

# Learning Image Demoiréing from Unpaired Real Data

**Yunshan Zhong<sup>1,2</sup>, Yuyao Zhou<sup>2,3</sup>, Yuxin Zhang<sup>2,3</sup>, Fei Chao<sup>2,3</sup>, Rongrong Ji<sup>1,2,3,4\*</sup>**

<sup>1</sup>Institute of Artificial Intelligence, Xiamen University.

<sup>2</sup>Key Laboratory of Multimedia Trusted Perception and Efficient Computing,  
Ministry of Education of China, Xiamen University.

<sup>3</sup>Department of Artificial Intelligence, School of Informatics, Xiamen University.

<sup>4</sup>Peng Cheng Laboratory.

## Abstract

This paper focuses on addressing the issue of image demoiréing. Unlike the large volume of existing studies that rely on learning from paired real data, we attempt to learn a demoiréing model from unpaired real data, i.e., moiré images associated with irrelevant clean images. The proposed method, referred to as Unpaired Demoiréing (UnDeM), synthesizes pseudo moiré images from unpaired datasets, generating pairs with clean images for training demoiréing models. To achieve this, we divide real moiré images into patches and group them in compliance with their moiré complexity. We introduce a novel moiré generation framework to synthesize moiré images with diverse moiré features, resembling real moiré patches, and details akin to real moiré-free images. Additionally, we introduce an adaptive denoise method to eliminate the low-quality pseudo moiré images that adversely impact the learning of demoiréing models. We conduct extensive experiments on the commonly-used FHDMI and UHDM datasets. Results manifest that our UnDeM performs better than existing methods when using existing demoiréing models such as MBCNN and ESDNet-L. Code: <https://github.com/zysxmu/UnDeM>.

## Introduction

Contemporary society is awash with electronic screens for presenting images, text, video, *etc.* With the widespread availability of portable camera devices such as smartphones, people have grown accustomed to using them for quick information recording. Unfortunately, a common issue arises from the intrinsic interference between the camera's color filter array (CFA) and LCD subpixel layout of the screen (Yu et al. 2022), resulting in captured pictures being contaminated with some rainbow-shape stripes, which are also known as moiré patterns (Sun, Yu, and Wang 2018; Yang et al. 2017b). These moiré patterns involve varying thickness, frequencies, layouts, and colors, which degrade the perceptual quality of captured pictures. Consequently, there has been considerable academic and industrial interest in developing demoiréing algorithms to rectify the issue.

Primitive research on demoiréing are mostly established upon image priors (Dabov et al. 2007; Cho et al. 2011)

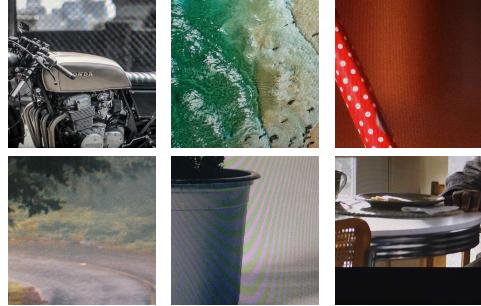


Figure 1: Illustration of image moiré. Natural moiré patterns are complex in varying thicknesses, frequencies, layouts, and colors across images and within an image.

or traditional machine learning methods (Liu, Yang, and Yue 2015; Yang et al. 2017a), which are demonstrated to be inadequate for tackling moiré patterns of drastic variations (Zheng et al. 2021). Fortunately, the fashionable convolutional neural networks (CNNs) have become a de facto infrastructure for the success of various computer vision tasks including the image demoiréing (He et al. 2020; Cheng, Fu, and Yang 2019; He et al. 2019; Liu et al. 2020; Sun, Yu, and Wang 2018; Yuan et al. 2019; Zheng et al. 2021; Yu et al. 2022; Liu, Shu, and Wu 2018; Gao et al. 2019). These CNN-based methods are typically trained on extensive pairs of moiré-free and moiré images in a supervised manner to model the demoiréing mapping. However, it is challenging to collect paired images given the fact in Fig. 1 that natural moiré patterns are featured with varying thicknesses, frequencies, layouts, and colors (Zheng et al. 2021). We can easily access to the moiré images as well as moiré-free images, but they are mostly unpaired. Although many studies try to capture image pairs from digital screens (He et al. 2020; Yu et al. 2022), their quality is barricaded by three limitations. First, acquiring high-quality image pairs involves professional camera position adjustments and even special hardware (Yu et al. 2022). Second, burdensome manpower is required to select well-aligned moiré-free and moiré pairs. Third, the captured moiré contents are very unitary under highly-controlled lab environments. However, image pairs full of more diverse moiré patterns are more expected for improving demoiréing models.

\*Corresponding Author

Copyright © 2024, Association for the Advancement of Artificial Intelligence ([www.aaai.org](http://www.aaai.org)). All rights reserved.

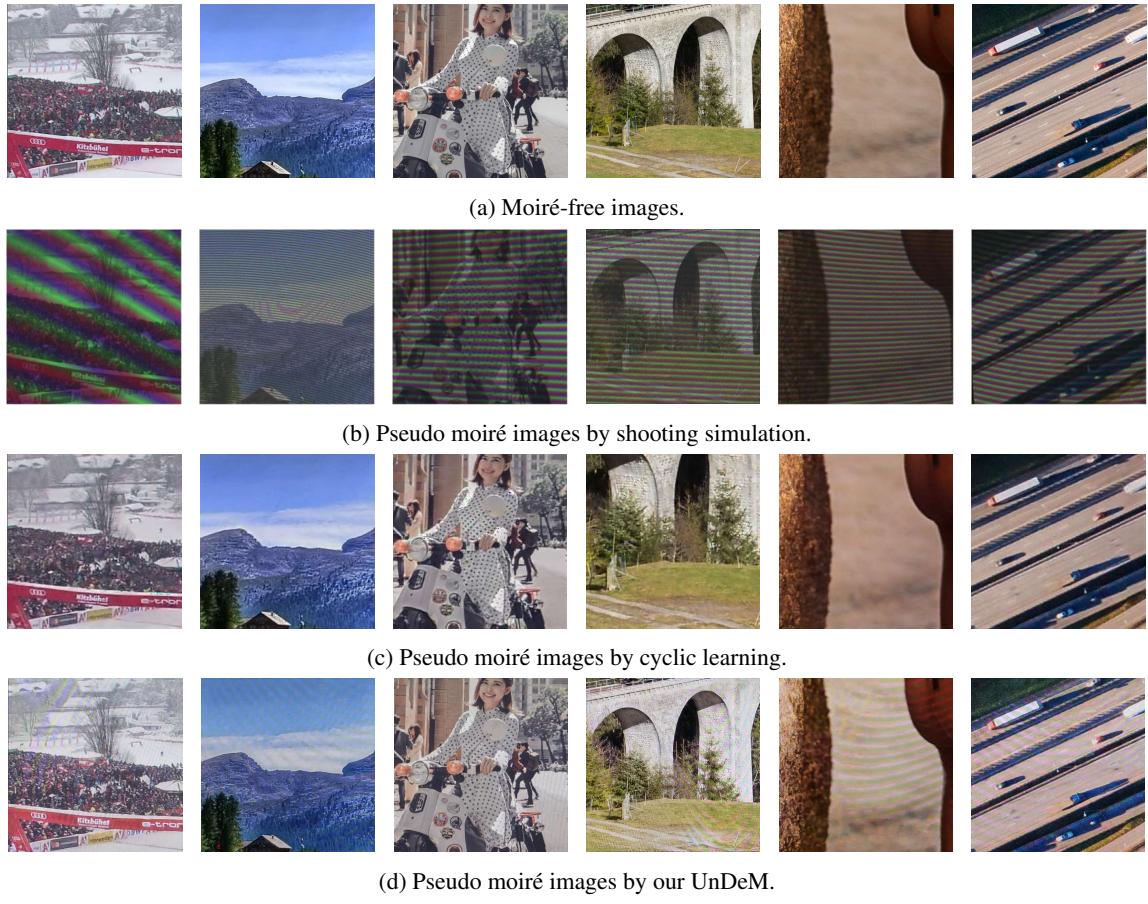


Figure 2: Visual examples of (a) real moiré-free images; (b) pseudo moiré images by shooting simulation (Niu, Guo, and Wang 2021); (c) pseudo moiré images by cyclic learning (Park et al. 2022); (d) pseudo moiré images by our UnDeM. Compared with details-missing and inauthentic pseudo moiré images by shooting simulation and (Park et al. 2022), ours result in more diverse moiré patterns and preserve more details of the moiré-free images. Best view by zooming in.

Synthesizing moiré images has therefore attracted increasing attention recently. Given Fig. 2a illustrated moiré-free screenshots, shooting simulation methods (Liu, Shu, and Wu 2018; Yuan et al. 2019; Niu, Guo, and Wang 2021) simulate the aliasing between CFA and screen’s LCD subpixel to produce corresponding paired moiré images in Fig. 2b. However, the synthetic images are insufficient to capture characteristics of real moiré patterns, leading to a large domain gap as we analyze from two aspects. First, the synthetic moiré images are much darker and cannot well seize the light quality, destroying the context of the viewing environment and obscuring the image details. Second, the synthetic moiré patterns lack authenticity, as the thicknesses, frequencies, layouts, and colors of moiré stripes are almost the same within an image. In Table 1 and Table 2 of the experimental section, we apply shooting simulated moiré images to train demoiréing CNNs. Our results manifest that the trained models generalize poorly to the natural-world test datasets. In (Park et al. 2022), Park *et al.* introduced a cyclic moiré learning method and we observe a better performance than shooting simulation in Table 1 and Table 2. However,

the generated pseudo moiré fails to accurately model real moiré patterns as illustrated in Fig. 2c, leading to limited performance. Therefore, it is desired to develop a better method to improve the synthetic moiré images.

In this paper, we present one novel method, dubbed as UnDeM, to learn demoiréing from unpaired real moiré and clean images that are fairly easy to collect, for example, by performing random screenshots and taking random photos from the digital screen. As displayed in Fig. 2d, the basic objective of our UnDeM is to synthesize moiré images that possess moiré features as the real moiré images and details as the real moiré-free images. The synthesized pseudo moiré images then form pairs with the real moiré-free images for training demoiréing networks. To this end, as shown in Fig. 3, we first split images into patches. These moiré patches are further grouped using a moiré prior that takes into consideration frequency and color information in each patch (Zhang et al. 2023). Consequently, moiré patches within each group fall into similar complexity such that they can be better processed by the individual moiré synthesis network. Specifically, the introduced synthesis network con-

tains four modules including a moiré feature encoder to extract moiré features of real moiré patches, a generator to synthesize pseudo moiré patches, a discriminator to identify real or pseudo moiré patches, and a content encoder to retain content information of real clean patches in synthesized pseudo moiré patches. The whole framework is conducted in an adversarial training manner (Goodfellow et al. 2014) for a better moiré image generation.

Before paired with real moiré-free images for training demoiréing networks, the synthesized moiré patches further undergo an adaptive denoise process to rule out these low-quality moiré patterns that bear image detail loss. Concretely, we find low-quality pseudo moiré leads to a large structure difference from its moiré-free counterpart, which therefore can be removed if the difference score is beyond a threshold adaptive to a particular percentile of the overall structure differences. Experiments in Table 1 and Table 2 demonstrate that, the proposed UnDeM improves the compared baseline by a large margin, on the real moiré image dataset. For example, when trained with a size of 384, MBCNN (Zheng et al. 2020) trained on the synthetic images from our UnDeM achieves 19.89 dB in PSNR on FHDMi (He et al. 2020), while obtaining 19.36 dB from cyclic moiré learning (Park et al. 2022) and only 9.32 dB from shooting simulation. Such results not only demonstrate our efficacy, but also enlighten a new moiré generation method for the demoiréing community.

## Related Work

### Image Demoiréing

Image demoiréing target at cleansing moiré patterns in taken photos. Earlier studies resort to some property presumptions of moiré patterns, such as space-variant filters (Siddiqui, Boutin, and Bouman 2009; Sun, Li, and Sun 2014), a low-rank constrained sparse matrix decomposition (Liu, Yang, and Yue 2015; Yang et al. 2017a), and layer decomposition (Yang et al. 2017b). Along with the surge of deep learning on many computer vision tasks, demoiréing also benefits from the convolutional neural networks (CNNs) recently. As the pioneering study, Sun et al. (Sun, Yu, and Wang 2018) developed DMCNN, a multi-scale CNN, to remove moiré patterns at different frequencies and scales. He et al. (He et al. 2019) proposed MopNet that is specially designed for unique properties of moiré patterns including frequencies, colors, and appearances. Zheng et al. (Zheng et al. 2020) introduced a multi-scale bandpass convolutional neural network (MBCNN) that consists of a learnable bandpass filter and a two-step tone mapping strategy to respectively deal with frequency prior and color shift. Liu et al. (Liu et al. 2020) designed WDNet that removes moiré patterns in the wavelet domain to effectively separate moiré patterns from image details. In (He et al. 2020), a multi-stage framework FHDe<sup>2</sup>Net is proposed. FHDe<sup>2</sup>Net employs a global to local cascaded removal branch to erase multi-scale moire patterns and a frequency-based branch to reserve fine details. Yu et al. (Yu et al. 2022) designed the ESDNet that utilizes the computationally-efficient semantic-aligned scale-aware module to enhance the network’s capability. However, all

these mentioned approaches require large amounts of moiré and moiré-free pairs. To solve this limitation, a cycle loss is further constructed to simultaneously train a pseudo moiré generator and a demoiréing network (Park et al. 2022; Yue et al. 2021). Very differently, our proposed UnDeM does not involve demoiréing network in the moiré synthesis stage.

### Moiré Dataset

Since data-driven CNN-based algorithms require large amount of paired moiré and moiré-free images to complete the training, many efforts have been devoted to constructing large-scale image pairs. Sun et al. (Sun, Yu, and Wang 2018) built the first real-world moiré image dataset from ImageNet (Russakovsky et al. 2015). He et al. (He et al. 2020) proposed the first high-resolution moiré image dataset FHDMi to satisfy the practical application in the real world. Yu et al. (Yu et al. 2022) further proposed the ultra-high-definition demoiréing dataset UHDM containing 4K images. Nevertheless, the data preparation process requires huge human efforts, and the resulting datasets are confined to limited scenes. To avoid the drudgery of collecting real-world paired moiré and moiré-free images, shooting simulation that simulates the camera imaging process becomes a more valuable approach (Liu, Shu, and Wu 2018; Yuan et al. 2019). However, the synthetic data fails to model the real imaging process and leads to a large domain gap between synthetic data and real data. As a result, demoiréing models trained on synthetic data are incapable of handling real-world scenarios.

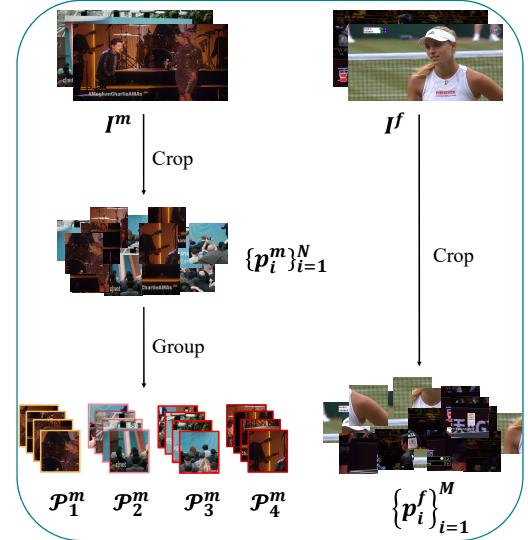


Figure 3: Image preprocessing. Both moiré images  $I^m$  and unpaired moiré-free images  $I^f$  are split into patches. Patches from moiré images are further grouped in compliance with the complexity of moiré patterns.

## Methodology

Our UnDeM contains image preprocessing, moiré synthesis network, and adaptive denoise, which are detailed one by one in the following.



Figure 4: An illustration of moiré images of each group. Each group has its own moiré patterns complexity.

## Image Preprocessing

Moiré patterns vary significantly even within one single image. It is challenging for one single network to learn all cases. To better learn from these different moiré patterns, we apply an isolated moiré synthesis network to deal with moiré patterns with similar complexity. We first split images in moiré set  $\mathcal{I}^m$  into non-overlapping patches, leading to a moiré patch set  $\mathcal{P}^m = \{p_i^m\}_{i=1}^N$ , where  $N$  is the number of patches for the whole moiré patch set. Similarly, we can have an  $M$ -size moiré-free patch set  $\mathcal{P}^f = \{p_i^f\}_{i=1}^M$  for  $\mathcal{I}^f$ . As illustrated in Fig. 3, we divide the moiré set  $\mathcal{P}^m$  into  $K$  subsets  $\mathcal{P}^m = \mathcal{P}_1^m \cup \mathcal{P}_2^m \cup \dots \cup \mathcal{P}_K^m$ . Each  $\mathcal{P}_j^m$  contains moiré patches with similar complexity and any two subsets are disjoint. Zhang *et al.* (Zhang et al. 2023) showed that a perceptible moiré pattern is highlighted by either high frequency or rich color information. Following (Zhang et al. 2023), given a moiré patch  $p^m \in \mathcal{P}$ , whose frequency is measured by a Laplacian edge detection operator  $\mathcal{F}(p^m)$  with kernel size of 3 (Marr and Hildreth 1980). In addition, the colorfulness, denoted as  $\mathcal{C}(p^m)$ , is the linear combination of the mean and standard deviation of the pixel cloud in the color plane of RGB colour space (Hasler and Suesstrunk 2003):

$$\begin{aligned} \mathcal{C}(p^m) &= \sqrt{\sigma^2(p_R^m - p_G^m) + \sigma^2(0.5(p_R^m + p_G^m) - p_B^m)} \\ &+ 0.3\sqrt{\mu^2(p_R^m - p_G^m) + \mu^2(0.5(p_R^m + p_G^m) - p_B^m)}, \end{aligned} \quad (1)$$

where  $\sigma(\cdot)$  and  $\mu(\cdot)$  return standard deviation and mean value of inputs, the  $p_R^m$ ,  $p_G^m$ , and  $p_B^m$  denote the red, green, and blue color channels of  $p^m$ .

We set  $K = 4$  and obtain four evenly-sized subsets of moiré patches, each of which has distinctive moiré features. The first group  $\mathcal{P}_1^m$  contains patches with the first  $N/4$  smallest  $\mathcal{F}(p^m) \cdot \mathcal{C}(p^m)$ , thus it has moiré patterns of low frequency and less color. We sort the remaining patches from the smallest to the largest with a new metric  $\mathcal{F}(p^m)/\mathcal{C}(p^m)$ . Then  $\mathcal{P}_2^m$  consists of the first  $N/4$  patches highlighted by low frequency but rich color. The middle  $N/4$  patches form

$\mathcal{P}_3^m$  featured with high frequency and rich color. The  $N/4$  smallest scored patches with high frequency but less color make up with  $\mathcal{P}_4^m$ . Fig. 4 gives some visual examples.

## Moiré Synthesis Network

Fig. 5 depicts an overall framework of our moiré synthesis network  $\mathcal{T}_i$  to learn moiré patterns from the group  $\mathcal{P}_i^m$ . It consists of a moiré feature encoder  $E^m$ , a generator  $G^m$ , a discriminator  $D^m$ , and a content encoder  $E^c$ . Given an unpaired moiré patch  $p^m \in \mathcal{P}_i^m$  and a moiré-free patch  $p^f \in \mathcal{P}^f$ , our motivation is to produce a pseudo moiré  $\tilde{p}^m$  that possesses the moiré pattern of  $p^m$  while still retaining image details of  $p^f$ , such that  $(\tilde{p}^m, p^f)$  forms moiré and moiré-free pairs to guide the learning of existing demoiré networks.

To fulfill this objective, the moiré feature encoder  $E^m$  extracts the moiré features of the real moiré patch  $p^m$ , denoted as  $F^m$ :

$$F^m = E^m(p^m). \quad (2)$$

Then, the generator  $G^m$  is to synthesize a pseudo moiré patch  $\tilde{p}^m$  with  $F^m$  and  $p^f$  as its inputs:

$$\tilde{p}^m = G^m(Con(F^m, p^f)), \quad (3)$$

where  $Con(\cdot, \cdot)$  indicates the concatenation operation.

The discriminator  $D^m$  cooperates with the generator  $G^m$  for a better pseudo moiré patch in an adversarial training manner (Goodfellow et al. 2014). The generator  $G^m$  is trained to trick the discriminator  $D^m$  by:

$$\mathcal{L}^{\text{dis-G}} = (D^m(\tilde{p}^m) - 1)^2. \quad (4)$$

The least squares loss function (Mao et al. 2017) is used for a better training stability. Also,  $D^m$  is trained to distinguish the pseudo moiré patch  $\tilde{p}^m$  from the real  $p^m$ :

$$\mathcal{L}^{\text{dis-D}} = (D^m(\tilde{p}^m))^2 + (D^m(p^m) - 1)^2. \quad (5)$$

The loss functions of Eq. (4) and Eq. (5) are optimized in a min-max game manner. As a result,  $D^m$  learns to distinguish the pseudo moiré and the real moiré images, while the moiré feature encoder  $E^m$  is forced to learn to extract the moiré feature appropriately and the generator  $G^m$  learns to synthesize real-like and in-distribution pseudo moiré images. In addition, we also require moiré feature of synthesized  $\tilde{p}^m$  to follow that of real  $p^m$  by:

$$\tilde{F}^m = E^m(\tilde{p}^m), \quad (6)$$

$$\mathcal{L}^{\text{fea}} = \|\tilde{F}^m - F^m\|_1, \quad (7)$$

where  $\|\cdot\|_1$  denotes the  $\ell_1$  loss. To well pair  $\tilde{p}^m$  and  $p^f$ ,  $\tilde{p}^m$  is also expected to have contents details of  $p^f$ . An additional content encoder  $E^c$  is introduced to align content features between  $\tilde{p}^m$  and  $p^f$  as:

$$\mathcal{L}^{\text{con}} = \|E^c(\tilde{p}^m) - E^c(p^f)\|_1. \quad (8)$$

Combining Eq. (4), Eq. (5), Eq. (7), and Eq. (8) leads to our final loss function as:

$$\mathcal{L} = \mathcal{L}^{\text{dis-G}} + \mathcal{L}^{\text{dis-D}} + \mathcal{L}^{\text{fea}} + \mathcal{L}^{\text{con}}. \quad (9)$$

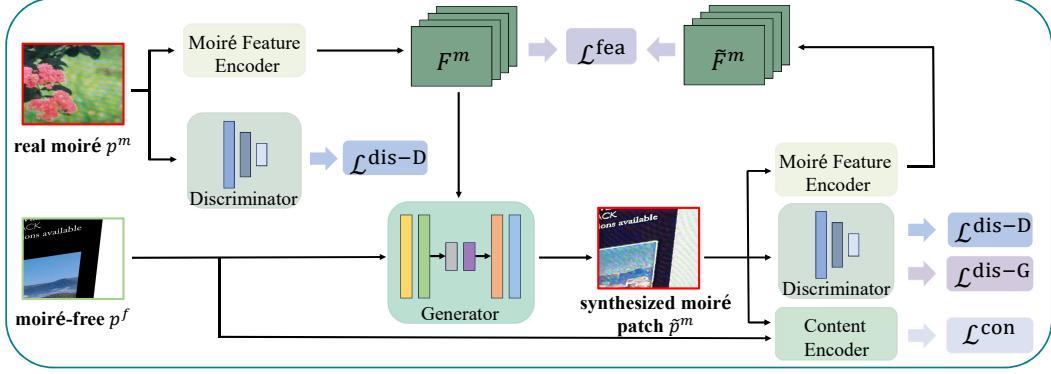


Figure 5: Framework of our moiré synthesis network.

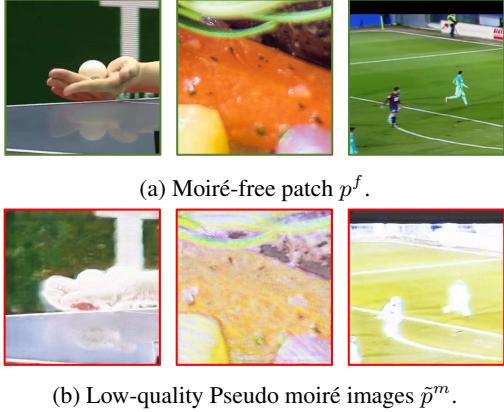


Figure 6: Examples of low-quality pseudo moiré images.

## Adaptive Denoise

After training our moiré synthesis networks  $\{\mathcal{T}_i\}_{i=1}^4$ , the pseudo moiré patch  $\tilde{p}^m$ , paired with corresponding moiré-free patch  $p^f$ , sets the dataset to train demoiréing network. Unfortunately, we find some pseudo moiré patches occasionally suffer from low-quality issues. Some examples are manifested in Fig. 6, where the contents and details of  $p^f$  are destroyed in  $\tilde{p}^m$ . Such noisy data hinders the learning of demoiréing models.

Fortunately, we observe in Fig. 6 that the ruined structure mostly attributes to the edge information. Therefore, we calculate the edge map of each patch by the Laplacian edge detection operator, and the structure difference is computed by summing up the absolute value of the edge difference between each pseudo pair. Low-quality pseudo moiré leads to a large score of structure difference, and we can rule out these pairs as long as the score is beyond a threshold that is adaptive to the  $\gamma$ -th percentile of structure differences in a total of  $N$  pseudo pairs. We conduct the above process for each synthesis network  $\mathcal{T}_i$  and set a corresponding  $\gamma_i$  to remove low-quality pseudo moiré. We find  $N = 6,400$  performs well already. Consequently, we obtain a better performance.

In summary, our UnDeM consists of 1) training a moiré synthesis network for synthesizing pseudo moiré images; 2)

training a demoiréing model based on the trained moiré synthesis network. This paper focuses on addressing moiré image generation. As for demoiréing models, we directly borrow from existing studies. Details of training algorithms are listed in the appendix.

## Experimentation

### Implementation Details

**Datasets.** Public demoiréing datasets used in this paper include the FHDMi (He et al. 2020) dataset and UHDM (Yu et al. 2022) dataset. The FHDMi dataset consists of 9,981 image pairs for training and 2,019 image pairs for testing with  $1920 \times 1080$  resolution. The UHDM dataset contains 5,000 image pairs with 4K resolution in total, of which 4,500 are used for training and 500 for testing. We use the training set to train the proposed moiré synthesis network. For image preprocessing, we crop the training images of FHDMi into 8 patches. For UHDM involving images with higher resolution, we crop the training images into 6 patches. During training, the moiré patch  $p^m$  and moiré-free patch  $p^f$  are selected from different original images (before image preprocessing) to ensure they are unpaired.

**Networks.** We implement our UnDeM using the Pytorch framework (Paszke et al. 2019). The architecture of moiré synthesis network is largely based on (Hu et al. 2019; Liu et al. 2021).  $E^m$  and  $E^c$  contain one convolutional layer and two residual blocks.  $G^m$  contains three convolutional layers, nine residual blocks, and two deconvolutional layers, and ends with a convolutional layer to produce the final output. The residual blocks constitute two convolutional layers that are followed by instance normalization and ReLU function (Ulyanov, Vedaldi, and Lempitsky 2016). The convolutional layer has 16 channels for  $E^m$  and  $E^c$  and 128 for  $G^m$ .  $D^m$  is borrowed from PatchGAN (Isola et al. 2017) and consists of three convolutional layers with a stride of 2, two convolutional layers with a stride of 1, and ends with an average pooling layer. For demoiréing models, we utilize MBCNN (Zheng et al. 2020) and ESDNet-L (Yu et al. 2022) (A large version of ESDNet).

The moiré synthesis network is trained using the Adam optimizer (Kingma and Ba 2014), where the first momentum and second momentum are set to 0.9 and 0.999, respec-

Model	C.S.	Method	PSNR↑	SSIM↑	LPIPS↓
MBCNN	192	Paired	22.49	0.815	0.191
		Shooting	10.66	0.477	0.570
		Cyclic	19.15	0.722	0.257
		UnDeM	19.45	0.732	0.230
ESDNet-L	384	Paired	22.73	0.819	0.182
		Shooting	9.32	0.513	0.572
		Cyclic	19.36	0.733	0.265
		UnDeM	19.89	0.735	0.226
ESDNet-L	192	Paired	22.86	0.823	0.143
		Shooting	10.06	0.558	0.487
		Cyclic	19.09	0.738	0.241
		UnDeM	19.38	0.749	0.228
ESDNet-L	384	Paired	23.45	0.834	0.134
		Shooting	9.81	0.553	0.512
		Cyclic	19.05	0.715	0.273
		UnDeM	19.66	0.747	0.205

Table 1: Quantitative results on the FHDMI dataset. The “C.S.” denotes size in the random crop and the “Paired” denotes real paired data.

tively. We use 100 epochs for training with a batch size of 4 and an initial learning rate of  $2 \times 10^{-4}$ , which is linearly decayed to 0 in the last 50 epochs. Besides, we perform different random crop sizes on the image patches after image preprocessing to validate the flexibility of our method for synthesizing pseudo moiré images. The crop sizes are set to  $192 \times 192$  and  $384 \times 384$  for FHDMI, and  $192 \times 192$ ,  $384 \times 384$ , and  $768 \times 768$  for UHDM, respectively. As for the demoiréing models, we retain the same training configurations as the original paper except that all models are trained for 150 epochs for a fair comparison. All networks are initialized using a Gaussian distribution with a mean of 0 and a standard deviation of 0.02. The  $\gamma_1$ ,  $\gamma_2$ ,  $\gamma_3$ , and  $\gamma_4$  for adaptive denoise are empirically set to 50, 40, 30, and 20, respectively<sup>1</sup>. All experiments are run on NVIDIA A100 GPUs.

**Evaluation Protocols.** We adopt the Peak Signal-to-Noise Ratio (PSNR), Structure Similarity (SSIM) (Wang et al. 2004), and LPIPS (Zhang et al. 2018) to quantitatively evaluate the performance of demoiréing models.

## Quantitative Results

**FHDMI.** We first analyze the performance on the FHDMI dataset by comparing our UnDeM against the baseline, *i.e.*, the results of the shooting simulation (Niu, Guo, and Wang 2021) and cyclic learning (Park et al. 2022). Table 1 shows that the performance of demoiréing models trained on data produced by shooting simulation is extremely poor. For example, MBCNN obtains only 10.66 dB of PSNR when trained with  $192 \times 192$  crop size, which indicates the existence of a large domain gap between the pseudo and real data. Both the cyclic learning method (Park et al. 2022) and our UnDeM exhibit much better results. Moreover, compared with (Park et al. 2022), our UnDeM successfully mod-

<sup>1</sup>Ablations on  $\gamma_i$  and each component in UnDeM are provided in the appendix.

Model	C.S.	Method	PSNR↑	SSIM↑	LPIPS↓
MBCNN	192	Paired	20.14	0.760	0.346
		Shooting	8.99	0.528	0.632
		Cyclic	17.42	0.663	0.464
		UnDeM	17.96	0.673	0.425
ESDNet-L	384	Paired	20.14	0.759	0.356
		Shooting	9.27	0.538	0.603
		Cyclic	17.68	0.665	0.476
		UnDeM	17.78	0.668	0.401
ESDNet-L	768	Paired†	21.41	0.793	0.332
		Shooting	9.33	0.543	0.605
		Cyclic	17.98	0.719	0.503
		UnDeM	18.13	0.723	0.360
ESDNet-L	192	Paired	21.30	0.786	0.258
		Shooting	9.80	0.606	0.544
		Cyclic	18.02	0.659	0.371
		UnDeM	18.30	0.662	0.365
ESDNet-L	384	Paired	21.18	0.785	0.257
		Shooting	10.27	0.604	0.522
		Cyclic	17.75	0.679	0.404
		UnDeM	18.18	0.688	0.361
ESDNet-L	768	Paired†	22.12	0.799	0.245
		Shooting	9.80	0.599	0.542
		Cyclic	18.00	0.697	0.423
		UnDeM	18.40	0.713	0.344

Table 2: Quantitative results on the UHDM dataset. The “C.S.” denotes size in the random crop and the “Paired” denotes real paired data. The “†” indicates results directly copied from (Yu et al. 2022).

els the moiré patterns and thus presents the highest performance. For instance, MBCNN respectively obtains 19.45 dB and 19.89 dB of PSNR when trained with crop sizes of 192 and 384. For ESDNet-L, the PSNR results are 19.38 dB and 19.66 dB, respectively. Correspondingly, the SSIM and LPIPS of our UnDeM also exhibit much better performance than shooting simulation and cyclic learning.

**UHDM.** The results on the UHDM dataset are provided in Table 2. Demoiréing models trained on shooting simulation still fail to deal with the real data, and cyclic learning provides better results. More importantly, our UnDeM surpasses these two methods over different networks and training sizes. Specifically, the UnDeM increases the PSNR by 0.54 dB, 0.10 dB, and 0.15 dB when training MBCNN with crop sizes of 192, 384, and 768, respectively. For ESDNet-L, the PSNR gains are 0.28 dB, 0.43 dB, and 0.40, respectively. To summarize from Table 1 and Table 2, we can conclude that the transferability of our produced moiré images to the downstream demoiréing tasks and the efficacy of our UnDeM over existing methods have therefore been well demonstrated.

## Qualitative Results

Qualitative comparisons of demoiréing images on UHDM dataset are presented in Fig. 7, with additional results provided in the appendix. As shown in Figure 7b, the



Figure 7: Visualization of demoiréing results of MBCNN (Crop Size: 768) on the UHDM dataset. For the convenience of demonstration, we crop the patch from the test image.

demoiréing results of shooting simulation exhibit unnaturally high brightness, leading to a loss of image detail. This decrease in visual quality can be blamed on the generally darker brightness of shooting simulation as shown in Fig. 2b, which makes the demoiréing model learning incorrect brightness relationship between the moiré and moiré-free images. As presented in Fig. 7c, the demoiréing model fails to remove moiré due to cyclic learning cannot model the moiré patterns as illustrated in Fig. 2c. Results in Fig. 7d demonstrate the efficacy of UnDeM in removing moiré patterns, reflecting the fact that UnDeM has the ability to successfully model the moiré patterns.

## Conclusion

In this paper, we present UnDeM that performs real image demoiréing using unpaired real data in a learning-based

manner. We synthesize pseudo moiré images to form paired data for training off-the-shelf demoiréing models. The proposed UnDeM contains three steps including image pre-processing, a moiré generation network, and adaptive denoise. The image preprocessing crops the real moiré images into multiple sub-image patches and groups them into four groups according to the moiré patterns complexity. A moiré generation network is applied to synthesize a pseudo moiré image that has the moiré feature as its input real moiré image and the image detail as its input moiré-free image. The adaptive denoise is introduced to rule out the low-quality synthetic moiré images for avoiding their adverse effects on the learning of demoiréing models. UnDeM is demonstrated to improve the quality of synthetic images and the demoiréing models trained on these images are experimentally shown to be superior in performance.

## Acknowledgments

This work was supported by National Key R&D Program of China (No.2022ZD0118202), the National Science Fund for Distinguished Young Scholars (No.62025603), the National Natural Science Foundation of China (No. U21B2037, No. U22B2051, No. 62176222, No. 62176223, No. 62176226, No. 62072386, No. 62072387, No. 62072389, No. 62002305 and No. 62272401), and the Natural Science Foundation of Fujian Province of China (No.2021J01002, No.2022J06001).

## References

- Cheng, X.; Fu, Z.; and Yang, J. 2019. Multi-scale dynamic feature encoding network for image demoiréing. In *Proceedings of IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, 3486–3493. IEEE.
- Cho, T. S.; Zitnick, C. L.; Joshi, N.; Kang, S. B.; Szeliski, R.; and Freeman, W. T. 2011. Image restoration by matching gradient distributions. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 34: 683–694.
- Dabov, K.; Foi, A.; Katkovnik, V.; and Egiazarian, K. 2007. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing (TIP)*, 16: 2080–2095.
- Gao, T.; Guo, Y.; Zheng, X.; Wang, Q.; and Luo, X. 2019. Moiré pattern removal with multi-scale feature enhancing network. In *Proceedings of the IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, 240–245.
- Goodfellow, I. J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; and Bengio, Y. 2014. Generative Adversarial Nets. In *Proceedings of the Advances in Neural Information Processing Systems (NeurIPS)*, 2672–2680.
- Hasler, D.; and Suesstrunk, S. E. 2003. Measuring colorfulness in natural images. In *Human vision and electronic imaging VIII*, volume 5007, 87–95.
- He, B.; Wang, C.; Shi, B.; and Duan, L.-Y. 2019. Mop moire patterns using mopnet. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2424–2432.
- He, B.; Wang, C.; Shi, B.; and Duan, L.-Y. 2020. FHDe<sup>2</sup>Net: Full high definition demoiréing network. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 713–729.
- Hu, X.; Jiang, Y.; Fu, C.-W.; and Heng, P.-A. 2019. Mask-ShadowGAN: Learning to remove shadows from unpaired data. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2472–2481.
- Isola, P.; Zhu, J.-Y.; Zhou, T.; and Efros, A. A. 2017. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 1125–1134.
- Kingma, D. P.; and Ba, J. 2014. Adam: A Method for Stochastic Optimization. In *Proceedings of the International Conference on Learning Representations (ICLR)*.
- Liu, B.; Shu, X.; and Wu, X. 2018. Demoiréing of Camera-Captured Screen Images Using Deep Convolutional Neural Network. *arXiv preprint arXiv:1804.03809*.
- Liu, F.; Yang, J.; and Yue, H. 2015. Moiré pattern removal from texture images via low-rank and sparse matrix decomposition. In *IEEE Visual Communications and Image Processing (VCIP)*, 1–4.
- Liu, L.; Liu, J.; Yuan, S.; Slabaugh, G.; Leonardis, A.; Zhou, W.; and Tian, Q. 2020. Wavelet-based dual-branch network for image demoiréing. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 86–102.
- Liu, Z.; Yin, H.; Wu, X.; Wu, Z.; Mi, Y.; and Wang, S. 2021. From shadow generation to shadow removal. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 4927–4936.
- Mao, X.; Li, Q.; Xie, H.; Lau, R. Y.; Wang, Z.; and Paul Smolley, S. 2017. Least squares generative adversarial networks. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2794–2802.
- Marr, D.; and Hildreth, E. 1980. Theory of edge detection. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 207: 187–217.
- Niu, D.; Guo, R.; and Wang, Y. 2021. Morié Attack (MA): A New Potential Risk of Screen Photos. In *Proceedings of the Advances in Neural Information Processing Systems (NeurIPS)*, 26117–26129.
- Park, H.; Vien, A. G.; Kim, H.; Koh, Y. J.; and Lee, C. 2022. Unpaired screen-shot image demoiréing with cyclic moiré learning. *IEEE Access*, 10: 16254–16268.
- Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. 2019. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In *Proceedings of the Advances in Neural Information Processing Systems (NeurIPS)*, 8026–8037.
- Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. 2015. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*, 115: 211–252.
- Siddiqui, H.; Boutin, M.; and Bouman, C. A. 2009. Hardware-friendly descreening. *IEEE Transactions on Image Processing (TIP)*, 19: 746–757.
- Sun, B.; Li, S.; and Sun, J. 2014. Scanned image descreening with image redundancy and adaptive filtering. *IEEE Transactions on Image Processing (TIP)*, 23: 3698–3710.
- Sun, Y.; Yu, Y.; and Wang, W. 2018. Moiré photo restoration using multiresolution convolutional neural networks. *IEEE Transactions on Image Processing (TIP)*, 27: 4160–4172.
- Ulyanov, D.; Vedaldi, A.; and Lempitsky, V. 2016. Instance normalization: The missing ingredient for fast stylization. *arXiv preprint arXiv:1607.08022*.
- Wang, Z.; Bovik, A. C.; Sheikh, H. R.; and Simoncelli, E. P. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing (TIP)*, 13: 600–612.

Yang, J.; Liu, F.; Yue, H.; Fu, X.; Hou, C.; and Wu, F. 2017a. Textured image demoiréing via signal decomposition and guided filtering. *IEEE Transactions on Image Processing (TIP)*, 26: 3528–3541.

Yang, J.; Zhang, X.; Cai, C.; and Li, K. 2017b. Demoiréing for screen-shot images with multi-channel layer decomposition. In *IEEE Visual Communications and Image Processing (VCIP)*, 1–4.

Yu, X.; Dai, P.; Li, W.; Ma, L.; Shen, J.; Li, J.; and Qi, X. 2022. Towards efficient and scale-robust ultra-high-definition image demoiréing. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 646–662.

Yuan, S.; Timofte, R.; Slabaugh, G.; Leonardis, A.; Zheng, B.; Ye, X.; Tian, X.; Chen, Y.; Cheng, X.; Fu, Z.; et al. 2019. Aim 2019 challenge on image demoiréing: Methods and results. In *Proceedings of IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, 3534–3545.

Yue, H.; Cheng, Y.; Liu, F.; and Yang, J. 2021. Unsupervised moiré pattern removal for recaptured screen images. *Neurocomputing*, 456: 352–363.

Zhang, R.; Isola, P.; Efros, A. A.; Shechtman, E.; and Wang, O. 2018. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 586–595.

Zhang, Y.; Lin, M.; Li, X.; Liu, H.; Wang, G.; Chao, F.; Ren, S.; Wen, Y.; Chen, X.; and Ji, R. 2023. Real-Time Image Demoiréing on Mobile Devices. In *Proceedings of the International Conference on Learning Representations (ICLR)*.

Zheng, B.; Yuan, S.; Slabaugh, G.; and Leonardis, A. 2020. Image demoiréing with learnable bandpass filters. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 3636–3645.

Zheng, B.; Yuan, S.; Yan, C.; Tian, X.; Zhang, J.; Sun, Y.; Liu, L.; Leonardis, A.; and Slabaugh, G. 2021. Learning frequency domain priors for image demoiréing. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 44: 7705–7717.

---

**Algorithm 1:** Overall process.

---

```

Input: Moiré image set  $\mathcal{I}^m$  and moiré-free image set  $\mathcal{I}^f$ .
1 Moiré synthesis network training:
2   Split the  $\mathcal{I}^m, \mathcal{I}^f$  into  $\mathcal{P}^m, \mathcal{P}^f$ , and group  $\mathcal{P}^m$ 
    into four groups  $\{\mathcal{P}_i^m\}_{i=1}^4$  according to their
    complexity (Sec.3.1 of the main paper).
3   for  $i = 1 \rightarrow 4$  do
4     Initialize a moiré synthesis network  $\mathcal{T}_i$  and
      train it with images of  $\{\mathcal{P}_i^m\}_{i=1}^4$  (Sec.3.2 of
      the main paper).
5     Determine a threshold  $\gamma_i$  for adaptive denoise
      (Sec.3.3 of the main paper).
6   return  $\mathcal{P}^f, \{\mathcal{P}_i^m\}_{i=1}^4, \{\mathcal{T}_i\}_{i=1}^4, \{\gamma_i\}_{i=1}^4$ .
7

Input:  $\mathcal{P}^f, \{\mathcal{P}_i^m\}_{i=1}^4, \{\mathcal{T}_i\}_{i=1}^4, \{\gamma_i\}_{i=1}^4$ .
8 Demoiréing model training:
9   Initialize a demoiréing model  $\mathcal{F}$ .
10  while not end do
11    Randomly select  $\gamma_i, \mathcal{P}_i^m$ , and  $\mathcal{T}_i$ .
12    Randomly select  $p^m$  from  $\mathcal{P}_i^m$  and  $p^f$  from
       $\mathcal{P}^f$ . Use  $\mathcal{T}_i$  to synthesize  $\tilde{p}^m$  with  $p^m$  and  $p^f$ 
      as its inputs.
13    Compute the structure difference score
      between  $\tilde{p}^m$  and  $p^f$ .
14    if  $\gamma_i \leq \text{score}$  then
15      Update  $\mathcal{F}$  with  $\tilde{p}^m$  and  $p^f$ .
16    else
17      Go to Line 12.
18  return  $\mathcal{F}$ 

```

---

### Algorithm Process

The detailed algorithm process of our method is presented in Alg. 1, which consists of 1) training a moiré synthesis network for synthesizing pseudo moiré images in Lines 1 – 6; 2) training a demoiréing model based on the trained moiré synthesis network in Lines 7 – 19.

### Ablation of Hyperparameters in Adaptive Denoise

We present the ablation study of hyperparameters in adaptive denoise, *i.e.*,  $\{\gamma_i\}_{i=1}^4$ . The experiments are conducted with MBCNN on the FHDMi dataset and the crop size is 192. Each  $\gamma_i$  is first empirically initialized. Then, we fix other hyperparameters when performing the ablation study of  $\gamma_i$  to search for the optimal value. From Fig. 8a to Fig. 8d, we can find that the best configuration for  $\gamma_1, \gamma_2, \gamma_3$ , and  $\gamma_4$  is 50, 40, 30, and 20, respectively. We use these configurations for all crop sizes and networks. Though they might not be optimal under different conditions, they already bring better performance than other methods.

Model	Components	PSNR↑	SSIM↑	LPIPS↓
MBCNN	All	19.45	0.732	0.230
	w/o Img Pre	16.62	0.645	0.309
	w/o Denoise	18.96	0.717	0.257
	w/o $\mathcal{L}^{\text{dis}}$	7.22	0.478	0.669
	w/o $\mathcal{L}^{\text{fea}}$	18.38	0.726	0.265
	w/o $\mathcal{L}^{\text{con}}$	17.90	0.674	0.280

Table 3: Ablation study of MBCNN on the FHDMi dataset. The crop size is 192. The “Img Pre” denotes image preprocessing and the “Denoise” denotes adaptive denoise.

### Ablation Study

We continue to perform ablation studies to analyze the proposed UnDeM. Table 3 suggests that each component in UnDeM plays an important role in the final demoiréing performance. In particular, removing image preprocessing and adaptive denoise respectively lead to PSNR drops of 2.83 dB and 0.49 dB. Also, the performance degrades if removing the losses used in the moiré synthesis network, *i.e.*, Eq. (4), Eq. (5), Eq. (7), and Eq. (8). Respectively, removing  $\mathcal{L}^{\text{dis}}$  ( $\mathcal{L}^{\text{dis-G}}$  and  $\mathcal{L}^{\text{dis-D}}$ ),  $\mathcal{L}^{\text{fea}}$ , and  $\mathcal{L}^{\text{con}}$  causes 12.23 dB, 1.07 dB, and 1.55 dB drop in PSNR. These results demonstrate the importance of adversarial training for producing real-like pseudo moiré images.

### More Qualitative Results

In this section, we present more qualitative comparisons. Fig. 9 and Fig. 10 present the qualitative results of MBCNN (Crop size: 384) and ESDNet-L (Crop size: 384) on FHDMi dataset, respectively. Fig. 11 presents the qualitative results of ESDNet-L (Crop size: 768) on UHDM dataset.

### Limitations and Discussions

Although the proposed UnDeM improves the performance of shooting simulation and cyclic learning by a large margin and provides a novel moiré generation paradigm for the research community, its performance still degrades a lot if compared with the results of real data. The further improvement of the quality of pseudo moiré still remains an open question. We analyze that, a more elaborated image preprocessing, dedicated moiré synthesis network and special design loss might be potential ways to improve the quality of pseudo images. Moving forward, further investigation into these potential solutions will be a major focus of our future research efforts.

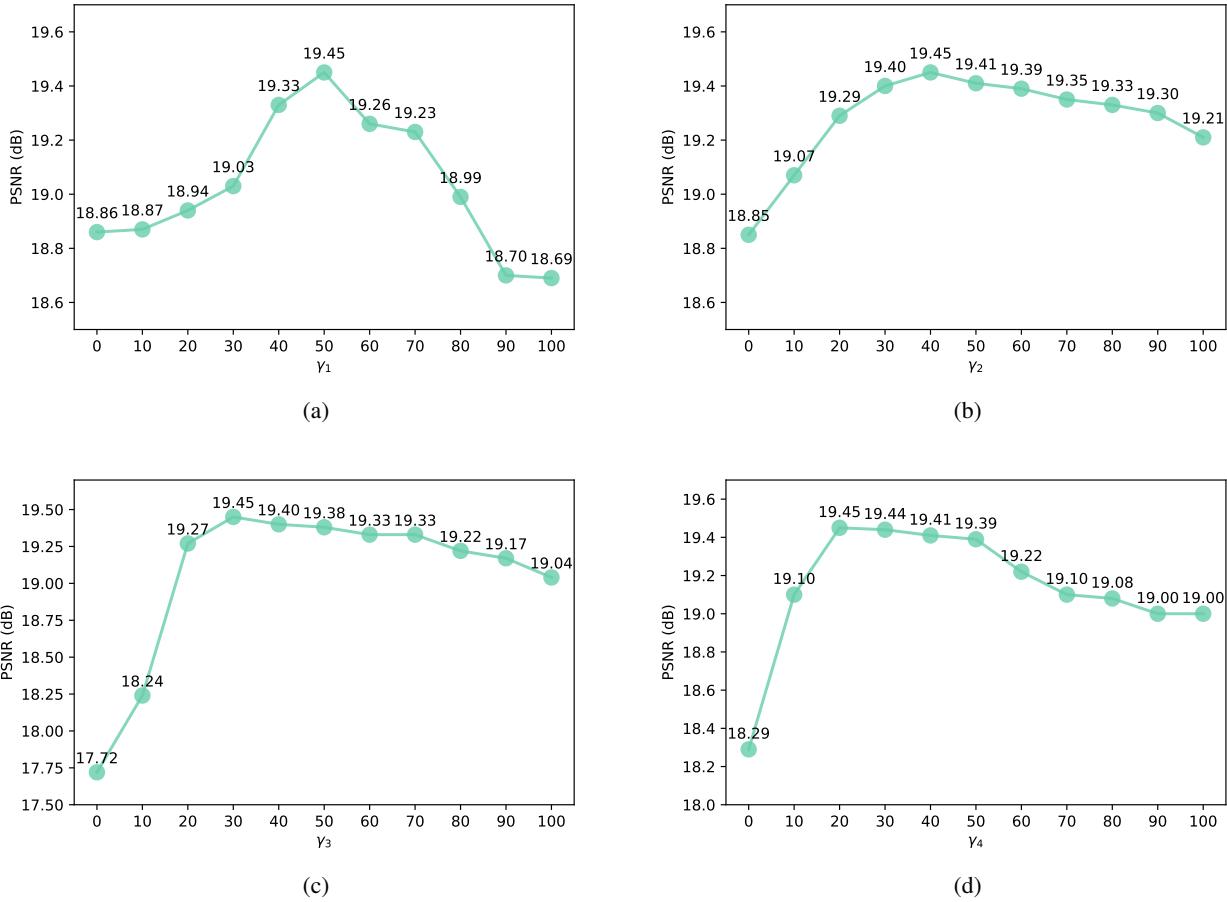


Figure 8: Ablation of  $\gamma_i$  in Adaptive Denoise.



(a) Moiré images.



(b) Demoiréing results by shooting simulation (Niu, Guo, and Wang 2021).



(c) Demoiréing results by cyclic learning (Park et al. 2022).



(d) Demoiréing results by our UnDeM.



(e) Moiré-free images.

Figure 9: Visualization of demoiréing results of MBCNN (Crop Size: 384) on the FHDMi dataset. For the convenience of demonstration, we crop the patch from the test image.

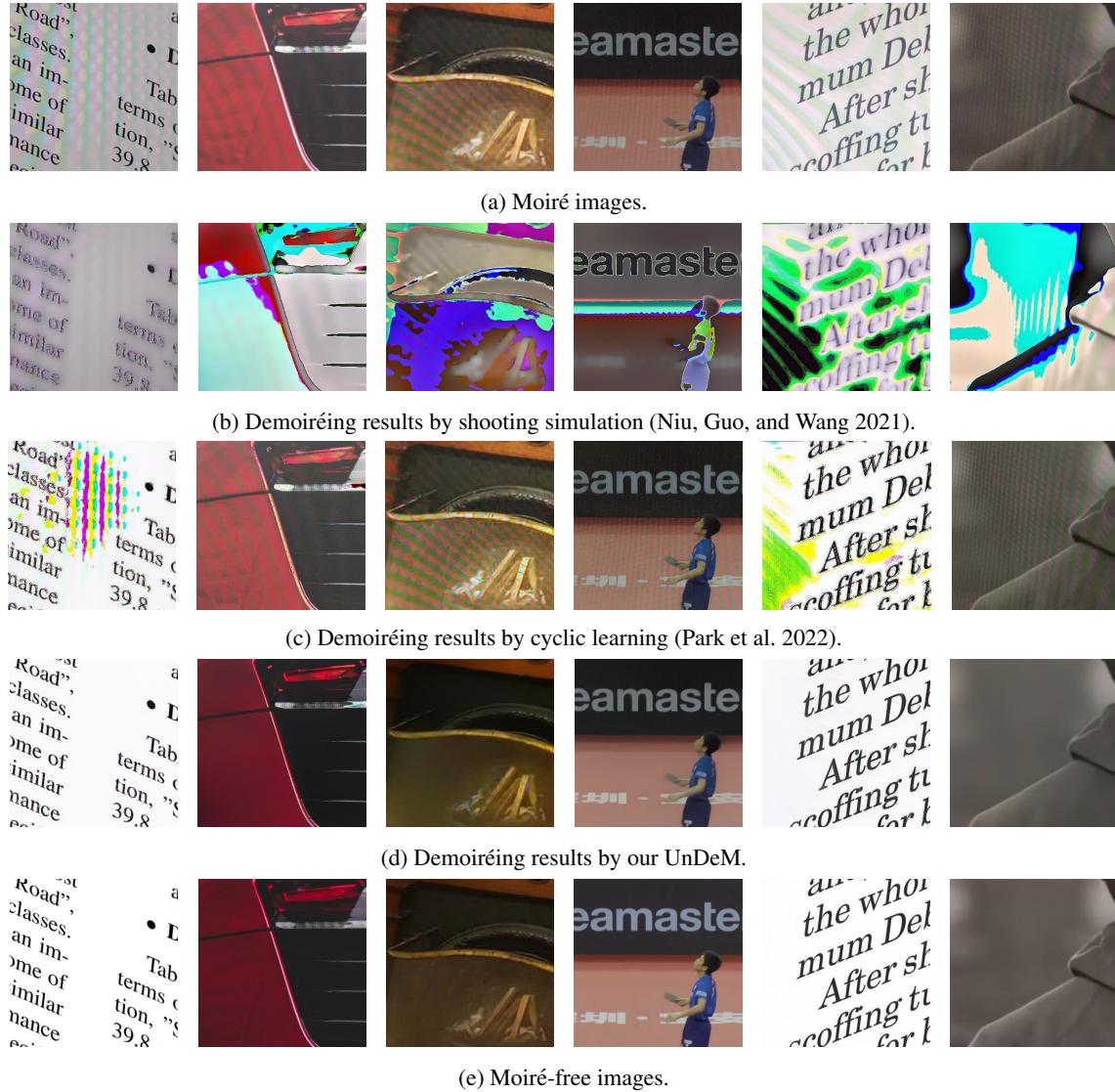
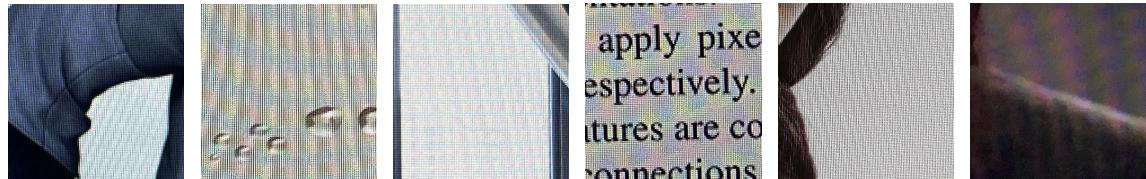


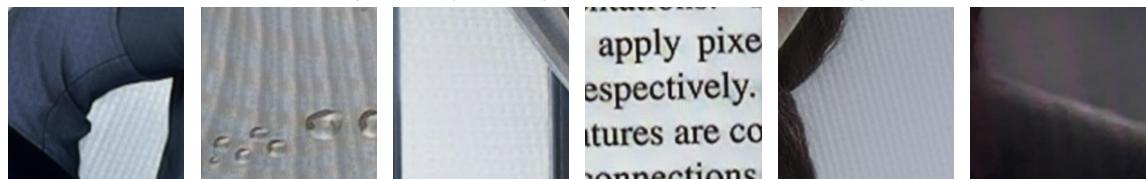
Figure 10: Visualization of demoiréing results of EDSNet-L (Crop Size: 384) on the FHDMi dataset. For the convenience of demonstration, we crop the patch from the test image.



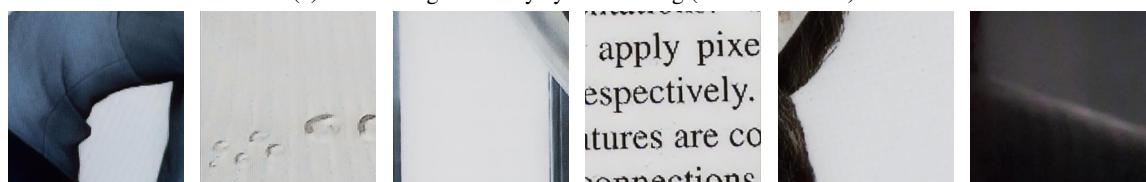
(a) Moiré images.



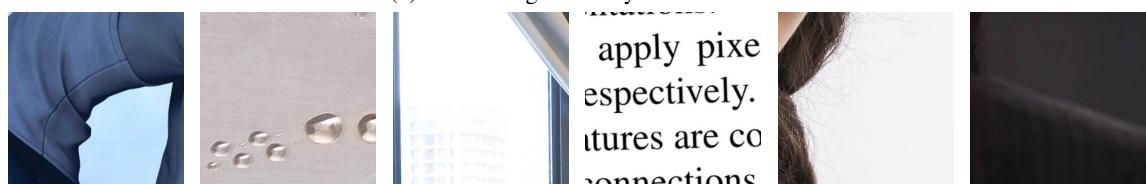
(b) Demoiréing results by shooting simulation (Niu, Guo, and Wang 2021).



(c) Demoiréing results by cyclic learning (Park et al. 2022).



(d) Demoiréing results by our UnDeM.



(e) Moiré-free images.

Figure 11: Visualization of demoiréing results of EDSNet-L (Crop Size: 768) on the UHDM dataset. For the convenience of demonstration, we crop the patch from the test image.