

Fabric Image Demoiréing Benchmark from Synthesis to Restoration

Pengchao Wei and Xiaojie Guo[✉]

Tianjin University, Tianjin, China

{pcwei.ai,xj.max.guo}@gmail.com

Project page: <https://weipengchao.top/PRISM-page>

Abstract. Fabric moiré is a sampling-induced aliasing artifact caused by the interaction between fine textile patterns and camera sensor grids, producing structured interference that severely degrades image quality. Unlike screen-induced moiré, which stems from strictly periodic display lattices, fabric moiré is intrinsically more challenging due to the broadband and semi-periodic nature of textile weaves. The heavy spectral overlap between intrinsic texture and aliasing components renders fabric demoiréing substantially more ill-posed. Consequently, existing models trained on screen moiré datasets generalize poorly to these complex textile patterns. Despite its practical importance, fabric image demoiréing remains underexplored and lacks standardized benchmarks. We present the first comprehensive benchmark for fabric image demoiréing. To address the difficulty of acquiring pixel-aligned real-world pairs, we develop a physically motivated synthesis framework and construct a large-scale dataset comprising 16,050 paired multi-resolution fabric images with controllable aliasing severity. Furthermore, we customize a baseline model, which establishes promising performance on the proposed benchmark dataset with strong generalization ability. Our benchmark provides a standardized platform for advancing research in fabric image demoiréing.

Keywords: Fabric moiré · Image demoiréing · Image restoration

1 Introduction

When a discrete imaging array samples high-frequency scene content below the Nyquist–Shannon limit [21, 25], spectral folding occurs, causing high-frequency components to alias into the baseband. The resulting interference manifests as structured patterns, such as periodic stripes, ripples, and chromatic distortions, commonly referred to as moiré artifacts. Formally, moiré arises from spectral overlap between intrinsic scene frequencies and sampling-induced replicas. When this overlap intersects with genuine image content, signal and artifact become spectrally entangled, rendering inverse recovery fundamentally ill-posed.

Fabrics are among the most challenging real-world sources of moiré. Textile weaves contain dense, anisotropic, and semi-periodic micro-patterns (*e.g.*,

✉ indicates corresponding author.



Fig. 1: Cause-to-appearance domain gap between screen moiré and fabric moiré. (a) Sampling sources: screen moiré originates from rigid planar displays with regular pixel/subpixel lattices, while fabric moiré arises on non-rigid, deformable garments with anisotropic weave microstructures, undergoing folds, stretching, and pose variations. (b) Sampling outcomes: screen moiré is typically global and content-agnostic, whereas fabric moiré is localized to high-frequency textile regions and strongly coupled with texture and deformation. (c) Domain gap: these differences hinder transferring screen-trained models, leading to moiré residual or over-smoothing.

fine stripes, grids, herringbone, and tight weaves), whose spectra are broadband and often approach the sensor’s Nyquist boundary. Unlike screen-induced moiré, which originates from strictly periodic display lattices with discrete spectral peaks, fabric textures have distributed frequency support. Consequently, aliasing overlaps heavily with intrinsic texture, producing strong spectral entanglement and substantially increasing restoration ambiguity. Additionally, fabrics are non-rigid surfaces that exhibit wrinkles, deformations, and pose variations, further complicating demoiréing.

Recent learning-based demoiréing methods have made significant progress in screen-camera scenarios, owing to paired datasets and architectures designed for periodic display patterns. Early works leveraged multi-scale decompositions to address the cross-frequency characteristics of moiré [3, 30], while subsequent methods incorporated structural priors and frequency-domain modeling, such as feature fusion for moiré attributes [6], wavelet-domain processing [14], and multi-band modeling [44]. While in high-resolution settings, designs focusing on efficiency and scale robustness have also emerged [7, 12, 34, 39]. However, these methods assume spectral separability between content and artifact, which fails in fabrics. Directly transferring screen-oriented models often results in over-smoothing, texture removal, or incomplete artifact suppression (see Fig. 1, which shows examples of removing fabric moiré using UHDM-trained ESDNet [39]),

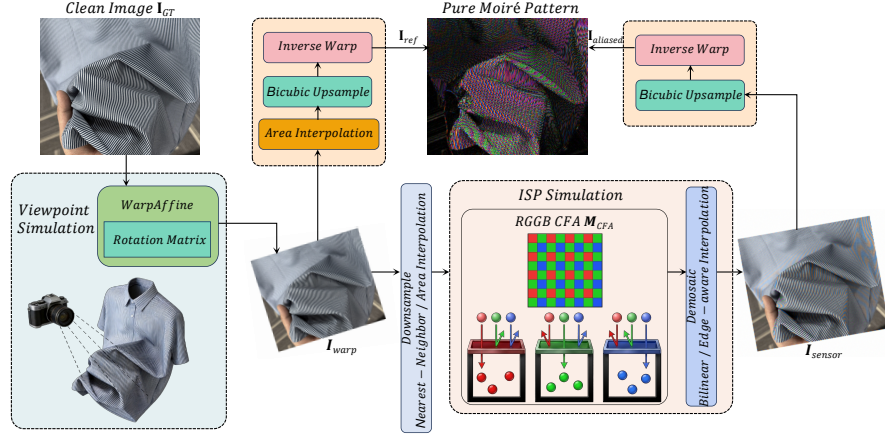


Fig. 2: Our synthesis pipeline.

demonstrating a fundamental domain gap. Moreover, supervised demoiréing is often sensitive to pixel-level alignment. Misalignment between moiré input images and their corresponding clean targets may cause unstable optimization and significantly degrade restoration quality. Even in the “screen is flat” ideal condition, constructing pixel-aligned data requires complex registration engineering. For example, the TIP2018 dataset designs high-contrast black borders and corner structures around the reference images and performs registration using homography [30]; the UHDM 4K dataset employs RANSAC [5] to estimate homography and manual selection to remove severely misaligned samples [39]. In addition, the FHDMi dataset acknowledges 5–10 pixel shifts caused by nonlinear camera distortions and adopts a learning-based tolerance strategy by employing the Contextual Bilateral (CoBi) loss [43] to match features within a local neighborhood [7]. More recent work further demonstrates that additional feature-matching-based alignment (such as SIFT [17]) on UHDM considerably improves training stability and restoration quality, highlighting the decisive role of alignment in demoiréing tasks [34]. While for fabric scenes, this alignment issue is further magnified, because non-rigid deformations and subtle pose changes make global alignment infeasible, thus preventing practical large-scale collection.

To overcome the above challenge, we introduce PRISM (Physics-based Residual Injection for Synthetic Moiré), a synthesis framework that models the imaging chain and generates paired fabric moiré with precise pixel-level alignment. PRISM extracts aliasing residuals via a round-trip construction and injects them into the original pixel grid, preserving genuine textile textures while simulating realistic aliasing. Using a one-to-many strategy, we generate a large-scale multi-resolution dataset of 16,050 fabric images covering diverse levels of aliasing. Additionally, a small real-world unpaired dataset is also collected to validate zero-shot generalization. Besides, building on the PRISM benchmark dataset, we further customize a fabric demoiré network, namely FaDeNet, to faithfully restore fabrics. FaDeNet explicitly decomposes the input into low-frequency base

and high-frequency detail components. The low-frequency branch uses a multi-scale U-shaped trunk [24] to predict residual corrections and a spatial confidence mask, and applies corrections only in the required regions using a mask-gated update rule. In parallel, a lightweight detail branch performs magnitude-bounded refinement on the high-frequency component under the same mask to suppress moiré interference while avoiding excessive modification of textile textures. Furthermore, we propose Spectral-Anisotropic Gated Blocks (SAGB) to capture anisotropic stripe structures in the spatial domain and narrow-band periodic cues in the frequency domain, enabling adaptive correction strength in moiré-dominant regions while suppressing unnecessary modifications in clean textures.

Our major contributions are summarized as follows:

- **First dedicated fabric demoiré benchmark.** We systematically investigate fabric image demoiré in real-world photography and introduce PRISM, a physics-based residual injection synthesis framework that generates realistic fabric moiré while guaranteeing perfect pixel-level alignment and preserving intrinsic textile textures. Based on PRISM, we construct a large-scale, multi-resolution paired dataset and validate its realism and transferability through zero-shot evaluation on real unpaired fabric images.
- **Benchmarking existing methods and a fabric-tailored model.** We benchmark representative screen-camera demoiré architectures on PRISM and further introduce FaDeNet, a conservative restoration network with base/detail decomposition, mask-gated low-frequency correction, magnitude-bounded detail refinement, and spectral-anisotropic gated blocks. FaDeNet achieves state-of-the-art performance on PRISM across PSNR, SSIM, and LPIPS, with low computational overhead, establishing competitive and reproducible baselines for future research.

2 Related Work

2.1 Screen-Camera Image Demoiréing

Moiré artifacts in screen-camera imaging arise from frequency aliasing between the display subpixel lattice and the camera sampling process. Before end-to-end learning became dominant, demoiréing was mostly studied through hand-crafted priors. Sidorov and Kokaram [28] proposed suppressing telecine-related moiré via spectral analysis. Sur and Grédiac [31] addressed periodic/quasi-periodic noise using frequency-domain statistics and automated notch-filter design. Liu et al. [13] formulated moiré removal as a patch-wise layer decomposition, regularizing textures with low-rank priors. Yang et al. [37] further studied textured image demoiréing via signal decomposition and guided filtering. Meanwhile, descreening methods for scanned halftone prints (*e.g.*, training-based descreening [27] and wavelet-based filtering [18]) also constitute an important line of early exploration. Subsequent learning-based methods typically adopt supervised training on paired moiré/clean images. Sun et al. [30] introduced DM-CNN, explicitly leveraging multi-resolution feature maps to remove moiré ar-

tifacts across different frequency bands. Later works further improved restoration quality through dynamic multi-scale feature encoding [3], property-oriented demoiréing with moiré-type cues [6], and frequency-aware modeling. In particular, Zheng et al. [44] formulated demoiréing as the combination of moiré texture removal and color/tone correction, and introduced learnable bandpass filters to better capture moiré-related frequency priors. Liu et al. [14] further exploited wavelet decomposition for targeted artifact suppression. To handle 4K resolutions, He et al. [7] proposed FHDe²Net, a cascaded global-to-local pipeline. Yu et al. [39] introduced the UHDM dataset and the lightweight ESD-Net with scale-aware fusion. For UHD deployment, Xiao et al. [34] developed P-BiC for patch-level restoration, while Lee et al. [12] explored large-kernel CNNs (MZNNet) to capture long-range structures. Recent works also explore RAW-domain demoiréing for screen recapture, showing that RAW observations can provide useful information before ISP-induced color and frequency mixing [35, 41]. In contrast to these screen-camera studies, we investigate fabric demoiréing and introduce PRISM and FaDeNet for texture-faithful restoration.

2.2 Fabric Demoiréing

While various moiré synthesis schemes have been proposed [22, 36, 38, 40, 45], they are mainly designed for screen-camera degradations, where moiré arises from the aliasing between display sub-pixel grids and camera sampling processes. These methods typically rely on the implicit assumption that moiré can be modeled as a content-independent foreground layer. However, this assumption is physically implausible for textiles, as fabric moiré is coupled with the intrinsic textile micro-structures. Consequently, demoiréing models optimized for screen-camera scenarios tend to over-smooth authentic textures or fail to suppress interference that is tightly entangled with the textile geometry. To our knowledge, research specifically addressing fabric moiré remains sparse. The most relevant work is by Liu et al. [15], which leverages focused/defocused observations for learning and highlights that screen-oriented data pipelines are ill-suited for high-texture demoiréing. Nevertheless, their approach necessitates a specialized hardware setup and a rigid acquisition pipeline to obtain paired inputs, posing significant challenges for practical data collection at scale. Compared with the above methods, we aim to enable a stable and scalable paired supervision for fabric demoiréing.

3 PRISM: Benchmark Dataset Construction

We investigate fabric image demoiréing under paired supervision. As illustrated in Fig. 1, due to significant differences in the physical microstructures of the sampling sources, the data distribution of screen-camera moiré diverges from that of fabric moiré. Specifically, fabric moiré patterns are more localized within the highly textured regions of fabric and exhibit greater pattern complexity. They tend to distort, degrade, or become heavily coupled with the intrinsic textures

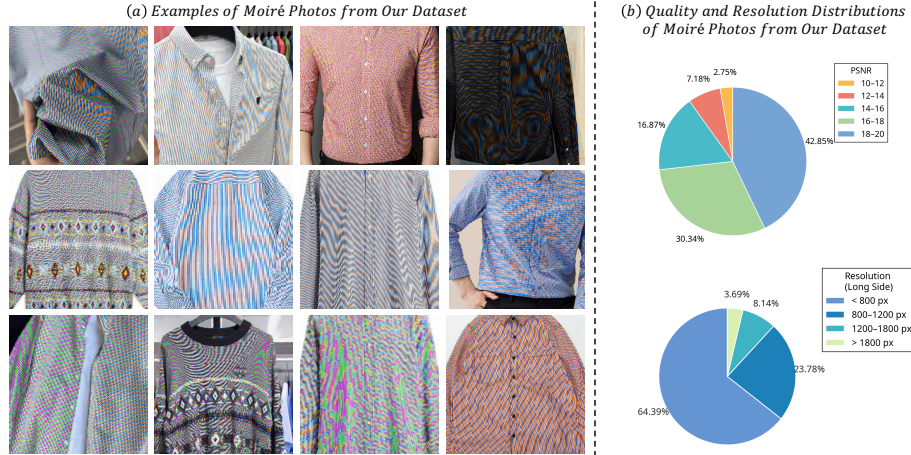


Fig. 3: Overview of the PRISM dataset. (a) Representative synthesized moiré images in PRISM. (b) Distributions of image quality and resolution.

of the fabric, making them more difficult to remove. Taking ESDNet [39] as an example, when the model is trained on the screen-camera dataset UHDM [39], it fails to effectively remove fabric moiré patterns and often suffers from unintended eradication of fabric textures. Therefore, a dedicated fabric moiré dataset is necessary. However, acquiring large-scale, pixel-aligned real-world fabric moiré datasets is costly and technically challenging.

From an alignment perspective, factors such as non-linear lens distortion, capture perturbations, and the interference of moiré patterns themselves introduce persistent alignment errors. In datasets like TIP18 [30], FHDMi [7], and UHDM, even with sophisticated registration engineering as we discussed in Sec. 1, noticeable alignment issues still exist in the data. Moreover, from a data acquisition perspective, screen-camera moiré pairs can be collected by displaying Internet images on a monitor and re-photographing the screen, where the original images serve as GT after standard processing. In contrast, fabric moiré presents a unique challenge as it must be captured from physical garments with high-frequency textures. The primary bottleneck lies in the inability to obtain perfectly aligned, moiré-free ground truth (GT) images from the identical camera perspective. This lack of pixel-perfect pairs makes the acquisition of high-quality training data a difficult task, significantly hindering the application of standard supervised learning methods. Motivated by these constraints, we propose PRISM (Physics-based Residual Injection for Synthetic Moiré) and construct a large-scale paired synthetic dataset for stable supervised learning.

3.1 Data Collection and Pair Synthesis

To better match real demoiré scenarios, we restrict fabric materials to high-frequency textures prone to moiré and manually remove samples with pre-existing artifacts. Under these constraints, we collect $\sim 2,000$ fabric images from

the Internet as GT. Based on the synthesis pipeline depicted in Fig. 2, we adopt a one-to-many strategy to better simulate moiré variations under diverse real-world capture conditions. In our release, PRISM contains 16,050 synthesized multi-resolution fabric moiré pairs.

3.2 PRISM Synthesis Pipeline

Problem Formulation The general synthesis workflow is illustrated in Fig. 2. Given a clean fabric image $\mathbf{I}_{\text{GT}} \in \mathbb{R}^{H \times W \times 3}$, PRISM aims to synthesize a moiré-corrupted observation while maintaining pixel-wise alignment with $\mathbf{I}_{\text{GT}}$.

We derive PRISM from a physics-based digital imaging chain. In real captures, moiré artifacts are shaped by multiple factors: (i) capture geometry (viewpoint, in-plane rotation, shooting distance) that changes the relative scale and orientation between the CFA sampling lattice and the sampled objects; (ii) discrete sensor sampling that can undersample high-frequency object content beyond the sensor’s Nyquist limit [21], causing spectral folding; and (iii) sensor photoelectric effects and ISP (CFA mosaicking, sensor noise, demosaicking [19], and color rendering) that modulate the visual appearance of moiré, especially chromatic fringes. A simplified imaging chain can be written as

$$\mathbf{I}[\mathbf{n}] = \mathcal{R}_{\text{ISP}} \left(\mathcal{S}_{\Delta} \left(\mathcal{T}_{\theta,s} (L(\mathbf{x})) \right) \right) + \eta, \quad (1)$$

where $\mathbf{I}[\mathbf{n}]$ is the discrete RGB image sampled on the pixel lattice indexed by $\mathbf{n} \in \mathbb{R}^2$, $L(\mathbf{x})$ denotes the continuous light field defined over spatial coordinate $\mathbf{x} \in \mathbb{R}^2$, $\mathcal{T}_{\theta,s}$ is a geometric projection parameterized by observation angle θ and distance s , $\mathcal{S}_{\Delta}$ is a discretization operator that samples and integrates the projected signal onto a sensor grid with pitch Δ , $\mathcal{R}_{\text{ISP}}$ denotes an ISP-like rendering pipeline, η models photoelectronic noise.

Although Eq. (1) characterizes the physical formation of moiré patterns, directly employing this process for image synthesis leads to an irreversible loss of high-frequency details due to the low-pass filtering nature of the discretization operator $\mathcal{S}_{\Delta}$. Such degradation hampers the model to learn the restoration of fine-grained textures. To provide the high-fidelity texture references, PRISM avoids reconstructing the entire degraded signal field. Instead, it adopts the residual injection strategy as:

$$\mathbf{I}_{\text{Moiré}} = \mathbf{I}_{\text{GT}} + \gamma \cdot \Delta\Phi(\mathbf{I}_{\text{GT}}; \theta, s), \quad (2)$$

where γ controls the injection strength. By superimposing the physical moiré residual field $\Delta\Phi$ onto $\mathbf{I}_{\text{GT}}$, we ensure that the underlying texture integrity is perfectly preserved. This strategy constructs an ideal training pair of clean texture, forcing the network to precisely decouple artifacts from complex backgrounds.

Imaging-Chain Simulation We generate moiré patterns by simulating the camera capture process, which includes the viewpoint relationship relative to the object and the subsequent ISP pipeline.

Viewpoint Simulation. We simulate the viewpoint relationship between the fabric and the sensor through a geometric manifold projection $\mathcal{T}_{\theta,s}$.

$$\mathbf{I}_{\text{warp}} = \mathcal{T}_{\theta,s}(\mathbf{I}_{\text{GT}}). \quad (3)$$

where the observation angle θ is sampled uniformly from $[-25^\circ, 25^\circ]$ and the distance factor s follows a Beta distribution $\mathcal{B}(1.2, 1.2)$ rescaled to $[0.25, 0.85]$.

Following the geometric warping, we apply a discrete sampling operator $\mathcal{S}_\downarrow$ to induce spectral folding:

$$\mathbf{I}_{\text{tiny}} = \mathcal{S}_\downarrow(\mathbf{I}_{\text{warp}}). \quad (4)$$

To further approximate the real ISP, $\mathbf{I}_{\text{tiny}}$ is subsampled according to an RGGB Bayer mosaic $\mathbf{M}_{\text{CFA}}$ with additive photon shot noise $\mathcal{N}(0, \sigma^2)$:

$$\mathbf{R}_{\text{raw}} = (\mathbf{I}_{\text{tiny}} \odot \mathbf{M}_{\text{CFA}}) + \mathcal{N}(0, \sigma^2). \quad (5)$$

ISP Simulation. Finally, a demosaic operator $\mathcal{R}_{\text{dm}}(\cdot)$ is employed. As implemented in our PRISM framework, we randomly select between standard bilinear and edge-aware interpolation to yield the RGB sensor observation $\mathbf{I}_{\text{sensor}}$:

$$\mathbf{I}_{\text{sensor}} = \mathcal{R}_{\text{dm}}(\mathbf{R}_{\text{raw}}). \quad (6)$$

This physics-based pipeline ensures that the synthesized moiré patterns exhibit realistic chromatic fringes and structural distortions.

Round-trip Alignment The conventional forward imaging chain (Eq. 1) is inherently destructive, as the discrete sampling operator $\mathcal{S}_\Delta$ irreversibly filters out textile details. To circumvent this, PRISM isolates pure moiré artifacts by contrasting two geometry-matched branches.

Aliased path. We upsample the rendered sensor output and map it back to the canonical high-resolution coordinate system:

$$\mathbf{I}_{\text{aliased}} = \mathcal{T}_{\theta,s}^{-1}(\text{interp}_\uparrow(\mathbf{I}_{\text{sensor}})), \quad (7)$$

where $\text{interp}_\uparrow$ is implemented via bicubic upsampling.

Reference Path. As an anti-aliased reference, we apply Nyquist-compliant down-sampling $\mathcal{S}_\downarrow^{\text{LPF}}$ via area interpolation to $\mathbf{I}_{\text{warp}}$:

$$\mathbf{I}_{\text{ref}} = \mathcal{T}_{\theta,s}^{-1}(\text{interp}_\uparrow(\mathcal{S}_\downarrow^{\text{LPF}}(\mathbf{I}_{\text{warp}}))). \quad (8)$$

By keeping the same operators as Eq. (7), $\mathbf{I}_{\text{ref}}$ effectively isolates the bias introduced by pure geometric resampling without introducing aliasing components.

Moiré extraction. The pure moiré residual field $\Delta\Phi$ is extracted by:

$$\Delta\Phi(\mathbf{I}_{\text{GT}}; \theta, s) = \mathbf{I}_{\text{aliased}} - \mathbf{I}_{\text{ref}}. \quad (9)$$

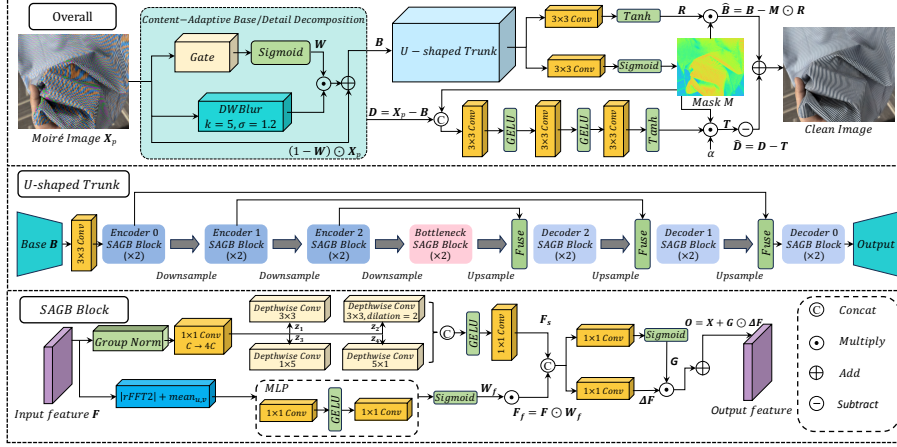


Fig. 4: Pipeline of our method.

4 FaDeNet: Proposed Fabric Demoiréing Network

We propose FaDeNet for fabric demoiréing. As illustrated in Fig. 4, the overall architecture starts with a content-adaptive base/detail decomposition, followed by two correction branches with a shared spatial confidence mask.

Given an input moiré image $\mathbf{X} \in \mathbb{R}^{H \times W \times 3}$, we first pad it to a multiple of 8 for pixel-rearrangement down/up-sampling: $\mathbf{X}_p = \text{Pad}_8(\mathbf{X})$. We then compute a content-adaptive base/detail split $(\mathbf{B}, \mathbf{D})$, where the base captures low-frequency variations (e.g., color/illumination shifts) and the detail keeps high-frequency structures (e.g., edges/textures).

Our 4-scale U-shaped trunk operates only on the base component and predicts (i) a base residual $\mathbf{R}$ and (ii) a spatial confidence mask $\mathbf{M} \in (0, 1)^{H \times W \times 1}$. Scale changes are implemented by PixelUnshuffle/PixelShuffle ($\times 2$) [26] to reduce aliasing. The base is updated by a masked conservative rule:

$$\hat{\mathbf{B}} = \mathbf{B} - \mathbf{M} \odot \mathbf{R}. \quad (10)$$

This enforces corrections only where the mask indicates moiré-dominant regions, improving stability and avoiding over-modification in clean areas.

To mitigate over-smoothing, we further apply a lightweight, mask-gated and magnitude-controlled correction on the high-frequency detail:

$$\hat{\mathbf{D}} = \mathbf{D} - \alpha \mathbf{M} \odot \tanh(\mathcal{R}([\mathbf{D}, \mathbf{M}])), \quad (11)$$

where $\mathcal{R}(\cdot)$ is a small residual predictor, and α is an empirically selected magnitude bound for controlling the detail correction strength. We set $\alpha=0.35$ by default to balance moiré suppression and texture preservation.

Finally, the restored output is recomposed and cropped back:

$$\hat{\mathbf{Y}} = \text{Crop}(\hat{\mathbf{B}} + \hat{\mathbf{D}}). \quad (12)$$

4.1 Content-Adaptive Base/Detail Decomposition

Moiré artifacts typically manifest as low-frequency color shifts coupled with high-frequency periodic interference. Instead of a fixed blur, we employ a content-adaptive depthwise Gaussian smoothing to split frequency bands:

$$\mathbf{W} = \sigma(\text{Gate}(\mathbf{X}_p)), \mathbf{B} = \mathbf{W} \odot \text{Blur}(\mathbf{X}_p) + (1 - \mathbf{W}) \odot \mathbf{X}_p, \mathbf{D} = \mathbf{X}_p - \mathbf{B}, \quad (13)$$

where $\mathbf{W}$ is a single-channel gating map; $\text{Blur}(\cdot)$ is depthwise Gaussian conv [9] with $(k=5, \sigma=1.2)$ and $\text{Gate}(\cdot)$ is a lightweight convolution stack.

4.2 Spectral-Anisotropic Gated Block (SAGB)

Moiré patterns typically manifest as anisotropic spatial stripes and narrow-band frequency peaks. SAGB is designed to capture both with low overhead.

Spatial anisotropic branch. Let $\mathbf{F}$ denote the input feature map of SAGB. Given the normalized features $\tilde{\mathbf{F}} = \text{GN}(\mathbf{F})$, we apply a $4C$ expansion followed by multi-branch anisotropic depthwise filtering, and finally fuse the features by:

$$\mathbf{F}_s = \mathcal{P}\left(\phi\left([\text{DW}_{3 \times 3}(\mathbf{Z}_1), \text{DW}_{3 \times 3}^{d=2}(\mathbf{Z}_2), \text{DW}_{1 \times 5}(\mathbf{Z}_3), \text{DW}_{5 \times 1}(\mathbf{Z}_4)]\right)\right), \quad (14)$$

where $[\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3, \mathbf{Z}_4]$ is a channel-wise split of the expanded feature $\text{Conv}_{1 \times 1}(\tilde{\mathbf{F}})$ (from C to $4C$), $\mathcal{P}$ denotes point-wise (1×1) conv and ϕ denotes GELU [8].

Spectral gate Prior work has shown that leveraging frequency-domain representations can benefit image restoration [4, 10]. To sense periodic interference, we compute a channel-wise gate from the 2D real FFT magnitude [2]:

$$\mathbf{W}_f = \sigma\left(\text{MLP}\left(\text{mean}_{u,v}(|\text{rFFT2}(\mathbf{F})|)\right)\right), \quad \mathbf{F}_f = \mathbf{F} \odot \mathbf{W}_f. \quad (15)$$

For efficiency, the spectral gate is applied only at coarse scales $(1/4, 1/8)$.

Gated residual update. We concatenate $\mathbf{C} = [\mathbf{F}_s, \mathbf{F}_f]$ and predict an update:

$$\mathbf{G} = \sigma(\text{Conv}_{1 \times 1}(\mathbf{C})), \quad \Delta\mathbf{F} = \text{Conv}_{1 \times 1}(\mathbf{C}), \quad \mathbf{O} = \mathbf{F} + \mathbf{G} \odot \Delta\mathbf{F}. \quad (16)$$

This formulation allows stronger corrections in moiré-dominant regions while suppressing unnecessary updates in clean, texture-rich areas.

4.3 Loss Function

We minimize an ℓ_1 loss with a Laplacian high-pass constraint [20, 29]:

$$\mathcal{L} = \|\hat{\mathbf{Y}} - \mathbf{Y}\|_1 + \lambda_{\text{hp}} \cdot \|\text{Lap}(\hat{\mathbf{Y}}) - \text{Lap}(\mathbf{Y})\|_1, \quad (17)$$

where $\lambda_{\text{hp}} = 0.3$ is set empirically, and $\text{Lap}(\cdot)$ denotes a Laplacian operator.

Table 1: Performance and efficiency comparison on PRISM. Bold and underlined values indicate the best and second-best results for PSNR/SSIM/LPIPS, respectively.

Method	Venue	Quality			Efficiency		
		PSNR $\uparrow$	SSIM $\uparrow$	LPIPS $\downarrow$	GFLOPs	Params (M)	FPS
Input	–	16.998	0.7268	0.3495	–	–	–
DMCNN [30]	TIP 2018	24.659	0.9307	0.0772	17.648	1.426	164.85
MDDM [3]	ICCVW 2019	24.916	0.9260	0.0755	57.489	8.009	17.74
WDNet [14]	ECCV 2020	26.141	0.9444	0.0524	34.195	3.919	65.00
MBCNN [44]	CVPR 2020	27.911	0.9617	0.0405	133.754	14.192	68.41
MopNet [6]	ICCV 2019	28.190	0.9648	0.0362	408.037	60.228	10.05
FHDe ² Net [7]	ECCV 2020	28.243	0.9654	0.0343	561.916	13.599	55.30
MZNet [12]	arXiv 2025	27.504	0.9598	0.0408	19.374	7.276	21.52
P-BiC [34]	ACM MM 2024	28.590	0.9689	0.0334	17.304	4.922	38.17
ESDNet-L [39]	ECCV 2022	<u>28.809</u>	<u>0.9711</u>	<u>0.0268</u>	129.674	10.623	19.22
FaDeNet (Ours)	–	32.159	0.9859	0.0169	133.681	7.051	38.10

5 Experiments

5.1 Experimental Setting

Datasets and Metrics. We perform benchmark evaluations on the proposed PRISM dataset, and we report PSNR, SSIM [33] and LPIPS [42] for quantitative comparison. In addition to restoration quality, we further report efficiency metrics, including computational cost (GFLOPs), model size (Params), and inference throughput (FPS) on PRISM. To assess real-world robustness, we further evaluate zero-shot transfer on unpaired fabric captures and report no-reference perceptual metrics on the real-world fabric moiré set, including ARNIQA [1] and CLIP-IQA+ (ViT-L/14, 512) [23, 32], complemented by human mean opinion score (MOS) ratings.

Implementation Details. All experiments were conducted on a single NVIDIA RTX 4090 GPU. During training, we randomly cropped 384×384 patches with a batch size of 4 and optimized the model for 200 epochs using Adam [11] ($\beta_1=0.9$, $\beta_2=0.999$). We set the initial learning rate to 2×10^{-4} and applied a cyclic cosine annealing schedule [16]. For efficiency comparisons, we benchmarked all methods using a fixed-resolution input of 384×384 . All compared methods were retrained from scratch on the PRISM dataset to ensure their optimal performance.

5.2 Comparison with Screen Demoiréing Methods on PRISM

Quantitative results. Tab. 1 reports quantitative comparisons on PRISM. FaDeNet achieves the best performance across all three quality metrics. The large PSNR margin indicates that FaDeNet more effectively restores structural fidelity under strong moiré interference, while the lower LPIPS suggests better perceptual alignment, particularly important for fabric images with dense and repetitive textures.

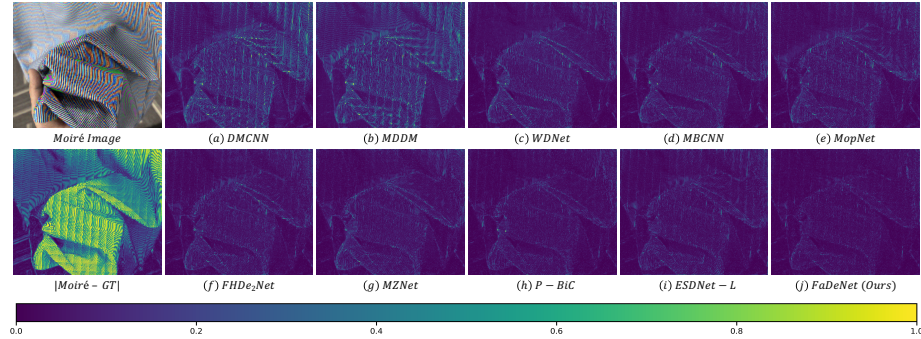


Fig. 5: Visual comparison of error maps on the PRISM dataset. The color transition from dark purple to bright yellow represents increasing absolute errors ($|\text{Output} - \text{GT}|$).

Qualitative results. Fig. 5 visualizes representative examples with absolute error maps. Compared with prior methods, FaDeNet produces the cleanest residual maps with fewer bright error regions, implying more complete removal of moiré components. These visual results are consistent with the quantitative gains, confirming that FaDeNet improves structural accuracy and visual realism simultaneously.

Computational cost analysis. Beyond restoration quality, FaDeNet maintains a competitive balance between architectural complexity and inference speed. As shown in Table 1, FaDeNet is relatively compact with 7.05M parameters, fewer than the second-best performing model ESDNet-L (10.62M). In terms of throughput, FaDeNet achieves 38.10 FPS, nearly doubling the speed of ESDNet-L (19.22 FPS) despite having a similar GFLOPs count. These metrics suggest that FaDeNet provides a favorable trade-off in efficiency and quality.

5.3 Generalization to Real-World Fabric Demoiréing

To evaluate real-world robustness, we directly apply PRISM-trained models to an unpaired real-world fabric moiré dataset without any fine-tuning. As shown in Fig. 6, PRISM-trained models produce visually meaningful results, effectively suppressing moiré artifacts while preserving plausible texture structures. To further quantify this zero-shot generalization, we report no-reference perceptual quality scores and human subjective results in Table 2.

No-reference perceptual quality. FaDeNet achieves the highest ARNIQA score, indicating superior distortion suppression and improved perceptual quality on real-world inputs. It also attains a competitive CLIP-IQA+ score, suggesting that its restorations remain visually natural and semantically consistent.

Table 2: No-reference perceptual quality comparison and blind MOS user study results on the real-world unpaired fabric moiré dataset. Bold and underlined values indicate the best and second-best results, respectively.

Method	No-Reference IQA		User Study (MOS)				
	ARNIQA $\uparrow$	CLIP-IQA+ $\uparrow$	MOS $\uparrow$	Std.	95% CI $\downarrow$	Avg. Rank $\downarrow$	Corr. p
DMCNN [30]	0.59850	0.43023	–	–	–	–	–
MDDM [3]	0.62414	0.42228	–	–	–	–	–
WDNet [14]	0.63115	0.43618	–	–	–	–	–
MBCNN [44]	0.62612	0.44005	–	–	–	–	–
FHDe ² Net [7]	0.62763	0.44991	–	–	–	–	–
MopNet [6]	<u>0.63358</u>	0.45106	2.86	1.21	± 0.14	3.30	2.64×10^{-8}
MZNet [12]	0.62660	0.44923	3.11	1.13	± 0.13	<u>2.63</u>	8.25×10^{-4}
P-BiC [34]	0.63327	0.48329	<u>3.17</u>	1.22	± 0.14	2.90	8.18×10^{-3}
ESDNet-L [39]	0.63001	0.45139	2.79	1.13	± 0.13	4.20	2.91×10^{-10}
FaDeNet (Ours)	0.63364	<u>0.47473</u>	3.39	1.07	± 0.12	1.97	–

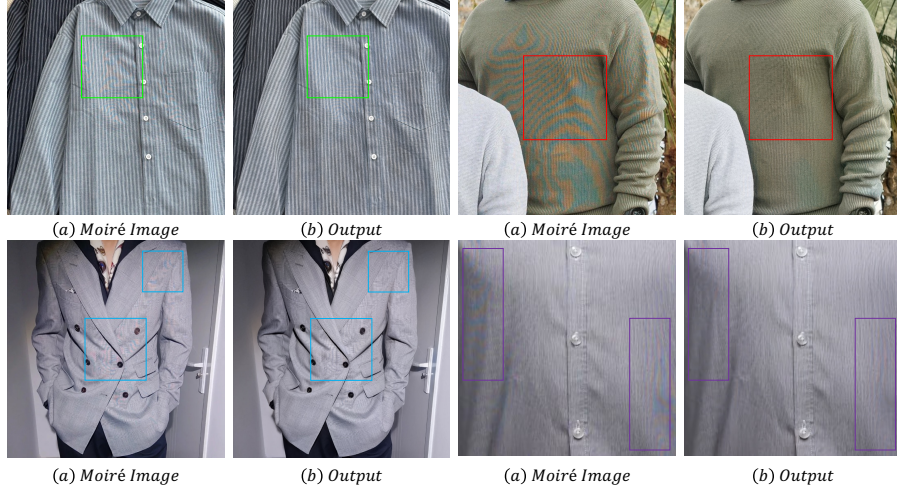


Fig. 6: Qualitative results on real-world cases by models trained only on PRISM.

Blind MOS user study. To further assess subjective visual appeal, we conducted a blind Mean Opinion Score (MOS) study. We selected five representative models and evaluated them across 15 groups of representative real-world images. The study involved 20 participants, including 10 domain experts in image processing and 10 volunteers. As shown in Table 2, FaDeNet outperforms other methods in human preference, achieving the best MOS of 3.39 and a leading Average Rank of 1.97. This indicates that FaDeNet provides reliable restoration quality across diverse real-world scenarios. To assess statistical significance, we further conduct paired Wilcoxon signed-rank tests between FaDeNet and each MOS baseline, followed by Holm-Bonferroni correction. The corrected p values in Table 2 are all below 0.01, indicating statistically significant preference for FaDeNet.

Table 3: Ablation results of FaDeNet on PRISM.

Configuration	PSNR $\uparrow$	SSIM $\uparrow$	LPIPS $\downarrow$
(A) w/o Decomposition	31.560	0.9837	0.0186
(B) w/o Mask ($\mathbf{M} \equiv 1$)	31.599	0.9840	0.0185
(C) w/o SAGB (param.-matched ConvBlock)	31.349	0.9831	0.0191
(D) SAGB w/o spectral gate	31.541	0.9837	0.0189
(E) SAGB w/o spatial anisotropic branch	31.355	0.9829	0.0194
(F) ℓ_1 loss only	31.777	0.9843	0.0183
Full (FaDeNet)	32.159	0.9859	0.0169

5.4 Ablation Studies

Base/Detail Decomposition. Configuration (A) removes the base/detail decomposition and the detail branch. The results in Tab. 3 indicate that separating low-frequency structure and high-frequency fabric details is beneficial.

Mask Gating. Configuration (B) disables mask gating by setting $\mathbf{M} \equiv 1$. The performance degrades in (B), suggesting that mask gating and spatially selective updates are important for real fabric demoiréing.

SAGB Design. Configurations (C–E) validate the SAGB’s architecture: replacing it with a standard ConvBlock or removing the spectral gate and spatial branch causes a consistent drop in performance. This confirms that the synergy between frequency-domain gating and directional spatial modeling is vital for addressing the complex, periodic nature of moiré patterns.

Training Objective. Configuration (F) employs ℓ_1 loss only. The results in Tab. 3 show a performance drop, particularly in LPIPS, indicating that pixel-wise supervision alone is insufficient for high-frequency fidelity.

6 Conclusion

In this paper, we have systematically investigated the underexplored problem of fabric image demoiréing. We identify that the fundamental performance bottleneck of existing methods lies in the spectral entanglement between intrinsic textile micro-patterns and sampling-induced aliasing. To bridge this gap, we proposed PRISM, a physics-based residual injection pipeline that synthesizes realistic fabric moiré while guaranteeing perfect pixel-wise alignment and preserving native fabric details. We further introduced FaDeNet, a fabric demoiréing architecture with base/detail decomposition and spectral-anisotropic gated blocks for jointly capturing spatial and frequency characteristics of moiré. Extensive experiments show that FaDeNet outperforms representative screen-camera demoiréing baselines on PRISM and generalizes well to real unpaired fabric captures. We hope PRISM facilitates future research on fabric image demoiréing.

Acknowledgements. This work was supported by the National Natural Science Foundation of China under Grant Nos. 62372251 and U2574209.

References

1. Agnolucci, L., Galteri, L., Bertini, M., Del Bimbo, A.: Arnika: Learning distortion manifold for image quality assessment. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. pp. 189–198 (2024)
2. Bracewell, R.N., Bracewell, R.N.: *The Fourier transform and its applications*, vol. 31999. McGraw-hill New York (1986)
3. Cheng, X., Fu, Z., Yang, J.: Multi-scale dynamic feature encoding network for image demoiréing. In: *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*. pp. 3486–3493. IEEE (2019)
4. Chi, L., Jiang, B., Mu, Y.: Fast fourier convolution. *Advances in Neural Information Processing Systems* **33**, 4479–4488 (2020)
5. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* **24**(6), 381–395 (1981)
6. He, B., Wang, C., Shi, B., Duan, L.Y.: Mop moire patterns using mopnet. In: *Proceedings of the IEEE/CVF international conference on computer vision*. pp. 2424–2432 (2019)
7. He, B., Wang, C., Shi, B., Duan, L.Y.: Fhde2net: Full high definition demoireing network. In: *European Conference on Computer Vision*. pp. 713–729. Springer (2020)
8. Hendrycks, D., Gimpel, K.: Gaussian error linear units (gelus). *arXiv preprint arXiv:1606.08415* (2016)
9. Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H.: Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861* (2017)
10. Jiang, X., Zhang, X., Gao, N., Deng, Y.: When fast fourier transform meets transformer for image restoration. In: *European conference on computer vision*. pp. 381–402. Springer (2024)
11. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014)
12. Lee, S., Baek, W., Kim, Y., Kim, E., Moon, H., Yoo, D., Park, E.: Moiré zero: An efficient and high-performance neural architecture for moiré removal. *arXiv preprint arXiv:2507.22407* (2025)
13. Liu, F., Yang, J., Yue, H.: Moiré pattern removal from texture images via low-rank and sparse matrix decomposition. In: *2015 Visual Communications and Image Processing (VCIP)*. pp. 1–4. IEEE (2015)
14. Liu, L., Liu, J., Yuan, S., Slabaugh, G., Leonardis, A., Zhou, W., Tian, Q.: Wavelet-based dual-branch network for image demoiréing. In: *European conference on computer vision*. pp. 86–102. Springer (2020)
15. Liu, L., Yuan, S., Liu, J., Bao, L., Slabaugh, G., Tian, Q.: Self-adaptively learning to demoiré from focused and defocused image pairs. *Advances in Neural Information Processing Systems* **33**, 22282–22292 (2020)
16. Loshchilov, I., Hutter, F.: Sgdr: Stochastic gradient descent with warm restarts. *arXiv preprint arXiv:1608.03983* (2016)
17. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *International journal of computer vision* **60**(2), 91–110 (2004)
18. Luo, J., De Queiroz, R., Fan, Z.: A robust technique for image descreening based on the wavelet transform. *IEEE Transactions on Signal Processing* **46**(4), 1179–1184 (1998)

19. Malvar, H.S., He, L.w., Cutler, R.: High-quality linear interpolation for demosaicing of bayer-patterned color images. In: 2004 IEEE international conference on acoustics, speech, and signal processing. vol. 3, pp. iii–485. IEEE (2004)
20. Marr, D., Hildreth, E.: Theory of edge detection. *Proceedings of the Royal Society of London. Series B. Biological Sciences* **207**(1167), 187–217 (1980)
21. Nyquist, H.: Certain topics in telegraph transmission theory. *Transactions of the American Institute of Electrical Engineers* **47**(2), 617–644 (1928)
22. Park, H., Vien, A.G., Kim, H., Koh, Y.J., Lee, C.: Unpaired screen-shot image demoiréing with cyclic moiré learning. *IEEE Access* **10**, 16254–16268 (2022)
23. Radford, A., Kim, J.W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., et al.: Learning transferable visual models from natural language supervision. In: *International conference on machine learning*. pp. 8748–8763. PmLR (2021)
24. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: *International Conference on Medical image computing and computer-assisted intervention*. pp. 234–241. Springer (2015)
25. Shannon, C.E.: Communication in the presence of noise. *Proceedings of the IRE* **37**(1), 10–21 (2006)
26. Shi, W., Caballero, J., Huszár, F., Totz, J., Aitken, A.P., Bishop, R., Rueckert, D., Wang, Z.: Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 1874–1883 (2016)
27. Siddiqui, H., Bouman, C.A.: Training-based descreeing. *IEEE transactions on image processing* **16**(3), 789–802 (2007)
28. Sidorov, D.N., Kokaram, A.C.: Suppression of moiré patterns via spectral analysis. In: *Visual Communications and Image Processing 2002*. vol. 4671, pp. 895–906. SPIE (2002)
29. Sobel, I., Feldman, G., et al.: A 3x3 isotropic gradient operator for image processing. a talk at the Stanford Artificial Project in **1968**, 271–272 (1968)
30. Sun, Y., Yu, Y., Wang, W.: Moiré photo restoration using multiresolution convolutional neural networks. *IEEE Transactions on Image Processing* **27**(8), 4160–4172 (2018)
31. Sur, F., Grediac, M.: Automated removal of quasiperiodic noise using frequency domain statistics. *Journal of electronic imaging* **24**(1), 013003–013003 (2015)
32. Wang, J., Chan, K.C., Loy, C.C.: Exploring clip for assessing the look and feel of images. In: *Proceedings of the AAAI conference on artificial intelligence*. vol. 37, pp. 2555–2563 (2023)
33. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing* **13**(4), 600–612 (2004)
34. Xiao, Z., Lu, Z., Wang, X.: P-bic: Ultra-high-definition image moiré patterns removal via patch bilateral compensation. In: *Proceedings of the 32nd ACM international conference on multimedia*. pp. 8365–8373 (2024)
35. Xu, S., Song, B., Chen, X., Liu, X., Zhou, J.: Image demoiréing in raw and srgb domains. In: *European conference on computer vision*. pp. 108–124. Springer (2024)
36. Yang, C., Yang, Z., Ke, Y., Chen, T., Grzegorzec, M., See, J.: Doing more with moiré pattern detection in digital photos. *IEEE Transactions on Image Processing* **32**, 694–708 (2023)
37. Yang, J., Liu, F., Yue, H., Fu, X., Hou, C., Wu, F.: Textured image demoiréing via signal decomposition and guided filtering. *IEEE Transactions on Image Processing* **26**(7), 3528–3541 (2017)

38. Yang, Z., Sun, Y., Peng, X., Yiu, S.M., Ma, Y.: Unidemoiré: Towards universal image demoiréing with data generation and synthesis. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. vol. 39, pp. 9354–9362 (2025)
39. Yu, X., Dai, P., Li, W., Ma, L., Shen, J., Li, J., Qi, X.: Towards efficient and scale-robust ultra-high-definition image demoiréing. In: *European Conference on Computer Vision*. pp. 646–662. Springer (2022)
40. Yuan, S., Timofte, R., Slabaugh, G., Leonardis, A., Zheng, B., Ye, X., Tian, X., Chen, Y., Cheng, X., Fu, Z., et al.: Aim 2019 challenge on image demoiréing: Methods and results. In: *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*. pp. 3534–3545. IEEE (2019)
41. Yue, H., Cheng, Y., Mao, Y., Cao, C., Yang, J.: Recaptured screen image demoiréing in raw domain. *IEEE Transactions on Multimedia* **25**, 5589–5600 (2022)
42. Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O.: The unreasonable effectiveness of deep features as a perceptual metric. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 586–595 (2018)
43. Zhang, X., Chen, Q., Ng, R., Koltun, V.: Zoom to learn, learn to zoom. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 3762–3770 (2019)
44. Zheng, B., Yuan, S., Slabaugh, G., Leonardis, A.: Image demoiréing with learnable bandpass filters. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. pp. 3636–3645 (2020)
45. Zhong, Y., Zhou, Y., Zhang, Y., Chao, F., Ji, R.: Learning image demoiréing from unpaired real data. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. vol. 38, pp. 7623–7631 (2024)