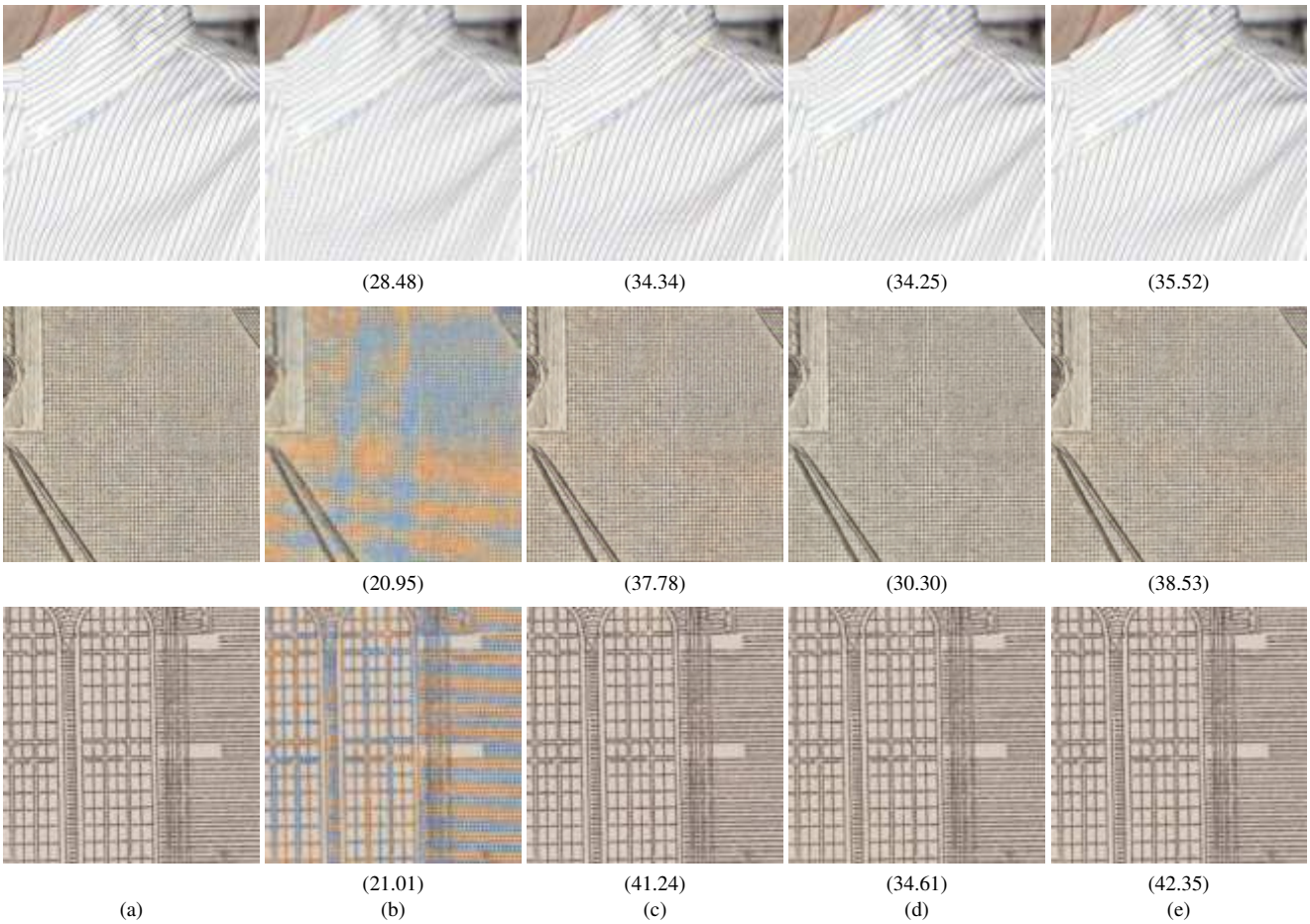

Figure 5: Single-image demoireing results for the test set. (a) Ground-truth, (b) moiré images, and outputs of (c) pixel network, (d) frequency network, and (e) fusion network. PSNR scores are provided below each image.

Table 1: Quantitative comparison of the demoireing performances of the proposed dual-domain network in the NTIRE 2020 Demoireing Challenge: Track 1 (Single image) and Track 2 (Burst). The boldfaced numbers denote the best results.

	Pixel	Frequency	Fusion
Track 1	38.07	35.61	**38.74**
Track 2	38.48	35.61	**39.22**

i.e., Track 1 (Single image) and Track 2 (Burst). We compute the average PSNR scores. Because the fusion network effectively combines the results from the pixel and frequency networks using the learnable filters, it shows the highest PSNR scores. In addition, using burst images with the proposed attention network further improves the performance by 0.48 dB by effectively exploiting full information from the images.

Figure 5 visually compares the demoireing results. The pixel network in Figure 5(c) yields blurring artifacts in highly-textured regions, although it effectively removes moiré artifacts. The frequency network in Figure 5(d) preserves fine details more faithfully compared with the pixel network; however it provides severe blurring artifacts. The fusion network in Figure 5(e) yields the highest-quality images, preserving the high-frequency details and effectively removing the moiré artifacts. This is because the fusion network learns the blending filters and merges the results from the two networks. We also verify the effectiveness of the attention network for burst-image demoireing using the dual-domain network. Figure 6 shows that burst demoireing results in a significantly higher image quality than single image demoireing in Figure 5, as demonstrated by the higher PSNR scores.

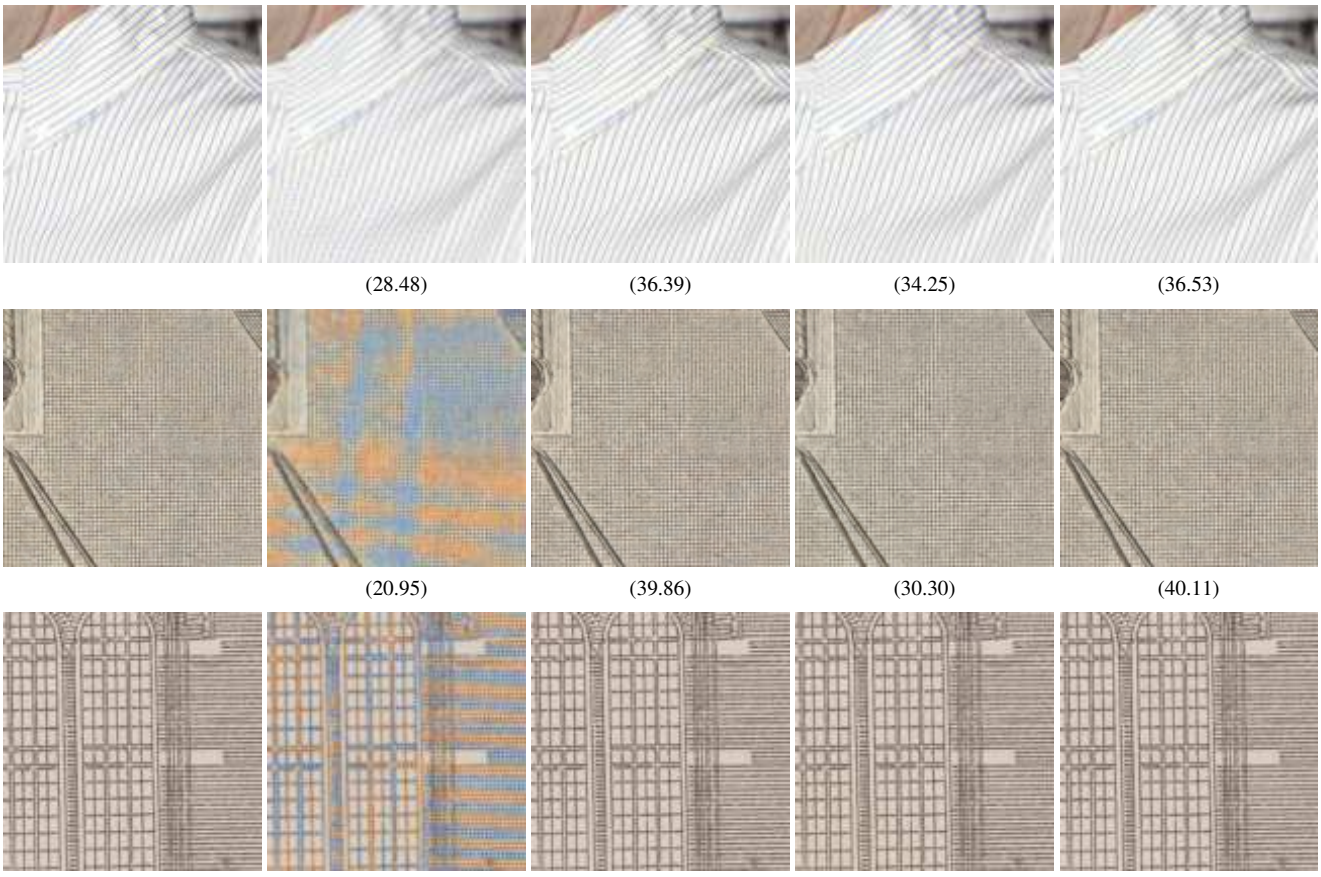

Figure 6: Burst-image demoireing results for the test set. (a) Ground-truth, (b) moiré images, and outputs of (c) pixel network, (d) frequency network, and (e) fusion network. PSNR scores are provided below each image.

4.2. Model Analysis

We conduct ablation studies to analyze the contributions of the key components in the pixel network, frequency network, and fusion network.

Pixel network: To analyze the optimal number of levels and the effectiveness of the edge extraction layer in a multi-resolution pixel network, we train the proposed pixel network using various settings. Table 2 compares the PSNR performances of the different settings. First, the number of levels has a significant impact on the demoireing performance; as the number of levels increases, the PSNR score gets higher accordingly. When the edge extraction layer is used, the demoireing performance significantly improves. More specifically, adding the edge extraction layer improves the PSNR score by 0.4 dB when the number of levels is 5.

Frequency network: We analyze the effects of the CBAM module in the frequency network by evaluating the number of RADBs and convolutional layers in each RADB. Table 3 shows the quantitative comparison of the different settings. We see that the frequency network with three RADBs and 10 convolutional layers achieves the highest PSNR score.

Table 2: Analysis of the pixel network.

No. of levels	Edge Extraction Layer	PSNR
1	-	37.25
1	✓	37.86
2	-	37.32
2	✓	37.67
3	-	37.46
3	✓	37.81
4	-	37.72
4	✓	38.05
5	-	37.88
5	✓	**38.28**

Table 3: Analysis of the frequency network.

No. of RADBs	No. of Convs in RADB	PSNR
3	3	34.02
3	5	34.49
3	10	**35.22**
5	3	34.31
5	5	34.49
5	10	35.04

Table 4: Analysis of the fusion network.

	Conv	Fusion Network
PSNR	38.72	**38.86**

Table 5: Analysis of the loss functions and residual learning.

Loss	Residual	PSNR
L_2	-	**38.86**
L_2	✓	38.77
L_1	-	38.84
L_1	✓	**38.86**

Table 6: Analysis of the attention network for burst-image demoireing.

Attention module	Global Module	PSNR
✓	-	37.82
-	✓	37.65
✓	✓	**37.93**

Fusion network: We analyze the effectiveness of the fusion network with a dynamic filter generation network over a 1×1 convolutional layer. Table 4 shows the quantitative comparison of these settings. The fusion network yields a PSNR score 0.14 dB higher than the convolutional layer, which confirms the effectiveness of the proposed fusion network.

Loss functions and residual learning: We analyze the performance of the proposed network trained with different loss functions [40] and residual connection. Table 5 shows the quantitative comparison of these combinations. First, training with the L_1 loss has less impact to directly learning moiré-free images without residual connection. When the residual connection is used, it achieves the highest score. Additionally, training with the L_1 loss yields 0.09 dB higher than that with the L_2 loss to learn moiré patterns with residual connection.

Burst-image demoireing: To analyze the effectiveness of the attention module and the global module in the proposed attention network, we train the proposed network to remove moiré artifacts in burst images using different combinations of the modules. Table 6 shows the quantitative comparison of the module combinations. We can achieve the highest demoireing performance when both attention and global modules are used.

5. Conclusions

In this work, we developed a dual-domain CNN consisting of a pixel network, frequency network, and fusion network to remove moiré artifacts in images. The pixel network converts moiré images directly into feature maps and processes these features at different resolutions. The frequency network removes moiré artifacts in the frequency domain by processing the DCT coefficients. Finally, the outputs of the pixel network and the frequency network are combined using the learned dynamic blending filters to generate a moiré-free image. We also showed that the proposed dual-domain network can be extended to remove moiré artifacts in multiple images with different geometric transformations of the same scene. We demonstrated the effectiveness of the proposed demoireing algorithm through the NTIRE 2020 Demoireing Challenge: Track 1 (Single image) and Track 2 (Burst).

References

[1] A. Abdelhamed, R. Timofte, M. S. Brown, et al. NTIRE 2019 challenge on real image denoising: Methods and results. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit Workshops*, Jun. 2019. 2

[2] M. Aittala and F. Durand. Burst image deblurring using permutation invariant convolutional neural networks. In *Proc. European Conf. Comput. Vis.*, 2018. 2

[3] S. Anwar and N. Barnes. Real image denoising with feature attention. In *Proc. IEEE Int. Conf. Comput. Vis.*, 2019. 2

[4] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao. DehazeNet: An end-to-end system for single image haze removal. *IEEE Trans. Image Process.*, 25(11):5187–5198, Nov. 2016. 2

[5] L. Cavigelli, P. Hager, and L. Benini. CAS-CNN: A deep convolutional neural network for image compression artifact suppression. In *Int. Joint Conf. Neural Networks*, 2017. 2

[6] X. Cheng, Z. Fu, and J. Yang. Multi-scale dynamic feature encoding network for image demoiréing. In *Proc. IEEE Int. Conf. Comput. Vis. Workshops*, 2019. 1, 2, 3

[7] T. Dai, J. Cai, Y. Zhang, S. Xia, and L. Zhang. Second-order attention network for single image super-resolution. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019. 2

[8] T. Dai, H. Zha, Y. Jiang, and S. Xia. Image super-resolution via residual block attention networks. In *Proc. IEEE Int. Conf. Comput. Vis. Workshops*, 2019. 2

[9] C. Dong, Y. Deng, C. C. Loy, and X. Tang. Compression artifacts reduction by a deep convolutional network. In *Proc. IEEE Int. Conf. Comput. Vis.*, 2015. 2

[10] C. Dong, C. C. Loy, K. He, and T. Xiaoou. Learning a deep convolutional network for image super-resolution. In *Proc. European Conf. Comput. Vis.*, 2014. 2

[11] H. Fan and J. Zhou. Stacked latent attention for multimodal reasoning. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018. 2

[12] S. Gu, Y. Li, L. V. Gool, and R. Timofte. Self-guided network for fast image denoising. In *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2019. 3

[13] B. He, C. Wang, B. Shi, and L. Duan. Mop moiré patterns using MopNet. In *Proc. IEEE Int. Conf. Comput. Vis.*, 2019. 1, 2, 3

[14] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger. Densely connected convolutional networks. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017. 2, 3

[15] X. Jia, B. D. Brabandere, T. Tuytelaars, and L. V. Gool. Dynamic filter networks. In *Proc. Adv. Neural Inf. Process. Syst.*, 2016. 4

[16] J. Kim, J. K. Lee, and K. M. Lee. Accurate image super-resolution using very deep convolutional networks. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016. 2

[17] G. Klambauer, T. Unterthiner, A. Mayr, and S. Hochreiter. Self-normalizing neural networks. In *Proc. Adv. Neural Inf. Process. Syst.*, 2017. 4

[18] W. Lai, J. Huang, N. Ahuja, and M. Yang. Fast and accurate image super-resolution with deep Laplacian pyramid networks. *IEEE Trans. Pattern Anal. Mach. Intell.*, 41(11):2599–2613, Nov. 2019. 2

[19] V. Lempitsky, A. Vedaldi, and D. Ulyanov. Deep image prior. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018. 2

[20] I. Loshchilov and F. Hutter. Decoupled weight decay regularization. In *Proc. Int. Conf. Learning Represent.*, 2019. 4

[21] J. Lu, C. Xiong, D. Parikh, and R. Socher. Knowing when to look: Adaptive attention via a visual sentinel for image captioning. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017. 2

[22] S. Nah, S. Son, R. Timofte, et al. AIM 2019 challenge on video temporal super-resolution: Methods and results. In *Proc. IEEE Int. Conf. Comput. Vis. Workshops*, 2019. 4

[23] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala. PyTorch: An imperative style, high-performance deep learning library. In *Proc. Adv. Neural Inf. Process. Syst.*, 2019. 4

[24] I. Pekkucuksen and Y. Altunbasak. Multiscale gradients-based color filter array interpolation. *IEEE Trans. Image Process.*, 22(1):157–165, Jan. 2013. 1

[25] M. Schöberl, W. Schnurrer, A. Oberdörster, S. Fössel, and A. Kaup. Dimensioning of optical birefringent anti-alias filters for digital cameras. In *Proc. IEEE Int. Conf. Image Process.*, 2010. 1

[26] W. Shi, J. Caballero, F. Huszár, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016. 3

[27] M. Suganuma, X. Liu, and T. Okatani. Attention-based adaptive selection of operations for image restoration in the presence of unknown combined distortions. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019. 2

[28] Y. Sun, Y. Yu, and W. Wang. Moiré photo restoration using multiresolution convolutional neural networks. *IEEE Trans. Image Process.*, 27(8):4160–4172, Aug. 2018. 1, 2

[29] T. Tong, G. Li, X. Liu, and Q. Gao. Image super-resolution using dense skip connections. In *Proc. IEEE Int. Conf. Comput. Vis.*, 2017. 2, 3

[30] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon. CBAM: Convolutional block attention module. In *Proc. European Conf. Comput. Vis.*, 2018. 2, 3

[31] Q. Yan, D. Gong, Q. Shi, A. van den Hengel, C. Shen, I. Reid, and Y. Zhang. Attention-guided network for ghost-free high dynamic range imaging. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019. 2, 4

[32] J. Yang, F. Liu, H. Yue, X. Fu, C. Hou, and F. Wu. Textured image demoiréing via signal decomposition and guided filtering. *IEEE Trans. Image Process.*, 26(7):3528–3541, July 2017. 1, 2, 3

[33] S. Yuan, R. Timofte, A. Leonardis, G. Slabaugh, et al. NTIRE 2020 challenge on image demoireing: Methods and results. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2020. 2, 4

[34] S. Yuan, R. Timofte, G. Slabaugh, et al. AIM 2019 challenge on image demoireing: Methods and results. In *Proc. IEEE Int. Conf. Comput. Vis. Workshops*, 2019. 1, 2

[35] H. Zhang and V. M. Patel. Densely connected pyramid dehazing network. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018. 2

[36] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Trans. Image Process.*, 26(7):3142–3155, July 2017. 2

[37] K. Zhang, W. Zuo, S. Gu, and L. Zhang. Learning deep CNN denoiser prior for image restoration. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017. 2

[38] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu. Residual dense network for image super-resolution. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018. 2, 3

[39] B. Zhao, X. Wu, J. Feng, Q. Peng, and S. Yan. Diversified visual attention networks for fine-grained object classification. *IEEE Trans. Multimedia*, 19(6):1245–1256, Jun. 2017. 2

[40] H. Zhao, O. Gallo, I. Frosio, and J. Kautz. Loss functions for image restoration with neural networks. *IEEE Trans. Compt. Imaging*, 3(1):47–57, Mar. 2017. 8