SURVEY

Deep Learning for Screen-Shot Image Demoiréing: A Survey

SHOUMING HOU[1], YABING WANG [1], KAI LI[1], YINGGANG ZHAO[1], BAOYUN LU[2], AND LIYA FAN[3]

[1]School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454000, China
[2]School of Software, Henan Polytechnic University, Jiaozuo 454000, China
[3]Department of Computer, Xi'an Jiaotong University City College, Xi'an 710018, China

Corresponding author: Liya Fan (fanly@xjtucc.edu.cn)

ABSTRACT Image demoiréing is an important image processing technology in computer vision, used to remove the moiré from images and improve the image quality. In recent years, the image demoiréing technique based on the deep learning method has gained more attention and achieved good results, but it still has some limitations. This paper aims to provide a review and perspective on the recent advances in deep learning-based image demoiréing techniques. First, the definition and production principle of the image moiré pattern are given. Common datasets and image quality evelation methods in demoiréing studies are analyzed. Then two internationally famous competitions in image demoiréing are introduced. Second, the research status of the supervised demoiréing technique is summarized from four dimensions: sampling method, model network design, baseline model, and training learning strategy. Recent progress made by the mainstream model of unsupervised deep learning in the field of image demoiréing is summarized. The typical application of the image demoiréing technique in panel defect detection and digital radiography is analyzed. The performance and the image quality of the above mentioned models based on different data sets are evaluated in detail. Finally, this paper analyzes and forelocks the problems to be solved in the coming years.

INDEX TERMS Image demoiréing, screen-shot image, deep learning, convolutional neural networks (CNN).

I. INTRODUCTION

The widespread use of digital cameras, especially smartphones, helps people record exciting moments in their work and life by taking photos for review and sharing. However, when the shooting object exist high-frequency texture similar to the resolution of the color filter array (CFA) grid, because the spatial frequency of the CFA pixel grid overlaps with the high-frequency texture of the subject, resulting in colorful water ripple-like texture called moiré [1], as shown in Fig. 1. According to the objects to be photographed, they can be divided into screen-shot images moiré for photographing

electronic displays and texture images moiré for photographing clothing with dense textures and long-distance building's tiles, etc., as shown in Fig. 1 and Fig. 2, respectively. The presence of moiré drastically degrades the visual quality of images and affects the people's visual experience.

In recent years, researchers have focused on the moiré problem and have studied the restoration of corresponding clean images from moiré images. Depending on the order of moiré removal, moiré removal methods can be divided into two main categories: pre-processing and post-processing. Pre-processing directly suppresses the appearance of moiré. In literature, various pre-processing methods have been proposed, including adding a low-pass filter to the camera [2], pixel merging-based methods, and interpolation-based
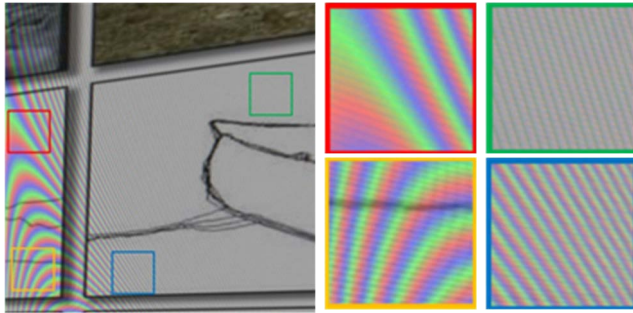

The associate editor coordinating the review of this manuscript and approving it for publication was Hengyong Yu [ID].

FIGURE 1. Moiré of different scales, frequencies, and colors.

FIGURE 2. Textured moiré images.

FIGURE 3. Differential shooting principle.

(a) pixel grid of the display (b) pixel grid of sensor (c) 0 degrees (d) 5 degrees (e) 10 degrees
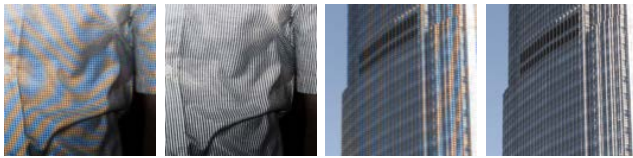
FIGURE 4. Simulation of screen-shots at different angles.

methods [3], [4]. On the contrary, post-processing improves the visual quality of the images by removing moiré using image processing. Liu et al. [5] found that the energy distribution of moiré in the frequency domain is focused and barely mixes with the image texture details and proposed image decomposition-based methods [5], [6], [7]. In addition, there are filter-based methods [8] and image processing methods using Adobe Photoshop. However, as the image resolution increases, the time consuming of removing moiré using image processing methods increases dramatically. For example, Fang et al. [9] take 10s to process an image with 256×256 resolution, so it requires faster methods.

With the rapid development of deep learning techniques in the past few years, deep learning-based image demoiréing methods have been actively explored. Researchers have proposed a variety of deep learning-based demoiréing models, ranging from convolutional neural networks (CNN) (e.g., DMCNN [10]) to generative adversarial nets (GAN) [11](e.g., MR-GAN [12] and cyclic GAN [13]), can remove moiré within 1s, and achieved the best score in image quality evaluation. In general, the methods of image demoiréing with deep learning differ in the following aspects: sampling methods [10], [14], [15], network design [14], [16], [17], baseline model [15], [16], [18], [21], [60], [61], and learning strategies [12], [18], [19], [20], [21], [22].

While there is a screen-shots image demoiréing survey [23], only two methods are involved, which is not comprehensive. Thus, this paper aims to comprehensively survey the latest developments in deep learning-based screen-shot image demoiréing. The rest is organized: Section II defines the moiré and introduces the moiré benchmark datasets and image quality evaluation metrics.

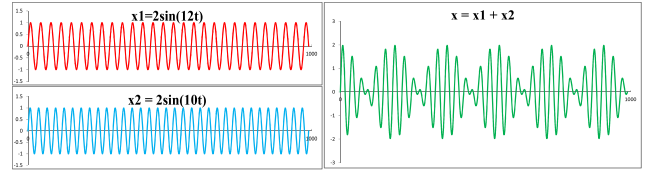Section III compares the supervised image demoiréing model in terms of network composition. Section IV provides an analysis of unsupervised image demoiréing methods. Section V introduces domain-specific image demoiréing applications. Section VI shows the demoiréing results of some models on different datasets, and Section VII summarizes current existing problems and proposes possible research directions in the future.
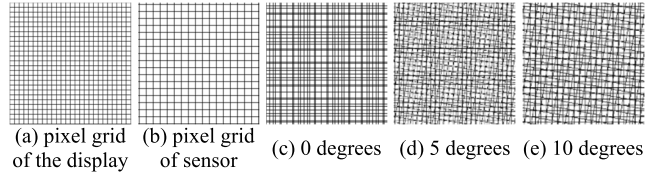
II. PROBLEM SETTING AND TERMINOLOGY

A. MOIRÉ DEFINITIONS

The moiré is a manifestation of the difference shot principle. Two equal amplitude sine waves of similar frequencies are superimposed, and the amplitude of the combined signal depends on the difference between the two frequencies [24], as shown in Fig. 3. Similarly, the difference shot principle also applies to the spatial frequency, i.e., moiré is generated when the spatial frequency of the camera's photosensitive unit is close to the spatial frequency of the photographed object [25]. Ideally, moiré can be avoided when the resolution of the camera lens is smaller than that of the camera sensor [26]. However, in practice, the resolution of the lens often exceeds the resolution of the sensor, which is why moiré is easily formed.

Image demoiréing can be classified as an image recovery task, aiming to restore the corresponding clean image from the moiré image. However, the distribution, appearance, and color of the moiré in the image vary with the type of camera and external factors such as the shooting distance between the subject, angle, and lighting. As shown in Fig. 4, moiré is generated in the simulated screen-shot image, and by changing the angle of the shot, the appearance of the moiré changes accordingly. Besides, as shown in Fig. 5, the moiré pattern is present both in the low-frequency region of the image and mixed with the high-frequency details of the image, which spans a wide range of frequencies. Therefore, image demoiréing is a considerable challenge.

B. DATASET FOR IMAGE DEMOIRÉING

Various datasets were created for image demoiréing studies with significantly different numbers of images,

TABLE 1. Public image datasets for demoiréing benchmarks.

Dataset	Amount	Resolution	Types	Moiré	Paired	Category Keywords
TIP2018[10]	135,000	384×384-874×874	screen-shot	real	yes	Image Net 2012
MoDI[17]	22,734	256×256	screen-shot	real	yes	document image
FHDMi[20]	12,000	1920×1080	screen-shot	real	yes	film clips, wallpapers, documents, sports video frames, etc.
MBRI[21]	65,495	128×128	screen-shot	real	yes	web pages, pictures, text, etc.
Liu[16]	80,000	512×512	synthesis	fake	yes	dialog boxes, texts, web pages, graphics, natural images, etc.
LCDMoiré[30]	10,200	1024×1024	synthesis	fake	yes	ICCV, ECCV, CVPR images
Yue[12]	25,000(18,000)	256×256	screen-shot	real	no	Web pages, texts, portraits, animals, landscapes, etc.
London[31]	460	2300×1700	texture	real	yes	bricks, windows, regular patterns
CFAMoiré[32]	11,100	128×128	synthesis	fake	yes	clothing, construction, etc.

25,000(18,000) indicates 25,000 moiré images and 18,000 clean images.

| clean image | frequency spectrum | moiré image | frequency spectrum |
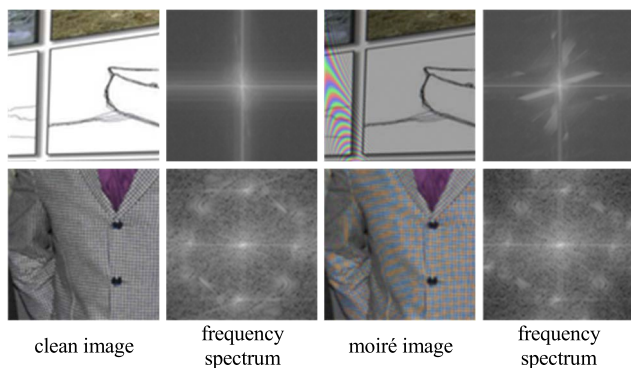
FIGURE 5. The frequency spectrum of the moiré image and the clean image.

resolution, moiré type, variety, etc. Some provide paired moiré images and corresponding moiré-free images, while others offer unpaired moiré images and clean images for training unsupervised models. Table 1 lists some frequently used moiré datasets and specifies their amount, resolution, moiré type, and category keywords.

C. IMAGE QUALITY ASSESSMENT

Researchers usually use objective evaluations as an assessment criterion to evaluate the model's power for demoiréing. The following are several commonly used objective evaluation methods: peak signal-to-noise ratio (PSNR) [27], structural similarity (SSIM) [28], and learned perceptual image patch similarity (LPIPS) [29].

1) PEAK SIGNAL-TO-NOISE RATIO

The PSNR is one of the most commonly used image quality evaluation metrics. For image demoiréing, the PSNR is defined by calculating the difference between the corresponding pixel points of the image. Given the ground truth image I (height and width H and W, respectively) and the reconstruction image I_d, the mean square error (MSE) [33] and PSNR between I and I_d are described as follows:

$$MSE = \frac{1}{H}\frac{1}{W}\sum_{i=1}^{H}\sum_{j=1}^{W}(I(i,j) - I_d(i,j))^2, \quad (1)$$

$$PSNR = 10 \cdot log_{10}\left(\frac{(2^n - 1)}{MSE}\right), \quad (2)$$

where n equals 8. Because PSNR only calculates the differences between corresponding pixels and does not consider the perceptual characteristics of the human eye, it may lead to poor performance in terms of the visual quality of the reconstructed image. However, owing to the absence of precise perceptual metric methods, PSNR remains the utmost commonly applied evaluation criterion for image demoiréing.

2) STRUCTURAL SIMILARITY

Considering that the sensitivity of the human visual system (HVS) to image noise depends on local brightness, contrast, and structure [34], so the SSIM is used to measures the structural similarity between images. The SSIM between I and I_d are described as follows:

$$SSIM(I, I_d) = \frac{(2u_I u_{I_d} + c_1)(2\sigma_{II_d} + c_2)}{(u_I^2 + u_{I_d}^2 + c_1)(u_I^2 + u_{I_d}^2 + c_2)}, \quad (3)$$

where u_I is the mean of I, σ_I^2 is the variance of I, u_{I_d} is the mean of I_d, $\sigma_{I_d}^2$ is the variance of I_d, σ_{II_d} is the covariance of I_d and I, c_1 and c_2 are minimal numbers used to avoid the denominator being equal to 0. The value domain of the SSIM is [0, 1], where a more considerable value indicates less image distortion and more similar images. Because of HVS characteristics, it is more in line with the needs of computer vision.

3) LEARNED PERCEPTUAL IMAGE PATCH SIMILARITY

LPIPS is derived from the weighted MSE distance between the depth features of an image and is used to measure the difference between images. The smaller LPIPS value indicates more similar images. The LPIPS between I_d and I are described as follows:

$$LPIPS(I, I_d) = \sum_{l}\frac{1}{H_l W_l}\sum_{h,w}\| \omega_l \odot \left(\widehat{I}_{hw}^l - \widehat{I}_{dhw}^l\right) \|_2^2, \quad (4)$$

where $\widehat{I}_{hw}^l$ is the l layer feature of $\widehat{I}_{hw}$, and ω_l is used to scale the activations channel-wise. LPIPS has an excellent correlation with human perceptual similarity and is closer to human visual perception than PSNR and SSIM [29].

D. IMAGE DEMOIRÉING CHALLENGES

Because of the wide application prospects of image demoiréing, the technology has also attracted the attention of some academic organizations, and the survey has found that there are two noted challenges about image demoiréing in recent years, namely, AIM [30], [35] and NTIRE [32].

The Advances in Image Manipulation (AIM) held in conjunction with ICCV 2019 contains several different challenges such as screen-shot image demoiréing [30], image and video super-resolution[36], [37], image denoising, image deblurring, etc. For image demoiréing, the AIM demoiréing challenge is based on the LCDMoiré [30] dataset. It consists of fidelity and perceptual quality tracks that aim to contribute to further development of screen-shot image demoiréing.

The New Trends in Image Restoration and Enhancement (NTIRE) held with CVPR 2020 contains several different challenges such as real-world image super-resolution [38], image and video deblurring [39], image dehazing [40], texture image demoiréing [32], etc. For image demoiréing, the NTIRE demoiréing challenge is based on the CFAMoiré [32] dataset and includes two tracks. Track 1 focuses on the single image demoiréing issue, which aims to remove moiré from a single texture image. Track 2 targeted the burst demoiréing problem, where a set of moiré images of the same scene as input, aiming to produce a single demoiréd image. The NTIRE demoiréing challenge intends to promote the development of texture image demoiréing.

III. SUPERVISED DEMOIRÉING

Nowadays, researchers have proposed various demoiréing methods based on deep learning. Most of these methods are based on supervised image demoiréing, trained with moiré and corresponding moiré-free images. Although the models differ significantly, they are essentially a combination of up-and-down sampling methods, network design, baseline model, and learning strategies. Therefore, the researchers mainly consider how to effectively organize and improve the above structures to achieve a better removal effect. In this section, we will analyze and summarize the advantages and limitations of various components of model design.

A. UP-AND-DOWN SAMPLING METHODS

Moiré exists, from low to high, in different frequency domains of an image, and has distinctive characteristics in different frequency domains. It is challenging to completely eliminate moiré at only a single scale [8], so researchers devoted their attention to removing moiré on multiple scales. By downsampling the moiré image several times to highlight the moiré in different frequency domains and then removing the moiré. Experiments show that multi-scale can remove moiré better than single-scale. This part will compare and analyze the commonly used sampling methods based on deep learning.

1) CONVOLUTION

Convolution can extract different features of the input layer and is combined with back-propagation to optimize the parameters of the convolution layer to extract the target features. In particular, convolution for downsampling is achieved by changing padding and stride to reduce the image resolution. Convolutional downsampling, which extracts the features of the image and reduces the resolution of the image, is widely used as a downsampling layer in image demoiréing studies [10], [14], [17], [19], [22], [41], [42].

After downsampling the image several times, the output of each branch must be upsampled to recover the image's original resolution. The commonly used upsampling methods include transposed and sub-pixel convolution. The transposed convolution, i.e., the deconvolution, performs the opposite operation of a standard convolution. Precisely, it scales the input image to twice the original resolution by inserting the corresponding zero values and performs the default convolution. The obtained output was twice as large as the input size. Because of its simplicity, transposed convolution is broadly used in multi-scale frameworks for demoiréing research [10], [17], [18], [22], [43]. However, it easily causes "uneven overlap," which leads to a checkerboard shape [44] of the result and degrades the quality of the image, affecting the result of moiré removal.

The sub-pixel convolution generates multiple channels by convolution and then reshapes the channels to achieve upsampling. In general, assuming that the input size is h × w × c, the first convolution is used to obtain the output size h × w × cs^2, where s is the enlargement factor. Subsequently, performing the reshaping operation to produce output size is hs × ws × c. Sub-pixel convolution can solve the checkerboard phenomenon in transposed convolution, and this method is widely used in demoiréing studies [14], [19], [41], [42], [45], [46]. Sub-pixel convolution has a larger perceptual field and provides more contextual information than transposed convolution. However, as the distribution of the receptive field during convolution at the image edges is uneven, it may result in artifacts at the boundary of the image.

2) WAVELET TRANSFORM

Because traditional convolution is irreversible, it causes information loss when used for upsampling and downsampling. The wavelet transform is reversible, has perfect reconstruction capability, and can avoid information loss during signal decomposition. Liu et al. [31] used discrete wavelet transform (DWT), and inverse discrete wavelet transform (IDWT) to replace downsampling and upsampling. In general, for the input features, the low-pass and high-pass filters are first used to filter the corresponding low-frequency part (high-frequency is filtered out) and high-frequency part (low-frequency is filtered out). Then, the low-pass and high-pass filters are used again for the obtained features, obtaining low-frequency information of the image and high-frequency information in horizontal, vertical, and diagonal directions.
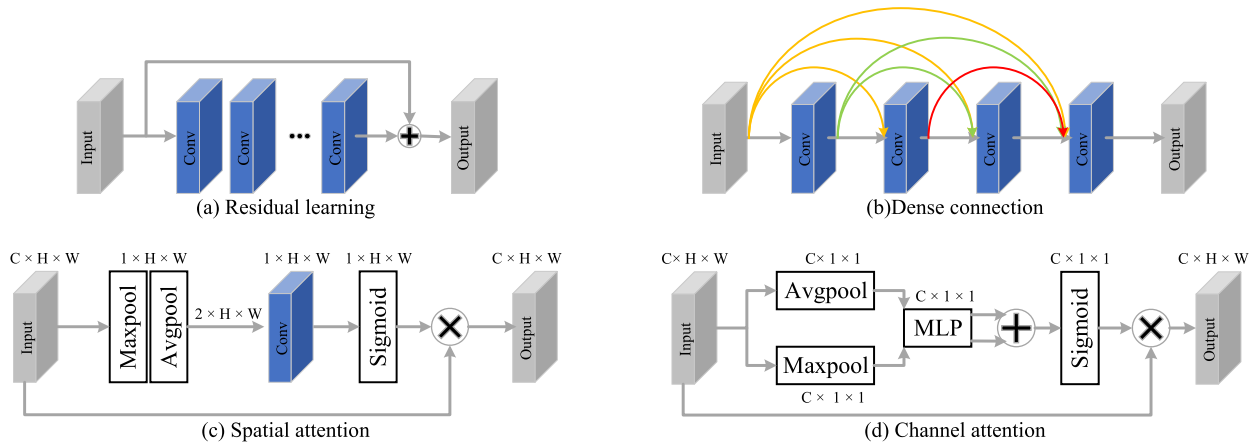
FIGURE 6. Network design strategies.

It is easy to extract the image's structural and texture detail information.

B. NETWORK DESIGN

One of the most important parts of deep learning is network design. In demoiréing studies, researchers have applied various components to build their multi-branch demoiréing networks. In this part, we analyze the advantages and limitations of the most commonly used essential principles and strategies for network design.

1) RESIDUAL LEARNING

Because of the unique construction of residual learning [47], as shown in Fig. 6a, deeper neural networks can be trained. Thus, residual learning is also used in demoiréing research [16]. Image demoiréing is an image transformation problem that converts a moiré image into a clean one. Because residual learning only learns the variation in the difference between clean and moiré images without learning the complete transformation between images, it reduces learning difficulty and complexity [22]. In addition, residual learning can avoid the gradient diffusion problem caused by adding layers to the network and increasing the network's performance and stability [15].

In practical applications, as well as stacking residual blocks, residual learning is used in combination with the attention mechanism [14], [17], [42] and dense connectivity [45], [46] to address the dynamic properties of moiré pattern and the variability between different branches.

2) DENSE CONNECTIONS

To address the problem of gradient disappearance that occurs as the number of network layers increases, Huang et al. [48] proposed a dense connection. A dense connection conveys the output features of each layer to all subsequent layers. In particular, in traditional convolutional neural networks, each layer has only one input and output to the following layer. With dense connections, each layer's inputs are the outputs of all previous layers, as shown in Fig. 6b. The dense connection alleviates the gradient disappearance problem, enhances feature transfer, and more effectively uses the features. Because the network does not need to learn additional feature maps, the number of parameters is reduced.

To fuse high-level features that contain more semantic information and less spatial information with low-level features that include more spatial information and less semantic information, Gao et al. [18] introduced a dense connection. The dense connection takes each branch's output as the previous branch's input so that each branch of the network learns more semantic and spatial information. Yang et al. [46] used dense connections to compose information exchange modules that combine features from different branches. Zheng et al. [19], Liu et al. [31], Zheng et al. [41], and Vien et al. [45] also employ dense connections in networks, and their dense connection are used between each branch to mitigate gradient disappearance and reduce network parameters.

3) ATTENTION MECHANISM

The human visual attention mechanism is a survival mechanism that humans developed during long-term evolution. Focusing on target regions and suppressing secondary information enables humans to use their limited attention to quickly filter the information that needs more attention from a massive amount of information. The attention mechanism in deep learning is derived from the human visual attention mechanism. In deep learning, through the attention mechanism, the network can choose only the information that needs to be focused on from a large amount of information received, saving most computational resources. Among them, the attention mechanisms can be divided into channel attention and spatial attention.

a: CHANNEL ATTENTION

Yang et al. [7] found that moiré artifacts pollute the R and B channels more than the G channel for RGB images. The

channel-based attention mechanism is modeling each channel to learn the interdependence between different channels, as shown in Fig.6c. For the dynamic characteristics of moiré at different scales, Cheng et al. [14], [42] combined residual blocks with channel attention encoding dynamically moiré pattern, which significantly improved the ability of demoiréing. Because most moiré shapes are curved and have distinct edges, He et al. [43] used a channel-based edge predictor based on channel attention to predict a moiré edge map to guide feature extraction. To avoid the effect of the moiré pattern on image color recovery, Zheng et al. [19], [41] proposed an improved channel attention-based color mapping module to achieve accurate color recovery.

b: SPATIAL ATTENTION

The area containing the moiré pattern requires more attention, and spatial attention is used to look for the target area, as shown in Fig. 6d. Liu et al. [31] designed a direction-aware module combined with spatial attention and dense connections to determine the location of moiré effectively. Guo et al. [17] combined channel attention and spatial attention and used channel-level attention maps to guide the production of spatial attention maps that accurately find the most critical information in multi-scale features. Sun et al. [15] proposed an efficient attention fusion module that combines channel attention, spatial attention, and local residual learning, adaptively learning the different features of each branch.

4) DILATION CONVOLUTION

Yu et al. [49] introduced dilation convolution, which increases the convolution kernel's perceptual field while maintaining the constant number of parameters. For 3×3 kernel convolution with a dilation rate of 2, the perceptual field of the convolution kernel after dilation is the same as that of the 5×5 convolution kernel, but the number of parameters is only equal to that of the 3×3 convolution kernel. To produce more realistic details in the demoiréing images, Guo et al. [17], Zheng et al. [19], and Zheng et al. [41] employed dilation convolution instead of standard convolution and increasing the perceptual field of the convolution kernel by more than a few times to get better demoiréing performance. Liu et al. [31] proposed that when dilation convolution is performed with a fixed dilation rate, some pixels not be involved, which may lead to the production of tessellated artifacts. Unlike using a continuous convolution rate r={1,2,3,4,5...}, Liu et al. [31] applied the Fibonacci series r={1,2,3,5,8...} as the dilation rate, which effectively reduced the tessellated artifacts.

5) GLOBAL-LOCAL STRATEGY

Because the moiré pattern is diverse, as shown in Fig. 1, ranging from curved stripes covering a large area with low periodicity to fine, densely distributed vertical stripes. It is difficult to remove all moiré adopting a network with the same receptive field. He et al. [20] proposed a cascaded moiré removal strategy that first removed large-scale moiré using

global branches and then removed residual moiré using local branches. Likewise, to solve the color degradation problem caused by moiré, Zheng et al. [19], [41] studied and observed a color shift between moiré images and moiré-free images and then proposed a two-step tone mapping strategy. The global tone mapping module was first applied to learn the global color shift between moiré and clean images. The local tone mapping module is then used to perform local color recovery, effectively solving the color degradation problem caused by moiré through global and local tone mapping.

6) MULTI-BRANCH FUSION STRATEGY

For the multi-branch demoiréing network, fuse the outputs of each branch to achieve better output results. This subpart will analyze and compare the weight distribution and fusion strategies in multi-branches.

a: DIRECT SUMMATION

For multi-scale fusion, the simplest method is to be equitable to the outputs of each branch and add the results of each branch directly, i.e., set the weights of each branch as 1 [10], [18]. The advantage of the direct summation method is that it is computationally simple. Still, the disadvantage is not considering the significance and effect of the output of each branch to image reconstruction.

b: BRANCH SCALING MODULE

Cheng et al. [14], [42] proposed that each branch's moiré texture and image details are different and have different importance for image recovery. Therefore, proposed a branch scaling module that automatically learns the importance of each branch by back-propagation assigns different importance weights and then sums.

c: SQUEEZE AND EXCITATION MODULE (SE)

Based on channel attention, Hu et al. [50] proposed an SE block that learns the weights of each channel by expressly modeling the mutual dependencies between channels and adaptively aligning the extracted features. The SE block proved to have a significant performance in image classification. He et al. [20], He et al. [43], and Yang et al. [46] also used the SE block for demoiréing studies. SE block reweighted the output of each branch, emphasis on branches with SE block dominant bands.

7) OTHER STRATEGIES

In addition to often used modules and strategies, there are also fewer used demoiréing modules. Migration learning is a machine learning method that uses pre-trained models in other tasks. Liu et al. [16] were first using a synthetic moiré dataset to train the remove moiré network. Then the network was further trained using a real moiré dataset. Zheng et al. [19], [41] employed a learnable bandpass filter for learning moiré prior to accurately distinguishing the moiré pattern from the image texture. Liu et al. [51] applied paired focused moiré images and scattered moiré-free images as datasets to

remove moiré images using a fuzzy kernel constraint. In this way, the authors proposed that the network does not require training and can be easily used on any smartphone or digital camera. Because images have different characteristics in different domains, for example, the spatial domain contains more structural information, and the frequency domain is rich in image details. Wang et al. [22] used a complementary dual-domain network (3DNet) that combined the spatial and frequency domains to remove moiré. 3DNet yields better PSNR values, 6.24 dB and 3.79 dB higher than a single spatial or frequency domain.

C. BASELINE MODEL
In this part, we analyze the baseline models used in demoiréing studies, including U-Net, and Transformer.

1) U-NET
U-Net was originally a multi-scale model for medical image segmentation [52], as shown in Fig. 7. Due to the symmetric contraction architecture of the U-Net model, precise localization while capturing contextual information. U-Net and its variants have achieved excellent results on other computer vision tasks [53], [54], [55]. Researchers have started to study the method of demoiréing based on the U-Net model. Gao et al. [18] proposed a Multi-scale feature enhancing network (MSFE) based on the U-Net architecture, residual learning, and dense connectivity. MSFE utilizes the multi-branch U-Net combined with residual learning to learn the differences between the moiré in different branches and make use of dense connections to fuse the semantic and spatial information in different frequencies. Sun et al. [15] used the U-Net architecture as a baseline model to perform a demoiréing study in the wavelet domain and achieved excellent results.
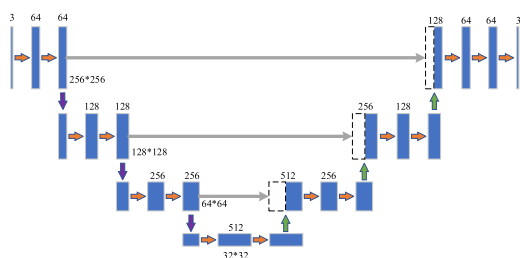

FIGURE 7. U-Net model.

2) TRANSFORMER
Transformers are a neural network framework featuring the self-attention mechanism [57]. They were initially applied in natural language processing and then in computer vision tasks, including low-level vision very recently [58], [59]. Chen et al. [59] show that the transformer is more advantageous than CNN in large-scale data pre-training low-level vision. Besides, the transformer significantly reduces the computational resources required. Thus Wang et al. [60] designed a transformer network UFormer based on the U-Net

model, which can handle local context information while efficiently capturing long-range inter-texture dependencies. Compared with the U-Net family, the computation cost and the number of model parameters of UFormer are significantly reduced. Liu et al. [61] take a self-attentive mechanism to separate generalized priors of images as additional priors for the transformer to investigate the nature of image restoration tasks rather than task-specific restoration. With the help of task-agnostic generalized image prior, it can be easy to address new image restoration tasks.

D. LEARNING STRATEGIES
1) LOSS FUNCTION
In deep learning, loss functions are used to guide model optimization and reduce errors in model prediction values. The proper loss function can help the model focus on the features that need attention and obtain the best and fastest convergence. In this part, we will look closely at the broadly used loss functions used in demoiréing studies.

a: PIXEL LOSS
Pixel loss is used to calculate the difference in the corresponding pixel points between the output image $\widehat{I}$ and the target image I, which includes L1 loss (MAE) [19], [31], [41], [45] and L2 loss (MSE) [10], [16], [17], [18]:

$$L1_{loss}\left(I, \widehat{I}\right) = \frac{1}{HWC} \sum_{i,j,k} \left| I_{i,j,k} - \widehat{I}_{i,j,k} \right|, \quad (5)$$

$$L2_{loss}\left(I, \widehat{I}\right) = \frac{1}{HWC} \sum_{i,j,k} \left(I_{i,j,k} - \widehat{I}_{i,j,k} \right)^2, \quad (6)$$

where H, W, and C are the image length, width, and the number of image channels, respectively. Besides, a variant of frequently used L1 loss, L1 Charbonnier loss [15], [16], [18], [42], [46]:

$$L1\,Cha_{loss}\left(I, \widehat{I}\right) = \frac{1}{HWC} \sum_{i,j,k} \sqrt{\left(I_{i,j,k} - \widehat{I}_{i,j,k} \right)^2 + \varepsilon^2}, \quad (7)$$

where ε is a minor constant used to stabilize training. The L1 loss calculates the absolute value of the difference between image pixels with a fixed penalty for any difference in size. Compared with L1 loss, L2 loss is insensitive to more minor errors (less than 1) and more concerned with more significant errors (greater than 1). L2 calculates the square of the pixel difference, which often results in the over-smoothing of the produced image. Because the image quality evaluation metric PSNR is negatively correlated with L2 loss (Section 2.3), i.e., the smaller the L2 loss, the larger the PSNR, and L2 loss is the more commonly utilized loss function. However, because pixel loss often does not consider image structural features, it leads to results with poor performance at perceptual quality.
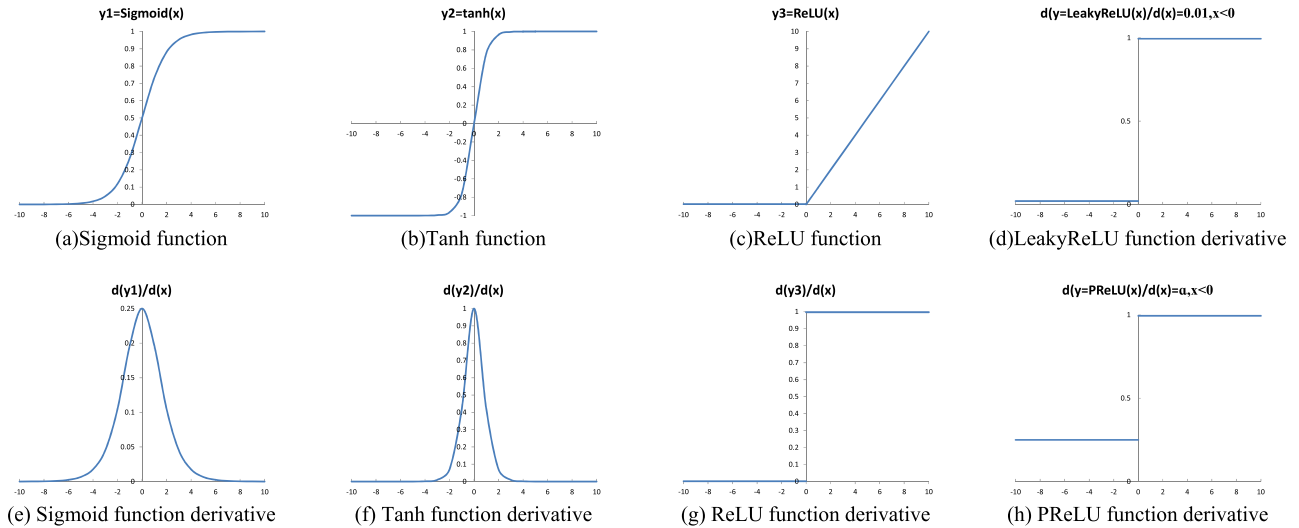
FIGURE 8. Activation functions and derivative functions.

b: PERCEPTUAL LOSS (CONTENT LOSS)

Because pixel loss only calculates the difference between the corresponding pixels, a slight misalignment between images can cause a considerable error value. In view of this, Johnson et al. [62] proposed a perceptual loss for comparing two different images that look similar, which is used to compare the semantic differences between images:

$$L_{percep}\left(I, \widehat{I}, \varphi, l\right) = \frac{1}{H_l W_l C_l} \sqrt{\sum_{i,j,k} \left(\varphi_{i,j,k}^{(l)}(I) - \varphi_{i,j,k}^{(l)}\left(\widehat{I}\right)\right)^2}, \tag{8}$$

where $\varphi_{i,j,k}^{(l)}$ refers to the l layer feature map of a VGG-16 or VGG-19 classification network pre-trained using ImageNet [63]. Perceptual loss does not minimize the differences between image pixels but rather a perceptual similarity and is extensively used in demoiréing studies [15], [20], [22], [31].

c: IMPROVED SOBEL LOSS

To explore the structural information of moiré images, Zheng et al. [19] proposed an improved Sobel loss (ASL):

$$ASL\left(I, \widehat{I}\right) = \frac{1}{HWC} \sum \left| Sobel^*(I) - Sobel^*\left(\widehat{I}\right)\right|, \tag{9}$$

where $Sobel^*$ denotes the improved Sobel filter. Zheng et al. [41] proposed a Sobel loss by adding dilation convolution, which adjusted the perceptual frequency domain through different dilation rates of the ASL:

$$D - ASL^{\{d_1, d_2, \ldots, d_n\}} = \sum_{i=1}^{n} ASL\big|_{dilation\_rate=d_i}, \tag{10}$$

where $ASL\big|_{dilation\_rate=d_i}$ denotes the ASL with dilation rate d_i.

d: WAVELET LOSS

Because the low-frequency and high-frequency details of the moiré are more easily observed in wavelet sub-bands,

Sun et al. [15] applied the L1 Charbonnier loss in the wavelet domain (11), as shown at the bottom of the next page, where ω denotes the wavelet decomposition, and w means the number of wavelet subbands. In addition to calculating the L1 Charbonnier loss in the wavelet domain, Liu et al. [31] added detail loss to prevent the wavelet coefficients from converging to zero. Detail loss can be expressed as follows:

$$L_{detail}\left(I, \widehat{I}\right) = \sum_{w=1}^{4} max\left(\alpha \left|I_w\right|^2 - \left|\widehat{I}_w\right|^2, 0\right), \tag{12}$$

where α is set to 0, avoid L_{detail} convergence to 0.

Additionally, there are various rarely used loss functions, such as contextual bilateral loss [20], Euclidean loss [43], L1 Sobel loss [46], and attentional loss [31]. In fact, researchers usually combine several loss functions by assigning different weights to different loss functions to help the model learn the mapping of moiré images to clean images from different perspectives. However, the application of loss functions and the size of the assigned weights must be explored to achieve higher performance for removing moiré.

2) ACTIVATION FUNCTION

The activation function is the key to distinguishing a perceptron from a neural network. With its activation function, the neural network has unlimited creativity. Activation functions are divided into saturating and non-saturating activation functions. This part will take a closer look at the advantages and limitations of the activation functions commonly used in demoiréing research.

a: SATURATING ACTIVATION FUNCTION

The saturating activation function is a function in which the derivative value tends to zero as x tends toward positive and negative infinity. The saturated activation functions include the Sigmoid function [14], [41] and Tanh function [16], [46].

As shown in Fig. 8a, the sigmoid function is continuous everywhere, which is easy to get the derivative and back-propagation. The disadvantage of the Sigmoid function is that its maximum derivative is 0.25. As shown in Fig. 8e, when the back-propagation has been performed, the gradient of each layer is multiplied, which easily leads gradients to zero, resulting in "gradient disappearance." In addition, the Sigmoid function is not centrosymmetric, and the output value is always positive, which makes back-propagation difficult and hardly converges with a deeper network. In order to solve the problem that the Sigmoid function is difficult to converge, proposed Tanh function, as shown in Fig. 8b. The Tanh function is centrosymmetric and converges easily compared to the Sigmoid function. However, as shown in Fig. 8f, the maximum derivative of the Tanh function is 1, which still does not solve the "gradient disappearance" problem.

b: NON-SATURATING ACTIVATION FUNCTION

The non-saturating activation function satisfies either left or right saturation or neither. In the study of demoiréing, the commonly used non-saturated activation functions are the ReLU function [10], [18], [19], [31], [46] and its variants LeakyReLU function [20], [22], and PReLU function [14], [16]. As shown in Fig. 8c (ReLU function), compared with the saturated activation function, the derivative of the ReLU function is always 1 in the range greater than zero, so it does not produce the gradient disappearance problem. However, in the region less than zero, as shown in Fig. 8g, the derivative of the ReLU function is zero, which causes the neuron to stop updating when the input of the ReLU function is negative during back-propagation. A variant of ReLU, LeakyReLU, was proposed to solve the problem of neurons stopping updating. In the negative region, the derivative of the LeakyReLU function is set to 0.01, as shown in Fig. 8d, which avoids the situation in which the derivative is zero and causes the neurons to stop updating. Similarly, another variant of ReLU, the PReLU function, has a small gradient in the negative area, i.e., the derivative, as shown in Fig. 8h, the negative value was updated with the momentum and learning rate [65]. The PReLU function was the ReLU function when the slope was equal to zero. The PReLU function is a LeakyReLU function when the slope equals 0.01.

IV. UNSUPERVISED DEMOIRÉING

Existing works on screen-shot images demoiréing studies mainly concentrated on supervised learning, i.e., training with pairs of moiré and corresponding clean images. However, it is difficult to collect clean images corresponding to moiré images in some scenes, such as outdoor LED displays and dense textures in realistic environments. the moiré
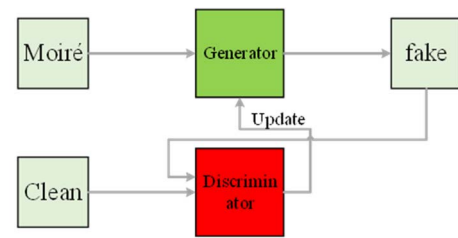

FIGURE 9. GAN model.

removal network training used existing screen-shot datasets and synthetic datasets can neither effectively remove the moiré produced by LED displays well, nor have excellent robustness. To solve this problem, researchers started to study unsupervised image demoiréing. Unsupervised learning refers to training with unpaired moiré images and clean images, and the networks are maybe more suitable for the moiré problem in real-world scenarios. GAN is a deep neural network based on game theory, is a commonly used unsupervised baseline based model that consists of two networks: a generator and a discriminator. as shown in Fig. 9, the generator network produces a pseudo-output image in image-to-image transformation by learning the mapping between the input and target images. The discriminator network tries to distinguish the pseudo-output image from the target image. The generator and discriminator play the game back and forth until the discriminator can no longer distinguish between the generator-generated image and the target image. In the following, we will analyze several existing deep learning-based unsupervised moiré removal methods, and more methods should be explored.

A. AMNet

Yue et al. [21] indicated that moiré not only decreases the visual quality but also affects the brightness of the image. Therefore, Yue et al. [21] proposed an additive and multiplicative network (AMNet) as the generator to improve the image brightness while removing moiré. Since moiré is distributed in different frequency domains, the additive module combines U-Net and atrous spatial pyramid pooling to extract image features from different frequency domains. inspired by SE [50], the multiplication module generates multiplication coefficients that, when combined with the additive module's output, improve the image's brightness. PatchGAN [56], which is good at distinguishing high-frequency details of images, is used as the discriminator to better distinguish between real and fake images. Experiments have proven that AMNet improves image brightness well while removing moiré. Still, the large brightness difference between the moiré

$$L_{wavelet}\left(I, \widehat{I}\right) = \frac{1}{HWC} \sum_{i,j,k} \sqrt{\left(\sum_{w=1}^{4} \omega\left(I_{i,j,k}\right) - \sum_{w=1}^{4} \omega\left(\widehat{I}_{i,j,k}\right)\right)^2 - \varepsilon^2}, \qquad (11)$$

images and its corresponding clean image leads to the lower PSNR values between the demoiréing image and the clean image, as shown in table 4.

B. MR-GAN

Yue et al. [12] used CycleGAN [66] as a benchmark model to build an unsupervised moiré removal model (MR-GAN). Unlike conventional GAN networks, MR-GAN contains two generator networks ($G_{moiré \to clean}$ and $G_{clean \to moiré}$) with discriminator networks (D_L and D_S). The moiré removal generator $G_{moiré \to clean}$ translates the moiré image I_m to the non-moiré image $\widehat{I}_c$, and the moiré synthesis generator $G_{clean \to moiré}$ re-translates the generated non-moiré image $\widehat{I}_c$ to the moiré image $\widehat{I}_m$, using a cycle of dual generators learning the mapping between the moiré and moiré-free images. Considering the multi-scale characteristics of moiré, MR-GAN uses dual discriminators D_L and D_S, which D_L is used to distinguish large-scale moiré pattern from image features and D_S focuses on distinguishing small-scale moiré pattern from image details. A set of self-supervised and adversarial loss functions to maintain the effect of learning effectiveness and training stability. The self-supervised loss includes cycle consistent loss (pixel layer, feature layer), identity loss, cosine similarity loss, and content leakage loss, which are guided in different ways to produce high-quality images. The self-supervised losses are as follows:

$$
\begin{aligned}
L_{pcyc} &= \| G_{clean \to moiré} \left(G_{moiré \to clean} \left(I_m \right) \right) - I_m \|_1 \\
&\quad + \| G_{moiré \to clean} \left(G_{clean \to moiré} \left(I_c \right) \right) - I_c \|_1 , \quad (13)
\end{aligned}
$$

$$
\begin{aligned}
L_{fcyc} &= \sum_i \big(\| \psi_i \left(G_{clean \to moiré} \left(G_{moiré \to clean} \left(I_m \right) \right) \right) - \psi_i \left(I_m \right) \|_1 \\
&\quad + \| \psi_i \left(G_{moiré \to clean} \left(G_{clean \to moiré} \left(I_c \right) \right) \right) - \psi_i \left(I_c \right) \|_1 ,
\end{aligned}
$$
$$
(14)
$$

$$
\begin{aligned}
L_{idt} &= \| G_{moiré \to clean} \left(I_c \right) - I_c \|_1 \\
&\quad + \| G_{clean \to moiré} \left(I_m \right) - I_m \|_1 , \quad (15)
\end{aligned}
$$

$$
\begin{aligned}
L_{cos} &= 1 - cos \left(G_{moiré \to clean} \left(I_c \right), I_c \right) \\
&\quad + 1 - cos \left(G_{clean \to moiré} \left(I_m \right), I_m \right), \quad (16)
\end{aligned}
$$

$$
L_{cl} = \| \psi_j \left(I_m^{syn} \right) - \psi_j \left(I_c \right) \|_1 , \quad (17)
$$

where L_{pcyc}, L_{fcyc}, L_{idt}, L_{cos}, and L_{cl} denote pixel-layer cycle consistent loss, feature-layer cycle consistent loss, identity loss, cosine similarity loss, and content leakage loss, respectively; ψ_i and ψ_j are the layer i and j features of the VGG-19 model trained using ImageNet, respectively, I_m^{syn} denotes a synthetic moiré image. The main contribution of the MR-GAN is the proposed self-supervised loss function, which proved to enhance the PSNR and SSIM effectively.

C. CYCLIC GAN

Inspired by Cycle GAN [66] and conditional generative adversarial network (CGAN) [67], Park et al. [64] proposed a cyclic GAN model that includes a moiré generation network and demoiréing network. Specifically, in the moiré generator network, the moiré image I_m is fed to the generator G_{De} to produce a pseudo-clean image $\widehat{I}_c$. Subsequently, $\widehat{I}_c$ is input to another generator G_M to restore the moiré image $\widehat{I}_m$. To reconstruct the realistic $\widehat{I}_m$, in addition to generating the moiré texture, it also considers the color degradation caused by moiré. The demoiréing network is based on CGAN and is trained in a supervised manner using clean images and the corresponding pseudo-moiré images $G_M \left(I_c \right)$.

To further improve the demoiréing capability of cyclic GAN, Park et al. [13] improved on [64]. Park et al. [13] found that the appearance of moiré in screen-shot images is accompanied by global pixel intensity degradation. When using the generator to generate the moiré pattern, a 256- dimensional image histogram was used to estimate the global pixel intensity degradation degree and added to the moiré generator network to make the generated moiré more realistic. It is experimentally proved the improved Cyclic GAN model improves 1.77 dB in PSNR compared with the Cyclic GAN model.

V. DOMAIN-SPECIFIC APPLICATIONS

In addition to people's daily use of the camera, which produces moiré and degrades image quality, moiré also occurs in some special applications, such as Mura defect detection in the LCD panel manufacturing process and X-ray maps in medical diagnosis.

A. DEFECT IMAGE DEMOIRÉING OF THIN-FILM TRANSISTOR LIQUID-CRYSTAL DISPLAY IMAGE

In manufacturing thin-film transistor liquid crystal displays (TFT-LCDs), displays are composed of multiple materials and bonded substrate layers. During bonding, various gaps, contaminants, air bubbles, and other defects may intersperse, resulting in Mura defects in the fabricated display. Common Mura defects include stains, dark clusters, and bright clusters, which result in non-uniform display brightness and affect display performance. Although automated optical inspection (AOI) was used to detect defective panels [68], AOI is also difficult to identify Mura defects [69]. Owing to the outstanding performance of deep learning, researchers use it to detect and identify panel defects. However, when using a digital camera to capture a monitor's panel, moiré is often present in the panel image captured by the camera, making it more challenging to detect Mura defects when mixed with moiré. In order to eliminate the effects of moiré, some researchers first explored removing moiré from the panel images and then detecting and classifying Mura defects.

For training model, in the process for checking semi-finished panel, Lu et al. [70] first collected moiré images with Mura defects. Then continue processing semi-finished

| Moiré | MDDM[14] | MopNet[43] | AMNet[21] | FHDe2Net[20] | MBCNN[19] | MR-GAN[12] | Clean |

FIGURE 10. Demoiréing results on the TIP 2018 dataset.

panels that should be scrapped. Finally, obtain Mura defects images without moiré by computer image acquisition. However, continue processing semi-finished panels that should have been scrapped would cause additional time and material costs, which leads the training dataset scale is limited. In addition, Lu et al. [70] used U-Net as a generator of GAN. They used a recurrent neural network combined with the attention mechanism to yield the moiré place graph that guides the U-Net model to remove moiré while preserving Mura defects. Because of the difficulty in collecting large-scale screen-shot panel moiré images, Lu et al. [71] used the migration learning method, that is, first training the CGAN model on a common image dataset to extract typical features suitable for all images, and then continuing the training with a screen-shot panel image moiré dataset to improve the demoiréing capability.

To solve the problem of limited moiré datasets containing Mura defects, Kim et al. [72] proposed to use a smartphone to display images with Mura defects and fixed the camera to collect moiré images containing Mura defects by moving or rotating the phone, and by inserting ArUco markers to align the Mura defect image and the corresponding moiré image. Besides, Kim et al. [72] used U-Net as a baseline model and combined it with Fourier transform to calculate the frequency domain loss between images to ensure the removal of moiré, which significantly improved the PSNR of the images.

B. DEMOIRÉING OF DIGITAL RADIOGRAPHY IMAGE

Digital radiography (DR) is an analytical instrument used in preventive medicine and public health, mainly for medical diagnosis. The core technology of DR is the flat panel detector. As X-rays pass through the body, they produce scattered rays of longer wavelengths and variable directions, resulting in blurred images from the flat panel detector. To solve this problem, using X-ray filter grids to absorb scattered rays can effectively reduce scattered rays and enhance image quality. However, X-ray filter grids lead to moiré in the image, which seriously affects clinical diagnosis. To address the moiré in DR images, Chen et al. [73] proposed to employ a CNN model based on U-Net. Confronted with the difficulty in collecting the DR imaging moiré dataset, they subsequently proposed a fast signal processing method to combine the moiré dataset for training. Experiments have proven that the proposed model significantly reduces the moiré in DR images and maintains the image resolution.

VI. SCREEN-SHOT IMAGE DEMOIRÉING MODELS' PERFORMANCE

To better show the real effect of demoiréing, as shown in Fig. 10, Fig. 11, and Fig. 12, we selected some supervised and unsupervised models to compare their performance in moiré removal based on the TIP 2018, LCDMoiré, and MBRI datasets respectively. Considering the limited computing power of mobile phones and embedded devices, the model's size is also a critical problem in addition to its demoiréing capability. In this section, we tabulate the performance and size of some of the previously discussed deep learning-based image demoiréing models on the TIP2018, LCDMoiré, and MBRI datasets.

Table 2 focuses on the TIP 2018 dataset, the first large-scale moiré benchmark dataset. In table 2, PSNR and SSIM are used to demonstrate the model's performance in demoiréing, while FLOPs and Parameters indicate the calculation volume and size of the model.

Table 3 shows the test results based on the LCDMoiré dataset, which is a synthesis moiré dataset for benchmarking example-based image demoiréing. Table 3 shows the performance of the model in terms of demoiréing. Compared to the real moiré pattern, synthetic moiré is easier to remove.
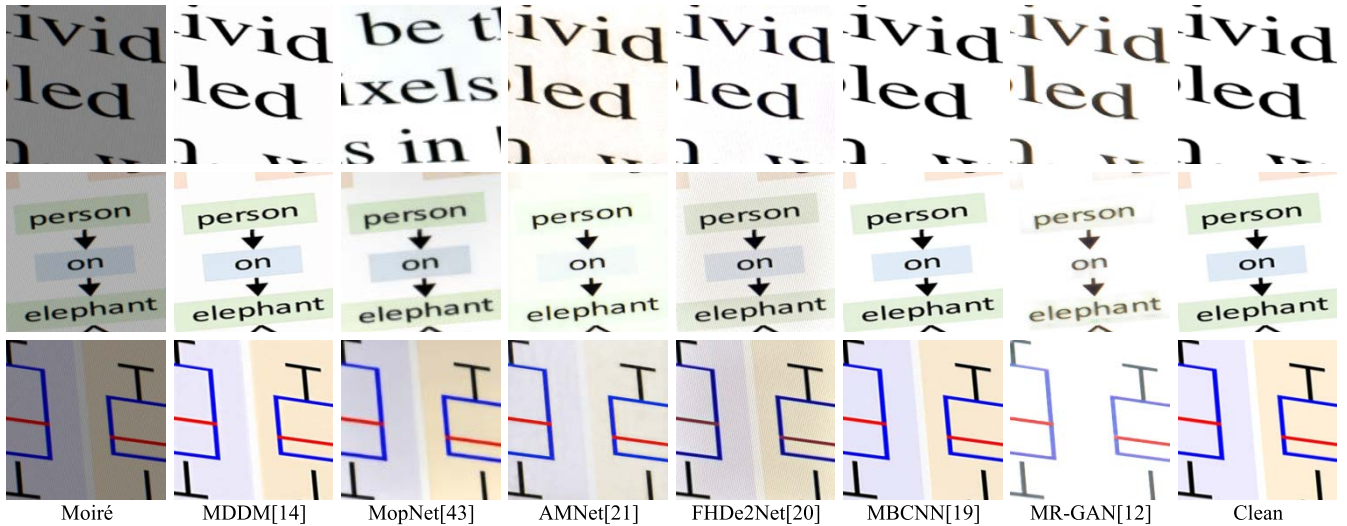
FIGURE 11. Demoiréing results on the LCDMoiré dataset.

FIGURE 12. Demoiréing results on the MBRI dataset.

TABLE 2. Performance comparison of screen-shot image demoiréing models on The TIP2018 dataset.

Model	SUPERVISED											UNSUPERVISED		
	DMCNN [10]	MDDM [14]	MSFE [18]	MopNet [43]	WDNet [31]	HRDN [46]	MBCNN [19]	FHDe²Net [20]	EAFN [15]	MDDM+ [42]	MBCNN+ [41]	MR-GAN [12]	Park [64]	Park [13]
Publication	2018	2019	2019	2019	2020	2020	2020	2020	2021	2021	2021	2021	2021	2022
PSNR	26.77	28.11	25.31	27.75	28.08	28.47	30.03	27.78	29.7	30.03	30.41	21.43	21.78	22.37
SSIM	0.871	0.895	0.91	0.895	0.904	0.860	0.893	0.896	0.912	0.897	0.900	0.7423	0.7632	0.7811
FLOPs	49.3G	26.34G	−	396.1G	42.9G	15.05G	125.3G	353.87G	−	27.73G	148.6G	47.22G	−	51.15G
Parameters	1.4M	8.01M	17.98M	12.4M	5.7M	7.79M	13.5M	51.9M	38M	3.57M	14.9M	29.1M	24.1M	51.7M

TABLE 3. Performance comparison of synthesis screen-shot image demoiréing models on the LCDMoiré dataset.

Model	SUPERVISED							UNSUPERVISED		
	DMCNN[10]	MDDM[14]	MSFE [18]	Vien[45]	MBCNN[19]	MDDM+[42]	MBCNN+[41]	MR-GAN[12]	Park[64]	Park[13]
PSNR	35.48	42.49	36.66	41.97	44.04	43.44	45.08	25.21	26.23	28.00
SSIM	0.9785	0.994	0.981	0.988	0.9948	0.996	0.9967	0.9207	0.9504	0.9515

TABLE 4. Performance comparison of screen-shot image demoiréing models on The MRBI dataset.

Model	SUPERVISED	UNSUPERVISED			
	DMCNN[10]	AMNet[21]	MR-GAN[12]	Park[64]	Park[13]
PSNR	16.71	19.68	15.68	16.34	17.47
SSIM	0.787	0.818	0.7324	0.7643	0.8051

Table 4 provides the test performance based on the MBRI dataset. Apart from demoiréing, it also focuses on improving the brightness reduction caused by moiré. However, as shown in Table 4, the large brightness difference between image pairs in the MBRI dataset leads to low PSNR of the removed moiré images.

In summary, image demoiréing models based on deep learning have received increasing attention and significantly improved both effectiveness in moiré removal and time efficiency. While some studies do not provide source code and lack reproducibility, as the research progresses, more and more researchers are open-sourcing their studies, promoting the development of image demoiréing research.

VII. CONCLUSION AND FUTURE DIRECTIONS

This paper extensively surveys recent advances in screen-shot image demoiréing based on deep learning. First, we present the definition of the moiré problem and introduce some moiré-related problems, including moiré datasets, image quality assessment methods, and moiré removal contests. Then, we focus on the current state of supervised and unsupervised learning image demoiréing and introduced domain-specific demoiréing applications. Finally, we show the moiré removal performance of different models based on different datasets. Despite the remarkable results that have been achieved in the research of screen-shot image demoiréing, some problems remain to be improved and solved. In the following, we will summarize the current problems and propose possible future research directions.

A. MODEL LIGHT-WEIGHTING

The performance of deep learning-based image demoiréing models has been dramatically improved with more profound research. However, the models built are also getting more complex and require more arithmetic power, which is unsuitable for the mobile and embedded devices. The trade-off between the moiré removal effectiveness and efficiency is a problem worth investigating. Recently, the issue of lightweight models [74], [75] has caught the attention of researchers. For example, deep learning-based lightweight models have achieved excellent results in image denoising [76], deraining [77], deblurring [78], and dehazing [79] studies. Thus constructing lightweight demoiréing models is a potential research direction.

B. LOSS FUNCTIONS

The loss function is considered to establish constraints between moiré and moiré-free images and guide the model for optimization. Different loss functions have different characteristics. In training, various loss functions are often combined by weighting, and the most suitable loss function for image demoiréing is still uncertain. Therefore, the most suitable loss functions and weight assignments should be explored in future work.

C. EVALUATION METRICES

Currently, the most commonly used evaluation metrics in computer vision research are the PSNR and SSIM. However, PSNR only calculates the inter-pixel differences in images without considering the perceptual quality of the images. Although HVS-based SSIM can evaluate images in terms of brightness, contrast, and structure, it cannot accurately measure the perceptual quality of images. The evaluation of demoiréing still faces the challenge of perceived quality and requires further exploration.

D. UNSUPERVISED IMAGE DEMOIRÉING

As analyzed in Section IV, in some cases, it is difficult to get clean images corresponding to moiré images and cannot be trained by supervised learning. Unsupervised learning can solve this problem, but its performance is lower than that of supervised learning. Better unsupervised demoiréing models should be further explored.

Furthermore, autoencoders are a class of neural networks used in semi-supervised and unsupervised learning, which use a back-propagation algorithm to make the output equal to the input. Researchers have already studied autoencoder-based image denoising [80], and image super-resolution [81] and have achieved some attainments. Thus, autoencoder-based image demoiréing studies are a promising direction for future development.

E. TEXTURE IMAGE DEMOIRÉING

Moiré also appears in images when a digital camera is used to capture the richly textured object in practical applications, such as tall buildings and clothing, as shown in Fig. 2. However, there are few deep learning-based studies on texture image demoiréing, except for the NTIRE2020 demoiréing challenge (moiré pattern is synthetic). Therefore, there is also a great demand for research in texture image demoiréing.

F. VIDEO DEMOIRÉING

In addition to moiré when taking photos, moiré can also occur in videos. Researchers have already studied video dehazing [82] and video deblurring [83] techniques and have achieved some progress. Likewise, video demoiréing is bound to be a field waiting to be explored in the future.

REFERENCES

[1] G. Oster and Y. Nishijima, "Moiré patterns," *Sci. Amer.*, vol. 208, no. 5, pp. 54–63, 1963.

[2] 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.*, Sep. 2010, pp. 4305–4308, doi: 10.1109/ICIP.2010.5651784.

[3] D. Menon and G. Calvagno, "Color image demosaicking: An overview," *Signal Process., Image Commun.*, vol. 26, nos. 8–9, pp. 518–533, Oct. 2011, doi: 10.1016/j.image.2011.04.003.

[4] I. Pekkucuksen and Y. Altunbasak, "Multiscale gradients-based color filter array interpolation," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 157–165, Jan. 2013, doi: 10.1109/TIP.2012.2210726.

[5] F. Liu, J. Yang, and H. Yue, "Moiré pattern removal from texture images via low-rank and sparse matrix decomposition," in *Proc. Vis. Commun. Image Process. (VCIP)*, Dec. 2015, pp. 1–4, doi: 10.1109/VCIP.2015.7457907.

[6] J. Yang, X. Zhang, C. Cai, and K. Li, "Demoiréing for screen-shot images with multi-channel layer decomposition," in *Proc. IEEE Vis. Commun. Image Process. (VCIP)*, Dec. 2017, pp. 1–4, doi: 10.1109/VCIP.2017.8305057.

[7] 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.*, vol. 26, no. 7, pp. 3528–3541, Jul. 2017, doi: 10.1109/TIP.2017.2698920.

[8] S. M. Hazavei and H. R. Shahdoosti, "A new method for removing the moiré pattern from images," 2017, *arXiv:1701.09037*.

[9] F. Fang, T. Wang, S. Wu, and G. Zhang, "Removing moiré patterns from single images," *Inf. Sci.*, vol. 514, pp. 56–70, Apr. 2020, doi: 10.1016/j.ins.2019.12.001.

[10] Y. Sun, Y. Yu, and W. Wang, "Moiré photo restoration using multiresolution convolutional neural networks," *IEEE Trans. Image Process.*, vol. 27, no. 8, pp. 4160–4172, Aug. 2018, doi: 10.1109/TIP.2018.2834737.

[11] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 27, 2014, pp. 139–144. Accessed: Mar. 24, 2022. [Online]. Available: https://proceedings.neurips.cc/paper/2014/hash/5ca3e9b122f61f8f06494c97b1afccf3-Abstract.html

[12] H. Yue, Y. Cheng, F. Liu, and J. Yang, "Unsupervised moiré pattern removal for recaptured screen images," *Neurocomputing*, vol. 456, pp. 352–363, Oct. 2021, doi: 10.1016/j.neucom.2021.05.099.

[13] H. Park, A. Gia Vien, H. Kim, Y. Jun Koh, and C. Lee, "Unpaired screen-shot image demoiréing with cyclic moiré learning," *IEEE Access*, vol. 10, pp. 16254–16268, 2022, doi: 10.1109/ACCESS.2022.3149478.

[14] X. Cheng, Z. Fu, and J. Yang, "Multi-scale dynamic feature encoding network for image demoiréing," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshop (ICCVW)*, Oct. 2019, pp. 3486–3493, doi: 10.1109/ICCVW.2019.00432.

[15] C. Sun, H. Lai, L. Wang, and Z. Jia, "Efficient attention fusion network in wavelet domain for demoireing," *IEEE Access*, vol. 9, pp. 53392–53400, 2021, doi: 10.1109/ACCESS.2021.3070809.

[16] B. Liu, X. Shu, and X. Wu, "Demoiréing of camera-captured screen images using deep convolutional neural network," 2018, *arXiv:1804.03809*.

[17] Y. Guo, C. Ji, X. Zheng, Q. Wang, and X. Luo, "Multi-scale multi-attention network for moiré document image binarization," *Signal Process., Image Commun.*, vol. 90, Jan. 2021, Art. no. 116046, doi: 10.1016/j.image.2020.116046.

[18] T. Gao, Y. Guo, X. Zheng, Q. Wang, and X. Luo, "Moiré pattern removal with multi-scale feature enhancing network," in *Proc. IEEE Int. Conf. Multimedia Expo Workshops (ICMEW)*, Jul. 2019, pp. 240–245, doi: 10.1109/ICMEW.2019.00048.

[19] B. Zheng, S. Yuan, G. Slabaugh, and A. Leonardis, "Image demoireing with learnable bandpass filters," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 3636–3645. Accessed: Dec. 5, 2021. [Online]. Available: https://openaccess.thecvf.com/content_CVPR_2020/html/Zheng_Image_Demoireing_with_Learnable_Bandpass_Filters_CVPR_2020_paper.html

[20] B. He, C. Wang, B. Shi, and L. Duan, "FHDe²Net: Full high definition demoireing network," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, vol. 2020, pp. 713–729.

[21] H. Yue, Y. Mao, L. Liang, H. Xu, C. Hou, and J. Yang, "Recaptured screen image demoiréing," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 1, pp. 49–60, Jan. 2021, doi: 10.1109/TCSVT.2020.2969984.

[22] H. Wang, Q. Tian, L. Li, and X. Guo, "Image demoiréing with a dual-domain distilling network," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2021, pp. 1–6, doi: 10.1109/ICME51207.2021.9428091.

[23] L.-W. Kang, C. Lo, and C.-H. Yeh, "Deep learning-based moiré pattern removal from a single image: A survey and comparative study," in *Proc. IEEE Int. Conf. Consum. Electron.-Taiwan (ICCE-Taiwan)*, Sep. 2020, pp. 1–2, doi: 10.1109/ICCE-TAIWAN49838.2020.9258241.

[24] W. Liang, "Solutions to the problem happened when taking photos of LED display by digital camera," *Adv. Display*, vol. 20, no. 7, pp. 34–37, 2014.

[25] K. Lu, "Probe into the problem of moiré pattern of studio LED large screen," *Telecom Power Technol.*, vol. 37, no. 6, pp. 245–247, 2020.

[26] X. Li, "Moiré pattern elimination on LED screen for TV studio," *Entertainment Technol.*, vol. 8, pp. 44–47, Jan. 2013.

[27] Q. Huynh-Thu and M. Ghanbari, "Scope of validity of PSNR in image/video quality assessment," *Electron. Lett.*, vol. 44, no. 13, pp. 800–801, Jun. 2008.

[28] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004, doi: 10.1109/TIP.2003.819861.

[29] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Salt Lake City, UT, USA, Jun. 2018, pp. 586–595, doi: 10.1109/CVPR.2018.00068.

[30] S. Yuan, R. Timofte, G. Slabaugh, and A. Leonardis, "AIM 2019 challenge on image demoiréing: Dataset and study," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshop (ICCVW)*, Oct. 2019, pp. 3526–3533, doi: 10.1109/ICCVW.2019.00437.

[31] L. Liu, J. Liu, S. Yuan, G. Slabaugh, A. Leonardis, W. Zhou, and Q. Tian, "Wavelet-based dual-branch network for image demoiréing," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Aug. 2020, pp. 86–102.

[32] S. Yuan, R. Timofte, A. Leonardis, and G. Slabaugh, "NTIRE 2020 challenge on image demoireing: Methods and results," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2020, pp. 460–461. Accessed: Dec. 5, 2021. [Online]. Available: https://openaccess.thecvf.com/content_CVPRW_2020/html/w31/Yuan_NTIRE_2020_Challenge_on_Image_Demoireing_Methods_and_Results_CVPRW_2020_paper.html

[33] U. Sara, M. Akter, and M. S. Uddin, "Image quality assessment through FSIM, SSIM, MSE and PSNR—A comparative study," *J. Comput. Commun.*, vol. 7, no. 3, pp. 8–18, 2019, doi: 10.4236/jcc.2019.73002.

[34] A. G. Vien and C. Lee, "Single-shot high dynamic range imaging via multiscale convolutional neural network," *IEEE Access*, vol. 9, pp. 70369–70381, 2021, doi: 10.1109/ACCESS.2021.3078457.

[35] S. Yuan et al., "AIM 2019 challenge on image demoireing: Methods and results," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshop (ICCVW)*, Oct. 2019, pp. 3534–3545, doi: 10.1109/ICCVW.2019.00438.

[36] S. Gu et al., "AIM 2019 challenge on image extreme super-resolution: Methods and results," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshop (ICCVW)*, Oct. 2019, pp. 3556–3564, doi: 10.1109/ICCVW.2019.00440.

[37] S. Nah et al., "AIM 2019 challenge on video temporal super-resolution: Methods and results," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshop (ICCVW)*, Oct. 2019, pp. 3388–3398, doi: 10.1109/ICCVW.2019.00421.

[38] A. Lugmayr, M. Danelljan, and R. Timofte, "NTIRE 2020 challenge on real-world image super-resolution: Methods and results," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2020, pp. 494–495. Accessed: Mar. 25, 2022. [Online]. Available: https://openaccess.thecvf.com/content_CVPRW_2020/html/w31/Lugmayr_NTIRE_2020_Challenge_on_Real-World_Image_Super-Resolution_Methods_and_Results_CVPRW_2020_paper.html

[39] S. Nah, S. Son, R. Timofte, and K. M. Lee, "NTIRE 2020 challenge on image and video deblurring," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2020, pp. 416–417. Accessed: Mar. 25, 2022. [Online]. Available: https://openaccess.thecvf.com/content_CVPRW_2020/html/w31/Nah_NTIRE_2020_Challenge_on_Image_and_Video_Deblurring_CVPRW_2020_paper.html

[40] C. O. Ancuti, C. Ancuti, F.-A. Vasluianu, and R. Timofte, "NTIRE 2020 challenge on nonhomogeneous dehazing," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2020, pp. 490–491. Accessed: Mar. 25, 2022. [Online]. Available: https://openaccess.thecvf.com/content_CVPRW_2020/html/w31/Ancuti_NTIRE_2020_Challenge_on_NonHomogeneous_Dehazing_CVPRW_2020_paper.html

[41] B. Zheng, S. Yuan, C. Yan, X. Tian, J. Zhang, Y. Sun, L. Liu, A. Leonardis, and G. Slabaugh, "Learning frequency domain priors for image demoireing," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 11, pp. 7705–7717, Nov. 2022, doi: 10.1109/TPAMI.2021.3115139.

[42] X. Cheng, "Improved multi-scale dynamic feature encoding network for image demoiréing," *Pattern Recognit.*, vol. 116, Aug. 2021, Art. no. 107970.

[43] B. He, C. Wang, B. Shi, and L.-Y. Duan, "Mop moiré patterns using MopNet," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, Oct. 2019, pp. 2424–2432. Accessed: Dec. 5, 2021. [Online]. Available: https://openaccess.thecvf.com/content_ICCV_2019/html/He_Mop_Moire_Patterns_Using_MopNet_ICCV_2019_paper.html

[44] A. Odena, V. Dumoulin, and C. Olah, "Deconvolution and checkerboard artifacts," *Distill*, vol. 1, no. 10, p. e3, Oct. 2016, doi: 10.23915/distill.00003.

[45] A. G. Vien, H. Park, and C. Lee, "Moiré artifacts removal in screenshot images via multiple domain learning," in *Proc. Asia–Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, Dec. 2020, pp. 1268–1273.

[46] S. Yang, Y. Lei, S. Xiong, and W. Wang, "High resolution demoire network," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2020, pp. 888–892, doi: 10.1109/ICIP40778.2020.9191255.

[47] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778. Accessed: Mar. 11, 2022. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html

[48] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4700–4708. Accessed: Mar. 11, 2022. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2017/html/Huang_Densely_Connected_Convolutional_CVPR_2017_paper.html

[49] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," 2016, *arXiv:1511.07122*.

[50] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141. Accessed: Mar. 14, 2022. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2018/html/Hu_Squeeze-and-Excitation_Networks_CVPR_2018_paper.html

[51] L. Liu, S. Yuan, J. Liu, L. Bao, G. Slabaugh, and Q. Tian, "Self-adaptively learning to demoire from focused and defocused image pairs," 2020, *arXiv:2011.02055*.

[52] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," 2015, *arXiv:1505.04597*.

[53] B. Mamidibathula and P. K. Biswas, "SVDocNet: Spatially variant U-Net for blind document deblurring," in *Proc. Workshop Document Intell. NeurIPS*, Sep. 2019, pp. 1–4. Accessed: Mar. 12, 2022. [Online]. Available: https://openreview.net/forum?id=Hyx3f65qLS

[54] Z. Lu and Y. Chen, "Single image super-resolution based on a modified U-Net with mixed gradient loss," *Signal, Image Video Process.*, vol. 16, pp. 1143–1151, Nov. 2021, doi: 10.1007/s11760-021-02063-5.

[55] S. Lee, M. Negishi, H. Urakubo, H. Kasai, and S. Ishii, "Mu-Net: Multi-scale U-Net for two-photon microscopy image denoising and restoration," *Neural Netw.*, vol. 125, pp. 92–103, May 2020, doi: 10.1016/j.neunet.2020.01.026.

[56] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

[57] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–11. Accessed: Jul. 1, 2022. [Online]. Available: https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html

[58] F. Yang, H. Yang, J. Fu, H. Lu, and B. Guo, "Learning texture transformer network for image super-resolution," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 5791–5800. Accessed: Jul. 1, 2022. [Online]. Available: https://openaccess.thecvf.com/content_CVPR_2020/html/Yang_Learning_Texture_Transformer_Network_for_Image_Super-Resolution_CVPR_2020_paper.html

[59] H. Chen, Y. Wang, T. Guo, C. Xu, Y. Deng, Z. Liu, S. Ma, C. Xu, C. Xu, and W. Gao, "Pre-trained image processing transformer," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 12299–12310. Accessed: Jul. 1, 2022. [Online]. Available: https://openaccess.thecvf.com/content/CVPR2021/html/Chen_Pre-Trained_Image_Processing_Transformer_CVPR_2021_paper.html

[60] Z. Wang, X. Cun, J. Bao, W. Zhou, J. Liu, and H. Li, "Uformer: A general U-shaped transformer for image restoration," 2021, *arXiv:2106.03106*.

[61] L. Liu, L. Xie, X. Zhang, S. Yuan, X. Chen, W. Zhou, H. Li, and Q. Tian, "TAPE: Task-agnostic prior embedding for image restoration," 2022, *arXiv:2203.06074*.

[62] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Proc. Eur. Conf. Comput. Vis.* Amsterdam, The Netherlands: Springer, Oct. 2016, pp. 694–711.

[63] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255, doi: 10.1109/CVPR.2009.5206848.

[64] H. Park, A. G. Vien, Y. J. Koh, and C. Lee, "Unpaired image demoiréing based on cyclic moiré learning," in *Proc. Asia–Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, 2021, pp. 146–150.

[65] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 1026–1034. Accessed: Mar. 13, 2022. [Online]. Available: https://openaccess.thecvf.com/content_iccv_2015/html/He_Delving_Deep_into_ICCV_2015_paper.html

[66] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Venice, Italy, Oct. 2017, pp. 2242–2251, doi: 10.1109/ICCV.2017.244.

[67] M. Mirza and S. Osindero, "Conditional generative adversarial nets," 2014, *arXiv:1411.1784*.

[68] A. Y. Jazi, J. J. Liu, and H. Lee, "Automatic inspection of TFT-LCD glass substrates using optimized support vector machines," *IFAC Proc. Volumes*, vol. 45, no. 15, pp. 325–330, 2012, doi: 10.3182/20120710-4-SG-2026.00054.

[69] S. Mei, H. Yang, and Z. Yin, "Unsupervised-learning-based feature-level fusion method for Mura defect recognition," *IEEE Trans. Semicond. Manuf.*, vol. 30, no. 1, pp. 105–113, Feb. 2017, doi: 10.1109/TSM.2017.2648856.

[70] H.-P. Lu, C.-T. Su, S.-Y. Yang, and Y.-P. Lin, "Combination of convolutional and generative adversarial networks for defect image demoiréing of thin-film transistor liquid-crystal display image," *IEEE Trans. Semicond. Manuf.*, vol. 33, no. 3, pp. 413–423, Aug. 2020, doi: 10.1109/TSM.2020.3005164.

[71] H.-P. Lu and C.-T. Su, "CNNs combined with a conditional GAN for Mura defect classification in TFT-LCDs," *IEEE Trans. Semicond. Manuf.*, vol. 34, no. 1, pp. 25–33, Feb. 2021.

[72] J.-H. Kim, K. Kong, and S.-J. Kang, "Image demoireing via U-Net for detection of display defects," *IEEE Access*, early access, Jun. 27, 2022, doi: 10.1109/ACCESS.2022.3186685.

[73] J. Chen, J. Zhu, Z. Li, W. Shi, Q. Zhang, Z. Hu, H. Zheng, D. Liang, and Y. Ge, "Automatic image-domain moiré artifact reduction method in grating-based X-ray interferometry imaging," *Phys. Med. Biol.*, vol. 64, no. 19, Oct. 2019, Art. no. 195013, doi: 10.1088/1361-6560/ab3c34.

[74] X. Zhang, X. Zhou, M. Lin, and J. Sun, "ShuffleNet: An extremely efficient convolutional neural network for mobile devices," 2017, *arXiv:1707.01083*.

[75] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*.

[76] T.-H. Le, P.-H. Lin, and S.-C. Huang, "LD-Net: An efficient lightweight denoising model based on convolutional neural network," *IEEE Open J. Comput. Soc.*, vol. 1, pp. 173–181, 2020, doi: 10.1109/OJCS.2020.3012757.

[77] X. Fu, B. Liang, Y. Huang, X. Ding, and J. Paisley, "Lightweight pyramid networks for image deraining," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 6, pp. 1794–1807, Jun. 2020, doi: 10.1109/TNNLS.2019.2926481.

[78] Y. Zhang, Y. Liu, Q. Li, M. Qi, D. Xu, J. Kong, and J. Wang, "Image deblurring based on lightweight multi-information fusion network," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Anchorage, AK, USA, Sep. 2021, pp. 1724–1728, doi: 10.1109/ICIP42928.2021.9506606.

[79] H. Ullah, K. Muhammad, M. Irfan, S. Anwar, M. Sajjad, A. S. Imran, and V. Hugo C. de Albuquerque, "Light-DehazeNet: A novel lightweight CNN architecture for single image dehazing," *IEEE Trans. Image Process.*, vol. 30, pp. 8968–8982, 2021, doi: 10.1109/TIP.2021.3116790.

[80] J. Jiao, L. Bao, Y. Wei, S. He, H. Shi, R. Lau, and T. S. Huang, "Laplacian denoising autoencoder," 2020, *arXiv:2003.13623*.

[81] D. Chira, I. Haralampiev, O. Winther, A. Dittadi, and V. Liévin, "Image super-resolution with deep variational autoencoders," 2022, *arXiv:2203.09445*.

[82] Y. Park and T.-H. Kim, "Fast execution schemes for dark-channel-prior-based outdoor video dehazing," *IEEE Access*, vol. 6, pp. 10003–10014, 2018, doi: 10.1109/ACCESS.2018.2806378.

[83] R. Jiang, L. Zhao, T. Wang, J. Wang, and X. Zhang, "Video deblurring via temporally and spatially variant recurrent neural network," *IEEE Access*, vol. 8, pp. 7587–7597, 2020, doi: 10.1109/ACCESS.2019.2962505.

KAI LI received the bachelor's degree from Henan Polytechnic University, China, in 2020, where he is currently pursuing the master's degree in computer science and technology software engineering. His current research interests include image processing and image enhancement.

YINGGANG ZHAO received the Ph.D. degree in computer science and technology from Zhejiang University, China, in 2007. His research interests include deep learning, computer vision, and hyperspectral image processing.

SHOUMING HOU received the B.S. degree from the Jiaozuo Mining Institute, in 1997, the M.S. degree in engineering from the Huazhong University of Science and Technology, in 2000, and the Ph.D. degree in engineering from Northeastern University, in 2010. He is currently a Professor with the School of Computer Science and Technology, Henan Polytechnic University, China. His research interests include engineering digitization and simulation, virtual reality, and intelligent systems.

BAOYUN LU received the Ph.D. degree in control theory and control engineering from the Chinese Academy of Sciences, in 2011. She is currently a Lecturer at the School of Software, Henan Polytechnic University. Her research interest includes data modeling and analysis.

YABING WANG received the bachelor's degree in electronic information from Henan Agricultural University, in 2019. He is currently pursuing the master's degree in software engineering with the School of Computer Science and Technology, Henan Polytechnic University. His current research interests include image processing and deep learning.

LIYA FAN received the B.S. degree in computer information management from Xidian University, Shanxi, China, in 2002, and the M.S. degree in management science and engineering from the Xi'an University of Technology, Shanxi, in 2007. From 2007 to 2022, she worked at The City College, Xi'an Jiaotong University. From 2020 to 2022, she worked as an Associate Professor with the Computer Department, The City College, Xi'an Jiaotong University. Her research interest includes the development of AR/VR technology and application.

● ● ●