Rain-like Layer Removal from Hot-Rolled Steel Strip Based on Attentive Dual Residual Generative Adversarial Network

Qiwu Luo, *Senior Member, IEEE*, Handong He, Kexin Liu, Chunhua Yang, *Fellow, IEEE*, Olli Silven, *Senior Member, IEEE*, and Li Liu, *Senior Member, IEEE*

Abstract—**Rain-like layer removal from hot-rolled steel strip surface is proved to be a workable measure for suppressing the false alarms frequently triggered in automated visual inspection (AVI) instrument. This paper extends the scope of "rain-like layer" from dispersed waterdrops to splashing water streaks and tiny white droplets. And a targeted method with both channel-wise and spatial-wise attention, namely attentive dual residual generative adversarial network (ADRGAN), is proposed. Meanwhile, a newly updated steel surface image dataset with typical natures of "rain-like layer" gathered from actual hot-rolling line, Steel_Rain, is opened for the first time. The comparison experimental results between our proposed network and *eleven* prestigious networks show that our ADRGAN-restored images are the closest to the ground-truth images on *six* public datasets, especially on the newly-opened industrial dataset Steel_Rain, it yeilds the best scores of 56.8627 peak signal to noise ratio (PSNR), 0.9980 structural similarity index (SSIM), 0.134 mean-square error (MSE) and 0.006 learned perceptual image patch similarity (LPIPS). In the final verification test, the concept of rain-like layer removal has been proved to perform best in defect inspection, where *four* traditional defect detection algorithms are involved. And as expected, defect detection methods assisted by ADRGAN yield the minimum false-alarms[1].**

Index Terms—**Automated visual inspection (AVI), hot-rolled steel strip, rain-like layer removal, generative adversarial network (GAN).**

I. INTRODUCTION

Steel is one of the fundamental materials for manufacturing enterprises, and its quality seriously affects the production of many subsequent industrial chains. The instrument of automatic visual inspection (AVI) has great practical value in ensuring the quality of steel products [1] [2] [3]. It is worth
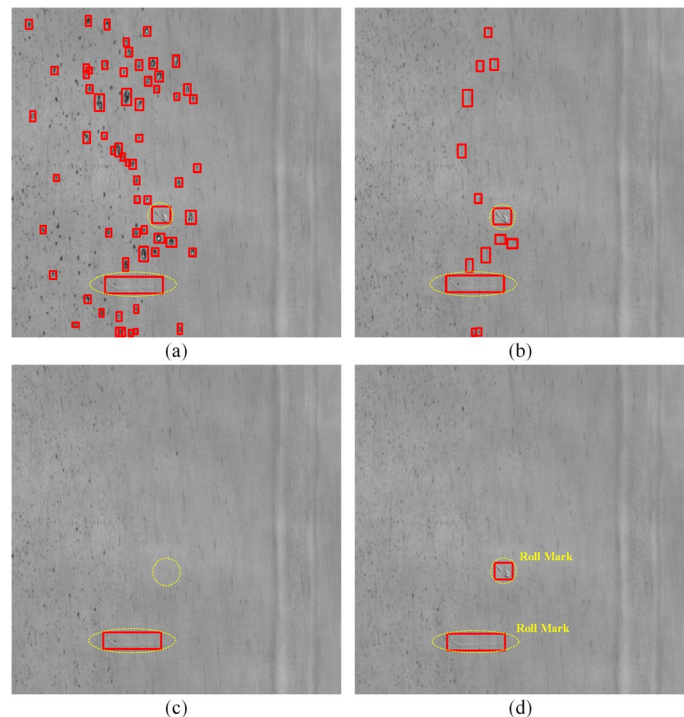
Fig.1 The influence of different de-raining algorithms on real defect detection. (a) The original hot-rolled steel strip image. (b) Detection result of PReGAN-restored images. (c) DuRN-PSP. (d) Ours. Some rain-like pseudo defects are incorrectly identified as real defects in (b), and some real defects are lost in (c) due to over removal of rain-like layer. Our proposed network (ADRGAN) can help the AVI instrument correctly locate the real defects.

noting that the layout of average hot-rolling production lines is pretty compact, thus there is always limited installation space for the AVI instrument. Taking Valin Steel for an example, the total length from the finishing mill F7 to the recoiling machine is about 18 meters, excluding the positions occupied by the laminar cooling devices and hot-rolling mill instruments, our AVI instrument can only be arranged around 6 meters closed to the recoiling machine, where is at the downstream site of the laminar colling devices. Therefore, it is inevitable that some waterdrops, water clothes or rain lines, essentially pseudo defects [4], are randomly distributed on the images of hot-rolled steel strip, triggering false alarms. Worse still, some real defects are partially or completely covered. Consequently, the defect inspection accuracy and efficiency of AVI instrument decrease sharply.

Many studies have attempted to detect [5]-[10] and classify [11] [12] [13] various defects directly from original steel strip images captured in the above-mentioned harsh environment,

but they have some limitations in practical application. Defects of hot-rolled steel strip can be generally divided into two categories: Periodic defects and occasional defects. The frequency of occasional defects is low, so it is generally difficult to accumulate enough image samples for the training of neural networks, while occasional defects cannot be ignored in quality inspection of steel strips. Therefore, algorithms based on statistical learning [14]-[17] are usually used in practical industrial production lines, which are disturbed by rain-like pseudo defects.

Aiming at solving the above problem, a fine-grained image restoration algorithm, PReGAN, was proposed in [18], making attempt to locate and remove waterdrops while preserving the real defects. It takes GAN [37] as the backbone and adopts the progressive image rainline removal network (PReNet) proposed by Ren *et al.* [38] as the generator. The achieved effects are mainly based a concept of considering the single pseudo defect of waterdrops as "rain-like layer". But the typical pseudo defects of splashing water streaks and tiny white droplets are totally ignored. Thus, it has removal limitation when being applied in the actual production line even using our improved dataset from Section IV for training. Since the distribution of cooling water is uneven in an image, PReGAN tends to have some under de-rain problem caused by the area with high rain density.

As declared in [19], the under and over de-raining are the major challenges for rain-like layer removal. As shown in Fig. 1, under de-raining increases the number of false alarms because rain-like pseudo defects are not adequately removed. Instead, over de-raining incorrectly removes some real defects and greatly affects the accuracy of defect detection. The primary task for solving this problem is to construct an image restoration method with strong robustness to prevent under (or over) de-rain problem.

This paper is an extensibility study of the previous works [18]. We attempt to propose a global search algorithm with both channel-wise and spatial-wise attention to remove hybrid rain-like pseudo defect clusters on the hot-rolled steel surface while preserving edge and texture details, as well as to improve image dataset construction method for rain-like layer removal in [18]. The paper highlights are listed as below.

(1) We *found* that the rain-like layer should not only focus on the *single* pseudo defect of waterdrops [18], but consider the splashing water streaks suspended in the imaging space and the tiny water droplets caused by the high-speed rolling of the steel strip. Therefore, we *generalize the concept of rain-like layer* to massive dispersed waterdrops, splashing water streaks and tiny white droplets, with typical rain-like characteristics, to make the removal of hybrid pseudo defect clusters on the surface of the steel strip *more universal*.

(2) For improving the *generalization* of the algorithm, *we propose a novel rain-like layer removal method of attentive dual residual generative adversarial network (ADRGAN), which is compatible with channel-wise attention and spatial-wise attention.* For channel-wise attention: A dual-residual-based periodic structure with double SE configuration, namely DuRN-PSP, is designed as the generator for tracking and restoring the generalized rain-like pseudo defects aggressively.

For spatial-wise attention: A conventional attention scheme with manual mask is imported between the generator and the discriminator to form a self-optimizing closed-loop in the training stage, and the prior knowledge of the generalized rain-like layer is implicitly but cleverly mined to deal with the occasional misidentification resulting from the aggressive searching stratagem based on pure data processing in the DuRN-PSP.

(3) The splashing water streaks and tiny white droplets are added to expand our previously-opened raindrop removal dataset only focusing waterdrops on steel surface [18]. These 1,455 pairs of images with rain-like pseudo defects might be an open testbench, stimulating the focused rain removal topic.

This paper consists of six sections. After an introduction, Section II reviews the related works of rain streaks and raindrop removal. Section III and IV demonstrate the details of the proposed ADRGAN and rain-like layer dataset respectively. Before concluding this paper in Section VI, extensive experiments are analyzed and discussed in Section V to compare the proposed ADRGAN network with stage-of-the-art de-raining methods.

II. RELATED WORKS

Rain-like layer removal, which mainly tests the global attention and the image restoration ability of networks, can be divided into water streak removal and waterdrop removal. In recent years, Rain Detection and Removal has long flourished in the field of image restoration and video quality enhancement [20]-[38], providing references for our works.

A. Rain streaks

Kang *et al.* [20] used sparse coding to extract the rain streaks components from the high-frequency (HF) parts of rainy images. For separating rain streaks from background images, Luo *et al.*[21] proposed a discriminative sparse coding algorithm. Later, Li *et al.* [22] considered the correlation between different stages of de-raining and proposed a neural network to remove rain streaks step by step. However, it has a limitation that the performance of the input images is affected by the number of iterations, which is not a uniform criterion. Hu *et al.* [25] proposed a new image recognition struction—Squeeze-and-Excitation (SE) Block, and enabled it to reinforce the features of important channels and weaken others, so as to effectively enhance the attention of the network. Deng *et al.* [28] put forward DRD-Net, a combination of rain residual network and detail repair network, within which the former combines SE operation with residual blocks to remove rain streaks by making full use of spatial context information. The experimental results demonstrated that implementing SE Block can efficiently advance the attention model (AM). Du *et al.* [30] proposed a method based on conditional variational auto-encoder to predict rainy images under different spatial locations and color channels. Jiang *et al.* [31] proposed IADN, which decomposed rain streaks into multiple rain layers and abstracted them at multiple levels. IADN can learn the correlation of overall spatial features to make use of similar features of rain information, and adopt a hybrid attention mechanism to guide its optimization direction,
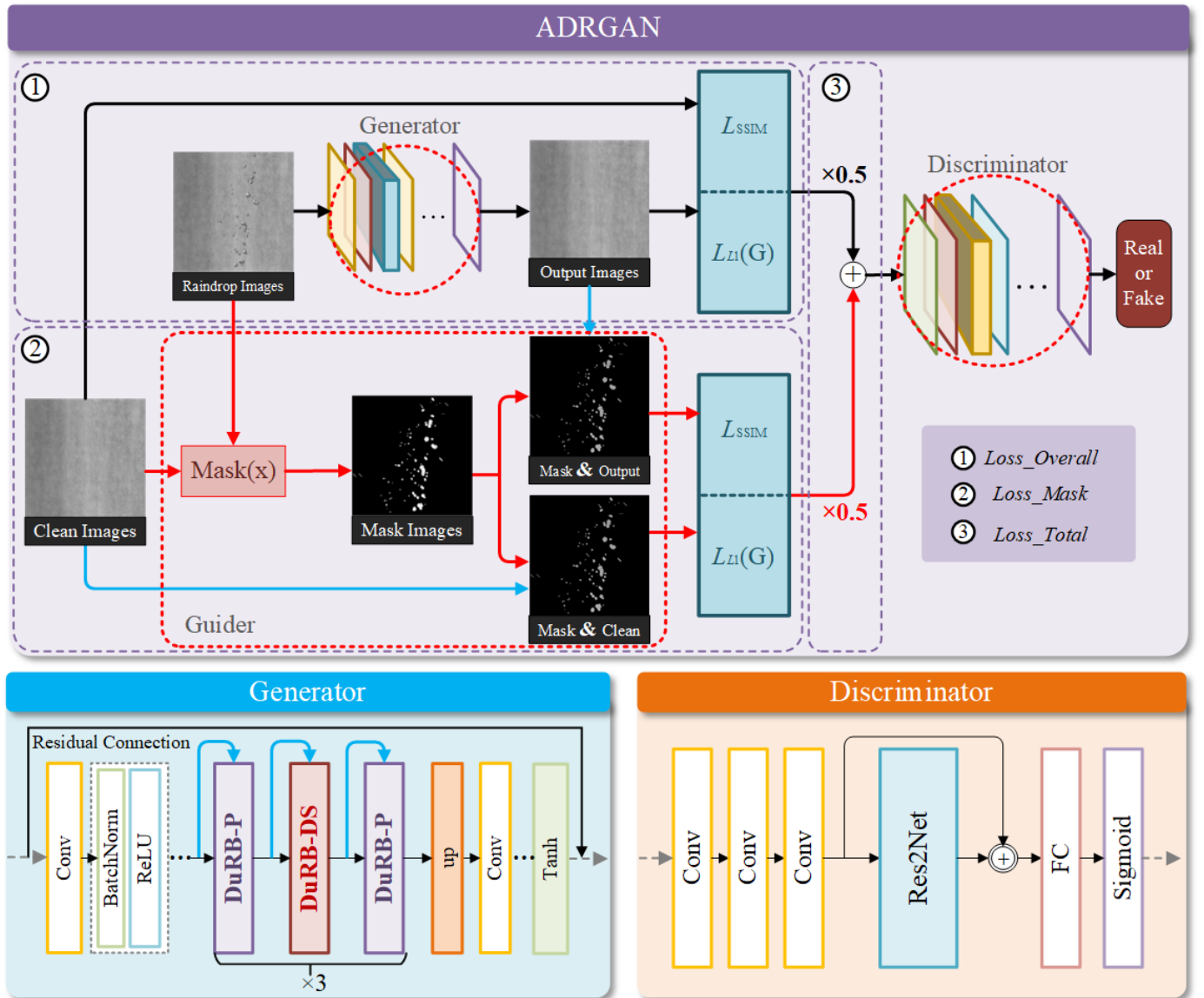
Fig.2 The overall structure of the proposed attentive dual residual generative adversarial network (ADRGAN). It basically contains three parts: generator, guider and discriminator. The generator restores input rain-like layer images to clean background images. The discriminator distinguishes the authenticity of generated images. The guider is an attention mechanism to guide the above two networks by making full use of the prior knowledge of rain-like pseudo defect distribution.

so as to ensure the removal of rain streaks while preserving rich textural details. Vincent *et al.* [51] proposed denoising auto-encoder (DAE) and combined it with a depth generation model, to obtain better images denoising results, which can also be used in rain streak removal.

B. Raindrops

Some methods also have been proposed for raindrop detection and removal. Roser and Geiger [32] came up with a way to compare synthetic raindrops with patches which might contain massive raindrops, but the mutability of raindrop morphology prevents it from being treated in the same way. For raindrop removal, Eigen *et al.* [33] used a pair of normal and rainy images with the same background to train convolutional neural network (CNN). This method only effectively removes sparse small raindrops, while being completely helpless for larger amd dense raindrops, which may be caused by the limited network capacity and the loss function insufficient to provide comprehensive limiting factors.

In [34], the performance of K-means clustering and median filtering in raindrop removal is also unappealing. The effectiveness of pairwise operation in various image processing tasks is proved by Liu *et al.* [35], they put forward an innovative way of residual connection, namely dual residual connection. They use this connection to design a structure called DuRN-S-P to solve the problem of hybrid rain layers, and achieve better results in comparison with Attentive GAN [23]. Pix2Pix [36] applies GAN to image-to-image translation, which can transform one image into another. This general mapping can also be used for the removal of raindrops.

C. Mask Mechanism

Mask is an attention mechanism that pays more attention to spatial information, which has great practical value in rain removal. In [23], Qian *et al* put forward Attentive GAN, which introduces the visual attention mechanism and injects it into both generation and discrimination networks. Ahn *et al.* [19] proposed a method including two sequential networks.
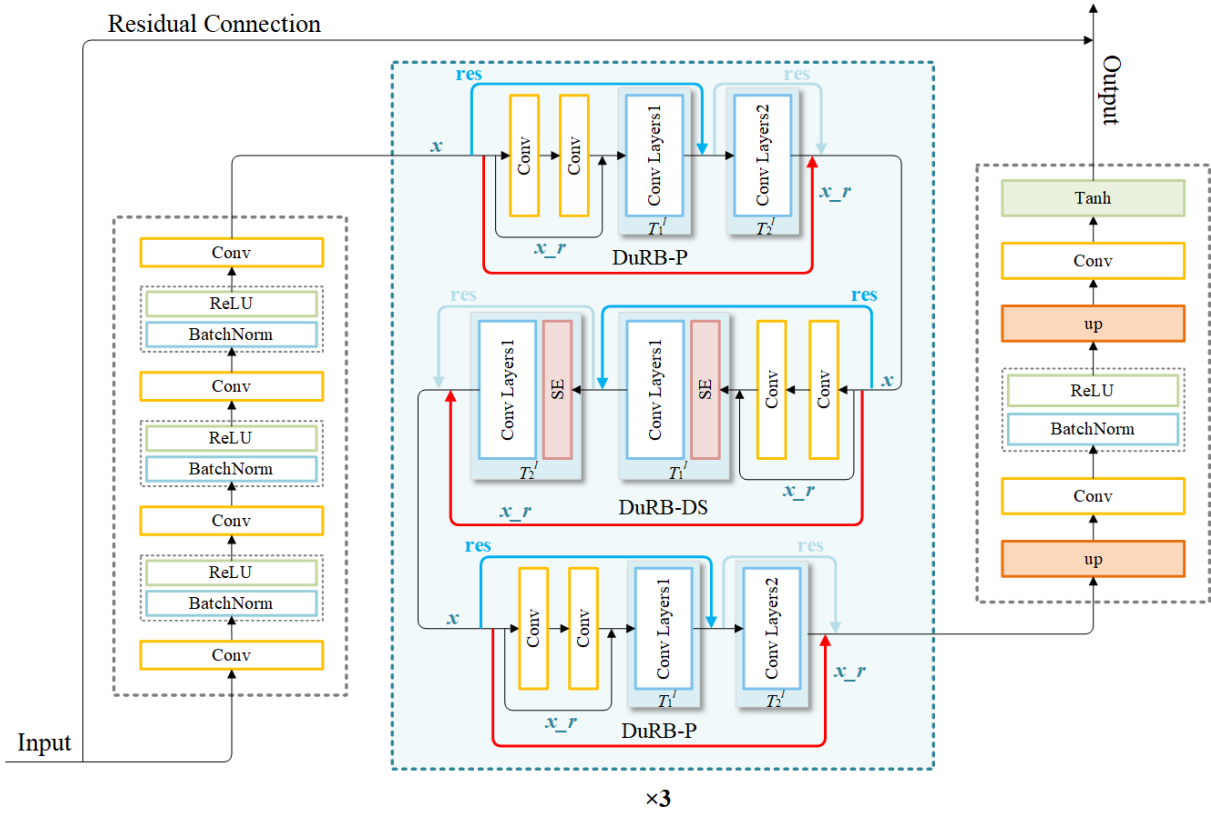
Fig.3 The primary structure of DuRN-PSP, a hybrid dual residual network composing with one DuRB-P, one DuRB-DS and one DuRB-P head to tail for three times. This dual-residual-based three-cycle search-repair structure is dedicated to make better balance between defect recognition and restoration, by avoiding the occasional misidentification resulting from aggressive searching stratagem especially in cases of vastly different sizes of water droplets.
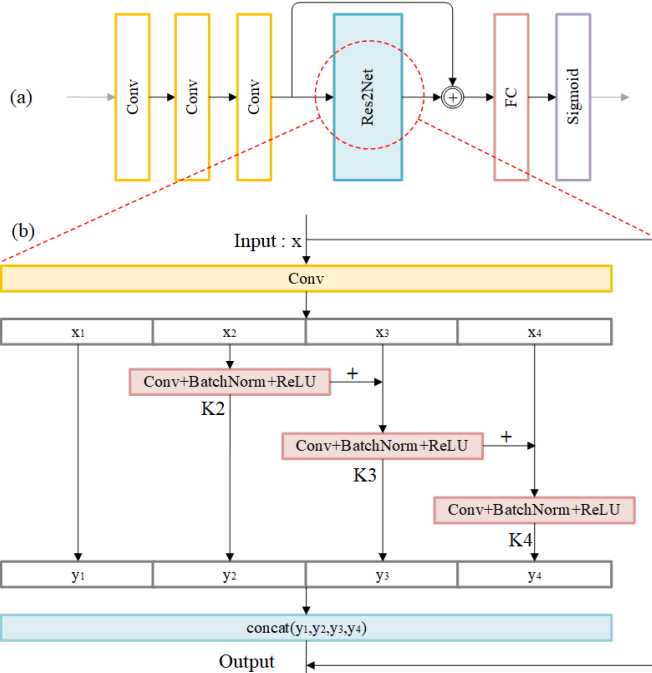

Fig.4 (a) is the architecture of discriminator; Conv is a convolutional layer with 3×3 kernels; (b) is the structure of Res2Net.

The network can estimate the rain maps and transmit them into the image restoration network for rain streaks removal. Yang *et al.* [29] proposed a recurrent dilated network, which can estimate the rain streaks image with the help of the rain mask and remove the rain region on each pixel. Those methods will produce an attention map indicating raindrop regions before removing the raindrops. But they have the limitation that the attention map cannot be optimized dynamically during rain degradation, leading to the inadequate removal of tiny raindrops. In addition, rain streaks images cannot fully reflect the density information, leading to over de-raining in some areas of the generated images. To alleviate the above problems, a generator with periodic structure is proposed in ADRGAN for attentive backtracking. The mask images used in our guider can instruct our generator to make attention corrections, dynamically update the rain distribution image during rain degradation.

The formation mechanism and morphology of rain-like layer are different from the real raindrops and rain streaks. The above-mentioned methods can hardly be applied to the rain-like layer removal from hot-rolled steel strip. Based on these research background, a global search algorithm called ADRGAN is designed in this paper to remove hybrid rain-like pseudo defects, while multiple attention mechanisms are introduced to enhance detail-preserving image restoration ability. In Section V, some experimental results using the same dataset to train are presented among our ADRGAN and some state-of-art methods for comparison.

III. RAIN-LIKE LAYER REMOVAL USING ATTENTIVE DUAL RESIDUAL GENERATIVE ADVERSARIAL NETWORK

Fig. 2 illustrates the overall structure of our ADRGAN framework. The generalized application of GAN from game theory to deep learning neural networks can produce clearer

and more realistic results by using adversarial training. Therefore, we develop our rain-like layer removal network as GAN [37] style, it basically contains three parts: generator, guider and discriminator. The goal function of GAN is defined as:

$$V(G,D) = \mathbf{E}_{R \sim p_{clean}} \left[\log(D(\mathbf{R})) \right]$$
$$+ \mathbf{E}_{I \sim P_{rain-like\ layer}} \left[\log(1 - D(G(\mathbf{I}))) \right] \quad (1)$$

where G and D refers to the generator and discriminator separately. $\mathbf{I}$, worked as the input of generator, is the sample with rain-like layer. $\mathbf{R}$ means ground-truth clean images. The goal of G is to generate near-real images to deceive D, while D needs to find the difference between generated images and real images as much as possible:

$$G^* = \arg \min_G \max_D V(G,D) \quad (2)$$

$$D^* = \arg \max_D V(D,G) \quad (3)$$

The whole network takes pairs of images as input, including ground-truth clean images and their corresponding images with rain-like layer. The generator we developed will generate pristine images to its largest realistic extent. The guider can instruct the generator and the discriminator by comparing the differences between generated images and their corresponding clean images. For the over de-rain problem caused by low frequency components, attention mask images are introduced to reflect the information about rain intensity in spatial dimension. In this way, the generator is guided to estimate the distribution of pseudo defects, so as to restore images purposefully. For the under de-rain problem caused by high frequency components, the generated images will be transferred to the discriminator for authenticity identification. The following introduces the detailed design and innovation of our network.

A. Generator

As shown in Fig. 3, an attentive dual residual network called DuRN-PSP is designed as the generator of GAN, which is based on encoder-decoder structure. As for its bottleneck, a periodic combinatorial structure is proposed, each consists of two DuRB-P and one DuRB-DS in between.

DuRB-P is an image restoration block,which makes the best of the context information and increase the capability of multi-scale feature representation by increasing the receptive fields of convolution. DuRB-DS is a channel-wise attention block, which can localize rain-like pseudo defects in a coarse-to-fine manner by decreasing dilation rates (12, 8, and 6) in the forward direction. The role of Dual Residual Blocks in each stage is explained as follows.

The fundamental framework of the dual residual blocks (DuRB) with dual residual connections are exhibited in Fig. 3. T_1^l and T_2^l are the containers for paired operations, which represent up-sampling and down-sampling respectively. x_r refers to residual information and its initial value is the input feature map x. Conv indicates convolution operation. The normalization operation and the rectified linear unit (ReLU) [39] are performed after each Conv.

The DuRB-P include three aspects:

(1) After receiving x, the Conv is performed twice and x_r is added to the output value, which can accelerate the convergence rate of the model. The final output is passed to T_1^l.

(2) The whole process of T_1^l mainly up-samples the input values and adds to res after that. It is worth noting that this final output of T_1^l will be the res input for the next dual residual block.

(3) Corresponding to the previous stage, the convolution layer of T_2^l performs down-sampling operation, and the resulting output combined with x_r will be the input x of the next residual block.

The proposed DuRB-DS is similar to the DuRB-P, except that the SE blocks [25] are set in T_1^l and T_2^l. The double SE configuration has better global attention than only one in the down-sampling process, which helps to infer the distribution of rain-like pseudo defects from actual steel strip images.

Further, our DuRN-PSP realizes the searching and restoring rain-like layer by repeating triple times in each epoch, for avoiding the occasional misidentification resulting from aggressive searching stratagem especially in cases of vastly different sizes of water droplets. Compared with [35], the fineness of the DuRN-PSP is greatly improved, to ensure tiny pseudo defects can be targeted for locating and eliminating.

Finally, the rain-like layer can be removed by repeatedly passing rainy image through the DuRN-PSP. However, due to the pure data processing in the DuRN-PSP, it is to some blind and unreliable occasionally, and some real defects may be misidentified as pseudo defects and removed (see Fig. 1(c) over de-rain). Hence, we design a reliable guider to deal with the occasional misidentification, and it is introduced in the upcoming subsection.

B. Guider

With the attention blocks [25], the DuRB-DS can obtain information of rain density actively, however, the information of rain streaks intensity in spatial dimension are ignored. Therefore, it is difficult for the generator to understand the correlation of features at different positions in a rainy image. As shown in Fig. 2, mask images are imported between the generator and the discriminator as the guider to form a self-optimizing closed-loop in the training stage, which exploits the potential of the prior knowledge from the generalized rain-like layer, to deal with the problems of edge blurring and loss of detail resulting from insufficient attention to spatial features in the DuRN-PSP. The generation method of mask images is a threshold-based binary classification strategy, the equation can be expressed as:

$$Mask(x) = \begin{cases} 0 & if \ \left| Pixel_{rain-like\ layer} - Pixel_{clean} \right| \leq 30 \\ 1 & else \end{cases} \quad (4)$$

Mask images strengthens the attention of ADRGAN to the rain-like layer, but the edge consistency of restoration areas is ignored. This paper constructed a weighted sum of $L1$ and $SSIM$ [41] loss to supervise the rain-like layer removal. In addition to generating realistic images to deceive the discriminator, the generator is also guided by the loss values of $L1$ and $SSIM$ to recover the local details of the mask and

consider the global features to ensure that the generated image is not distorted. The composition of each loss value is as follows:

$$Loss\_Total = \frac{1}{2}*Loss\_Mask + \frac{1}{2}*Loss\_Overall \quad (5)$$

where *Loss_Mask* indicates the loss value of rain-like layer and *Loss_Overall* represents the loss value of the whole image. Eq. (5) is used to calculate the total loss of L1 or SSIM.

Structural Similarity (*SSIM*) can solve the problem of image distortion, which is an indicator used to measure the similarity of pictures. The image distortion is smaller when its value is larger. Here lists the SSIM loss function in the generator of ADRGAN:

$$L_{\text{SSIM}} = \mathbf{E}_{R \sim p_{clean}, I \sim p_{rain-like\ layer}} [-0.5*SSIM(G(\mathbf{I}), \mathbf{R}) \\ - 0.5*SSIM(G(\mathbf{I}) \circ \mathbf{M}, \mathbf{R} \circ \mathbf{M})] \quad (6)$$

where $\circ$ represents elementwise multiplication, G indicates the generative network. R and I represent rain images and clean images, respectively, and M indicates the corresponding mask images.

Noted as the Least Absolute Deviation (LAD) or the LeastAbsolute Error (LAE), the *L*1 loss function is the error obtained by the sum of the absolute difference between the target value and the estimated value. It comes as:

$$L1(y, y) = \sum_{i=0}^{m} |y^{(i)} - y^{(i)}| \quad (7)$$

where $y^{(i)}$ means each pixel's value on the target image, and $y^{(i)}$ indicates that on the estimation image. Therefore, *L*1 loss function in ADRGAN generator is formulated as:

$$L_{L1}(G) = \mathbf{E}_{R \sim p_{clean}, I \sim p_{rain-like\ layer}} [0.5*||\mathbf{R} - G(\mathbf{I})||_1 \\ + 0.5*||\mathbf{R} \circ \mathbf{M} - G(\mathbf{I}) \circ \mathbf{M}||_1] \quad (8)$$

where $||\cdot||_1$ indicates L1-norm, other variables are defined in the same way as Eq. (6).

Since the GAN objective function is mixed with other losses, the prediction results of the model are comprehensively measured, and the operating performance is improved. Accordingly, the loss function of the final generator on the basis of fusion strategy is expressed as:

$$G^* = \arg \min_G \max_D V(G, D) + (r_1 \cdot L_{L1} + r_2 \cdot L_{\text{SSIM}}) \quad (9)$$

where r_1 is set to 0.75. r_2 is set to 1.1.

The quantitative and qualitative experiments in Section V show that, the guider is a reliable mechanism of visual attention. Under its guidance, our generator can identify pseudo defects correctly, not as blind as before, which greatly enhances the image restoration ability of our ADRGAN.

C. Discriminator

The generator only realizes low frequency loss and perception loss, and the recovery effect of high frequency pseudo defects is not ideal during image generation. Therefore,

there is an urgent need for the discriminator to distinguish the authenticity of high frequency components in generated images, which will also make the training of the whole GAN more stable with faster convergence speed. Inspired by above-mentioned theories, our discriminator adopts the idea of Res2Net [40]. Fig. 4(a) shows the structure of the discriminator.

This discriminator with multi-scale features at granular level gradually raises the receptive fields of each network layer. In Fig. 4(b), input feature map is divided into four blocks after a convolution and each part is denoted as x_{i} ($i \in \{1,2,3,4\}$), with the form of Convolution-BatchNorm-Relu as corresponding modules [42] represented by $K_i()$. The characteristic subset x_i is added to the output of $K_{i-1}()$ and fed into $K_i()$. Here is the equation:

$$y_i = \begin{cases} x_i & i = 1 \\ K_i(x_i) & i = 2 \\ K_i(x_i + y_{i-1}) & 2 < i \leq s \end{cases} \quad (10)$$

In order to increase the number of channels while reducing the parameters, we omit the form of x_i, which can also be viewed as the reuse of features. Finally, we use $y_i()$ to represent the output of $K_i()$ and concat them for the result of final output.

Res2Net enjoys significant advantages. It can improve multi-scale feature extraction without increasing the computational cost, make the texture of the generator recover image more finely and the trained model focus more on details. Its final objective function is the same as Eq. (1).

IV. RAIN-LIKE LAYER DATASET OF HOT-ROLLED STEEL STRIP SURFACE: STEEL_RAIN

In this paper, the generator in GAN needs to learn and generate images by referring to the training set, while the discriminator needs ground-truth images with clean background to guide the generator to obtain more realistic images. All of these are inseparable from an excellent image dataset, which has a direct link with the quality of the final generated images.

However, the image acquisition at the scene of hot-rolled steel strip suffers with cooling water dispersion, mechanical vibration and high temperature, etc. This terrible working condition requires strictly on data collection, resulting in high acquisition costs. Moreover, due to the rapid production rhythm, it is impossible to get a strict pair of training images with and without waterdrops for a certain surface position of steel strip at the same time. This challenging has also been claimed in our previous paper [18]. That is why we need to generate rain or streak mask to synthesize the other half images with raindrop pollution, other than directly taking rain-drop-polluted images from the raw captured images.

All the 1,455 pairs of clean images with a resolution of 1000 ×1000 pixel are cropped from the raw images (4096×1024 pixel) captured by line-scan cameras. And the actual size of the image samples is 500×500 mm because the pixel physical resolution is 0.5×0.5mm per pixel. For obtaining absolutely clean raw surface images, we need to select the operation condition very carefully and patiently.
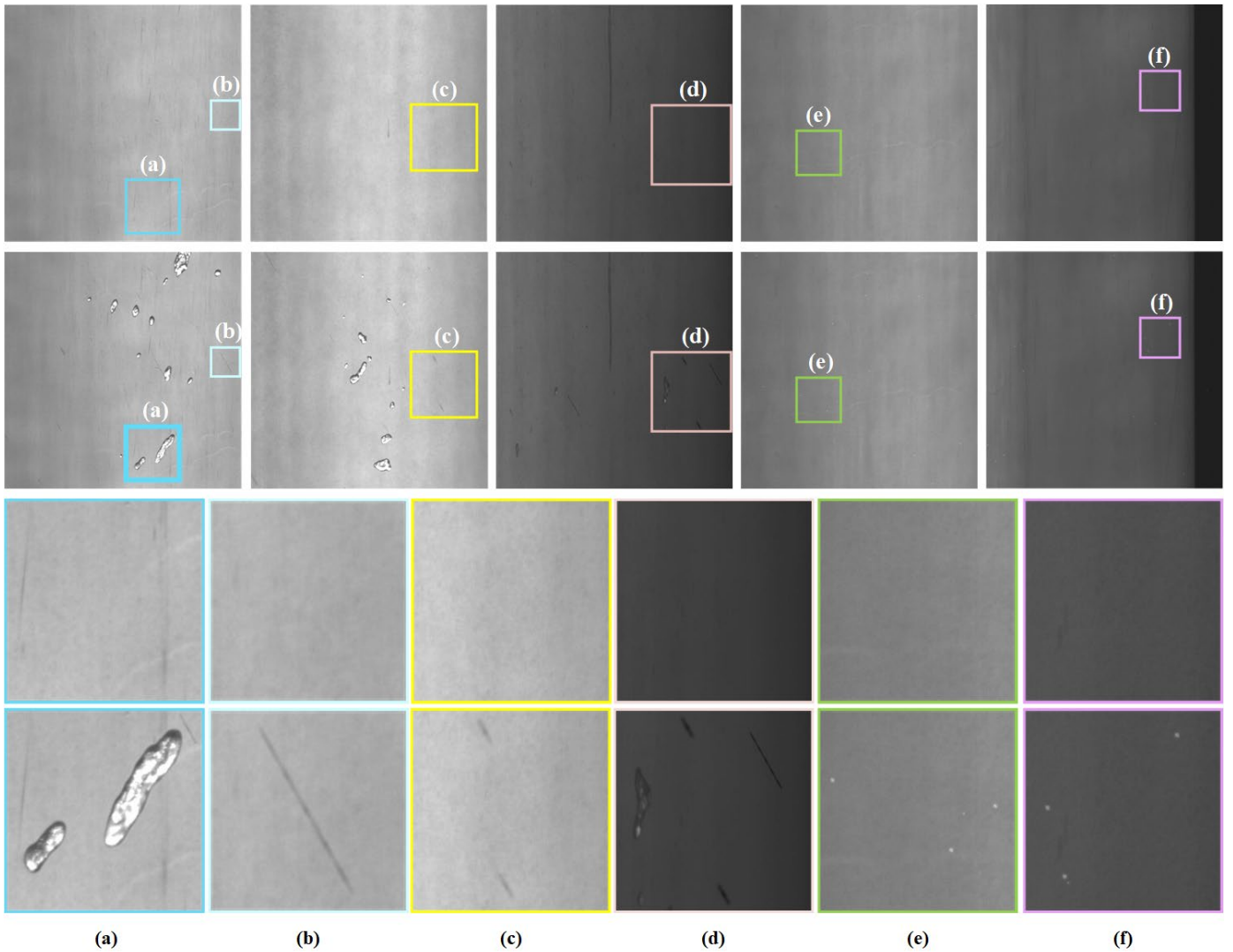
Fig.5 Samples in our high-resolution dataset for the rain-like layer removal in the field of automatic steel surface inspection. (a) are the images with scattered waterdrops; (b)~(d) are the images with splashing water streaks of diverse angles and sizes, while (e) and (f) are the images with tiny white droplets.

TABLE I THE NUMBER OF IMAGES FOR EACH PSEUDO DEFECT IN OUR DATASET

Steel_Rain	Rain-like pseudo defect			
	Scattered waterdrops	Splashing water streaks	Tiny white droplets	Hybrid
Training set (1305 pairs)	674	220	280	131
Test set (150 pairs)	56	24	34	36

The other half (i.e., 1,455 images with raindrop pollution) are artificially made images by pasting different masks of waterdrops and water streaks with disparate angles and sizes on the clean images. The main types of pseudo defects are scattered waterdrops, splashing water streaks and tiny white droplets:

(1) **Scattered waterdrops:** for the improvement of the generalization of the model, the real waterdrops extracted from original images are mixed with synthetic raindrops, and then pasted them into clean background images.

(2) **Splashing water streaks:** the rain-like water streaks with 4 kinds of incline angles and 6 kinds of aspect ratios were made by gaussian noise, and they were superimposed with the background map in proper proportion.

(3) **Tiny white droplets:** after a careful study of the pixel-wise presentation of tiny white droplets caused by the high-speed rolling of the steel strip in the original images, it was simulated to produce pseudo defects of different sizes.

We call this fresh dataset as Steel_Rain, Fig. 5 shows some samples of the dataset. The images in the first row are clean background images while the other row are corresponding images with different types of pseudo defects. TABLE I shows the proportion of pseudo defects for each type in the Steel_Rain dataset, which strictly follows their respective probabilities of occurrence in the actual production line. After testing with real rain-like layer images in those experiments from Section V.E, it is proved that the dataset is suitable for the training requirements. In addition, we use real-world images to prove the practicability of our ADRGAN.

V. EXPERIMENTS

A. Evaluation Criteria and methodology

Quantitative evaluation: The peak signal-to-noise ratio (PSNR) [43], structural similarity index (SSIM) [41], mean-square error (MSE) and learned perceptual image patch

Fig.6 Failure cases of our ADRGAN on *six* datasets. (a) and (b) are failure cases of removing inadequately, (c), (d) and (e) are failure cases of losing details.

Fig.7 The images generated by the models proposed in development of ADRGAN. From left to right: Real rain-like image (input image), DuRN-DS, DuRN-P, DuRN-S-P, DuRN-PSP, DuRN-MaskPSP and ADRGAN.

similarity (LPIPS) [50] are adopted for quantitative analysis of the network performance in the removal of rain-like layer on the steel surface.The image quality becomes better as the PSNR gets larger. SSIM considers the strong correlation between natural image pixels, which can make up for the deficiency that PSNR cannot measure the similarity of image structure. MSE and LPIPS are used to measure the difference between the generated image and the ground truth image, and LPIPS is more consistent with human perception. In addition,

we introduced runtime, model size and floating-point operations (FLOPs) to assess the complexity of models.

Qualitative evaluation: It is the subjective qualitative evaluation of the image by human observer. According to their own knowledge and understanding, human beings evaluate the degree of image restoration and distinguish whether a cleaner image can be obtained after removing the rain layer. For space saving, we only gave analysis of failure cases as the qualitative evaluation on the public datasets, and

TABLE II COMPARISON TEST WITH STATE-OF-THE-ART ALGORITHMS ON SIX PUBLIC DATASETS

Method		DAE	Attentive GAN	DID-MDN	JORDER-E	RESCAN	Pix2Pix	PReNet	IADN	DuRN-S-P	PReGAN	DRD-Net	ADRGAN
Rain 100L	PSNR	31.80	33.69	33.64	36.61	37.52	36.76	36.49	37.44	37.63	37.65	**37.89**	37.66
	SSIM	0.911	0.947	0.953	0.952	0.953	0.957	0.961	0.961	0.973	0.962	0.965	**0.977**
	MSE	42.96	27.80	28.12	14.19	11.51	13.71	14.59	11.72	11.22	11.17	**10.57**	11.15
	LPIPS	0.047	0.031	0.033	0.027	0.024	0.026	0.028	0.024	0.022	0.023	**0.019**	0.021
Rain 100H	PSNR	26.32	27.03	27.44	27.12	28.69	29.16	29.13	29.81	29.60	30.02	29.97	**30.85**
	SSIM	0.783	0.810	0.822	0.873	0.840	0.893	0.892	0.899	0.893	0.896	0.890	**0.912**
	MSE	151.73	128.85	117.24	126.21	87.92	78.90	79.45	67.93	71.30	64.73	65.48	**53.47**
	LPIPS	0.107	0.089	0.082	0.084	0.078	0.077	0.077	0.073	0.074	0.069	0.072	**0.066**
Rain 200H	PSNR	24.19	25.71	25.95	26.37	27.81	28.83	28.84	28.68	28.22	28.31	**28.89**	28.75
	SSIM	0.778	0.802	0.804	0.849	0.834	0.887	0.891	0.893	0.887	0.891	0.892	**0.895**
	MSE	247.79	174.61	165.23	150.00	107.67	85.13	84.93	88.12	97.97	95.96	**83.96**	86.71
	LPIPS	0.144	0.128	0.125	0.121	0.116	0.104	0.103	0.104	0.108	0.105	0.102	**0.093**
Rain 12	PSNR	27.96	29.87	30.07	34.99	35.53	35.45	35.38	35.59	35.63	35.73	**35.90**	35.76
	SSIM	0.841	0.889	0.895	0.924	0.922	0.919	0.921	0.936	0.944	0.930	0.941	**0.961**
	MSE	104.01	67.00	63.99	20.61	18.20	18.54	18.84	17.95	17.79	17.38	**16.71**	17.26
	LPIPS	0.110	0.072	0.069	0.047	0.040	0.042	0.041	0.039	0.038	0.038	**0.035**	0.036
DDN	PSNR	28.72	30.14	30.34	30.59	33.18	33.18	33.19	33.36	33.07	33.14	**33.50**	33.16
	SSIM	0.835	0.846	0.871	0.894	0.916	0.909	0.915	0.922	0.916	0.911	0.932	**0.935**
	MSE	87.31	62.96	60.13	56.76	31.27	31.27	31.19	30.00	32.07	31.56	**29.05**	31.41
	LPIPS	0.083	0.068	0.064	0.062	0.048	0.047	0.046	0.044	0.048	0.047	0.043	**0.041**
DID-MDN	PSNR	28.57	29.08	29.33	31.41	33.24	33.96	34.42	34.47	34.50	34.38	34.51	**34.63**
	SSIM	0.841	0.837	0.855	0.970	0.891	0.887	0.891	0.919	0.922	0.914	0.930	**0.931**
	MSE	90.38	80.37	75.87	47.00	30.84	26.13	23.50	23.23	23.07	23.72	23.02	**22.39**
	LPIPS	0.114	0.101	0.098	0.081	0.074	0.067	0.063	0.061	0.060	0.062	**0.059**	0.059

TABLE III ABLATION STUDY

Structure	Method	Metric		
		Time (s/img)	PSNR	SSIM
6DS	DuRN-DS	0.0960	47.5667	0.9778
6P	DuRN-P [35]	0.0878	48.1995	0.9864
(3S) + (6P)	DuRN-S-P [35]	0.0921	51.4286	0.9956
3(P + DS + P)	DuRN-PSP	0.0889	52.2203	0.9958
DuRN-PSP using the guider	DuRN-MaskPSP	**0.0794**	54.8607	0.9965
DuRN-MaskPSP using the GAN	ADRGAN	0.0826	**56.8627**	**0.9980**

more on the Steel_Rain dataset involves.

Real-world verification: The removal of rain-like layer on steel surface is dedicated to the reduce the false alarms in the final stage of defect detection triggered by rain-like pseudo defects. The generated images can be considered as waterless if the false detections are less enough and within the acceptable range of AVI instruments. *Four* simple and typical defect detection algorithms were applied on the steel surface images before and after rain-like layer removal to verify the benefits brought by the proposed methodologies.

B. Comparison SOTA methods and experimental settings

To verify the generalization of the method, we include *eleven* state-of-the-art (SOTA) algorithms for the comparative experiments, which are the denoising auto-encoder (DAE) [51], Attentive GAN [23], Pix2Pix [36], PReNet [38], IADN [31], DuRN-S-P [35], DID-MDN [45], JORDER-E [29], RESCAN [22], DRD-Net [28] and PReGAN [18]. And *six* public datasets in the field of rain removal and *one* fresh dataset in the actual hot-rolling line are selected as test

benches, which are Rain100L, Rain100H, Rain200H [29], [46], DDN [47], [48], DIDMDN [45], Rain12 [22], and the Steel_Rain set up in the Section IV. To be specific:

- **Rain100L & Rain100H & Rain200H**. We choose 2,400 pairs of synthetic clean images for training (RainTrainH and RainTrainL) and 400 images for testing (Rain100L, Rain100H, and Rain200H) to evaluate the performance under the conditions with extreme rain density and rain streaks intensity.

- **DDN**. We utilize totally 12,600 and 1,400pairs of images for training and testing, respectively. These images are synthesized by combining rain streaks with different directions and shapes on the same background.

- **DIDMDN**. We tend to verify the generalization of the algorithms by selecting 13,200 composite images with heavy, middle, and light rain density.

- **Rain12**. We also select the challenging dataset containing 12 synthetic images with extreme rain density for testing.

All experiments are carried out on 24GB Nvidia RTX3090 TURBO GPU, 3.7GHz Intel Xeon W-2255 CPU and 64GB RAM.

C. Experiments on public datasets

Quantitative Evaluation: The quantitative test results of the *twelve* competitors on *six* public datasets are given in the TABLE II. In most cases, our ADRGAN possesses the competitive PSNR to those of others. Although the PSNR of ADRGAN is slightly lower than that of DRD-Net sometimes, it yields the highest SSIM in any cases. Moreover, on the Rain100H, Rain200H and DID-MDN datasets, ADRGAN obtained the best scores on both MSE and LPIPS metrics, indicating that our proposed method showed great superiority under the condition of heavy rain. These metrics proves that the proposed algorithm produces the images closest to the
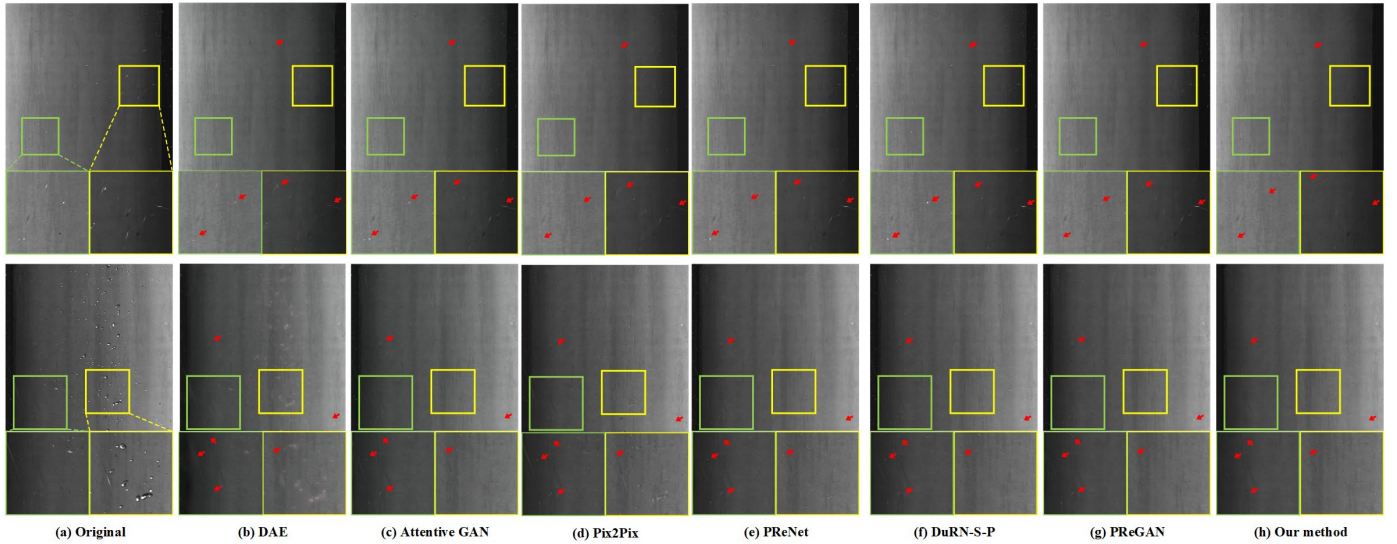
Fig.8 Rain-like layer removal effects for different methods. From left to right: Real images with rain-like layer(input images), DAE [51], Attentive GAN [23], IADN [31], PReNet [38], Pix2Pix [36], DuRN-S-P [35], PReGAN [18] and our method. Nearly all rain-like pseudo defects are removed by our method despite the diversity of their sizes and shapes.

TABLE IV COMPARISON TEST ON THE STEEL RAIN DATASET

Method	Metric							Reference
	PSNR	SSIM	MSE	LPIPS	Time (s/img)	Size (MB)	FLOPs (G)	
DAE [51]	38.9035	0.8258	8.370	0.033	0.0840	**131.60**	**12.60**	JMLR 2010
Attentive GAN [23]	40.8463	0.8554	5.351	0.031	0.0975	572.07	52.79	CVPR 2018
DID-MDN [45]	43.6727	0.9470	2.791	0.027	0.3159	307.88	29.49	CVPR 2018
JORDER-E [29]	45.9128	0.9653	1.666	0.023	0.4795	192.42	18.43	TPAMI 2019
RESCAN [22]	47.2806	0.9704	1.216	0.021	0.5536	267.94	26.66	ECCV 2018
Pix2Pix [36]	47.6178	0.9551	1.125	0.020	0.1251	366.91	35.14	CVPR 2016
PReNet [38]	47.6311	0.9772	1.122	0.020	0.1774	376.12	36.02	CVPR 2019
IADN [31]	49.1751	0.9804	0.786	0.018	0.1332	379.63	36.36	TCSVT 2021
DuRN-S-P [35]	51.4286	0.9956	0.468	0.014	0.0921	355.02	33.20	CVPR 2019
PReGAN [18]	52.7235	0.9714	0.347	0.011	0.1663	413.29	39.08	TIM 2021
DRD-Net [28]	54.2948	0.9877	0.242	0.009	0.3754	447.35	42.84	CVPR 2020
ADRGAN	**56.8627**	**0.9980**	**0.134**	**0.006**	**0.0826**	385.03	36.83	Ours

ground truth. The qualitative results with vivid images can be reached via the open codes, where the rain removal effects are available especially on heavy rains. The possible reason why the DRD-Net wins higher PSNR in some cases is that it introduces two parallel sub-networks with a comprehensive loss function which synergize to de-rain and recover the lost details caused by de-raining.

Analysis of Failure cases: As shown in Fig. 6, some difficult images still posed challenged to our method. We can see from Fig. 6(a)~(b) that our method is lack of integrity in restoring the high-density area of partial images containing white stripes. Fig. 6(c)~(e) indicate that some edge details without obvious color difference will be blurred. The main reason for those cases is the lack of training images with backgrounds similar in shape and color to raindrops. We will work on these problems in the future.

D.   Experiments on fresh datasets

We randomly select 1,305 pairs of images from our newly-built Steel_Rain dataset as the network training samples, and the batch size is set to 24. The remaining 150 pairs are used as the testing samples. Two evaluation aspects are involved:

Ablation study: As shown in TABLE III, we evaluate the *six* steps gradually to learn about the performance of each part of the ADRGAN by removing them progressively. From the values top to bottom, the PSNR is gradually increased from 47.5667 of the most basic DuRN-DS to 56.8627 of the most advanced ADRGAN, this improvement on PSNR proves the proposed method has really workable on image enhancement by tackling the pseudo defect removal. In addition, the DuRN-MaskPSP enhances the performance of raindrop removal and reduces running times through a compatible structure with both channel-wise and spatial-wise attention. Although ADRGAN adds few computational overheads, the pixels covered by the waterdrops are effectively restored, which has been verified from the final striking SSIM of 0.998.

To be specific, as shown in Fig. 7, benefited from the proposed attentional block of double SE, no matter the large or small water droplet will attract the attention by this sub-block. Hence, the rain removal effect in Fig. 7(b) is more satisfactory than that in Fig. 7(c), which indicates the DuRN-DS performs better than DuRN-P on recognition ability of pseudo defects.

TABLE V THE AVERAGE NUMBER OF FALSE ALARMS DETECTED IN 60 REAL RAIN-LIKE LAYER IMAGES BEFORE AND AFTER IMAGE RESTORATION

Method	Original image	Attentive GAN	Pix2Pix	PReNet	IADN	DuRN-S-P	PReGAN	Ours	our method / real raindrop image
Thresholding	43	29	22	18	10	12	10	5	**0.13**
LBP	39	31	20	21	19	7	17	6	0.14
Canny	33	27	24	21	22	18	13	9	0.27
Gabor filtering	46	37	19	24	17	14	21	11	0.24

However, its image restoration ability is relatively weak, many repairing traces (marked with red arrows) could be found if you zoom the green image patch in the upper figure or yellow image patch in the down lower figure in Fig. 7(b), that is why the DuRN-DS get the lowest scores on both PSNR and SSIM in the TABLE III. In contrast, such kind of repairing traces is absent in the Fig. 7(c) when the DuRN-P arrives, which shows its better detail maintaining ability. This phenomenon is highly consistent with the concept proposed in DuRN-S-P [35] that combining the blocks of DuRN-DS and DuRN-P is beneficial to obtain better object identification ability and prevent well the edge details being destroyed. Inheriting well this concept, we propose the three-cycle search-repair structure (refer to Fig. 3) to handle the occasional misidentification resulting from aggressive searching stratagem in the disposable structure of DuRN-PSP. The scores of the initial structure of DuRN-PSP are slightly higher than those of DuRN-S-P in TABLE III. This result can also be reachable in the Fig. 7(d) and Fig. 7(e). All of the preliminary results prove that the dual-residual-based periodic structure with double SE configuration can track and restore the rain-like pseudo defects effectively.

Further, the guider insists the generator to avoid over de-rain problem by utilizing the spatial features of the rain-like layer, obtaining higher scores on PSNR and SSIM one more time. Then we introduce the framework of GAN to remove residual high-frequency pseudo defects in Fig. 7(f), and ultimately obtain a sufficiently clean image (see Fig. 7(g)).

Comparison test: As shown in Table IV, our ADRGAN beats all the *eleven* competitors on *four* criteria. Notably, the SSIM is pushed to nearly equal to 1, which means images polluted by the rain-like pseudo defects have been restored to the previous ground truth incredibly. Moreover, MSE and LPIPS measure the difference between clean and generated images from the pixel level and perceptual level respectively, and our method both achieved the minimum values. For model complexity, our model takes 0.0826s average interference time of per image, 385.03 MB model size and 36.83G FLOPs. Compared to the DAE, RESCAN, DID-MDN and JORDER-E, our ADRGAN pays more attention to the image restoration accuracy, but at the expense of increasing additional parameters costs. Our method also achieves better score in average interference time. These comparative results proved that the multiple cycles of searching and restoring are really beneficial to the dynamical removal of rain-like pseudo defects and the detail preservation of actual defects, which have pretty application potential to promote the performance of AVI instrument for surface defect inspection, especially under harsh production environment.

Qualitative Evaluation: For space saving, we only take DAE [51], Attentive GAN [23], Pix2Pix [36], PReNet [38], DuRN-S-P [35] and PReGAN [18]. from TABLE III for the qualitative evaluation in Fig. 8. To be specific, Fig. 8(a) are the real images polluted by the so-called rain-like layer, As shown in Fig. 8(b), DAE can wipe out main water streaks, but the recovery of image texture is extremely limited, which probably due to its lack of attention mechanisms and scalability for high-dimensional features, resulting in the loss of many valid details. Fig. 8(c) shows the images generated by Attentive GAN, which can barely see large waterdrops, but it fails to suppress the water streaks and tiny white droplets, let alone the global distortion. As can be seen from Fig. 8(d), Pix2Pix cannot completely remove waterdrops, although the generated images have no apparent color difference. In addition, it might remove some real defects erroneously such as tiny roll marks, which result in undetected problem at the subsequent stage of defect inspection. In Fig. 8(e), PReNet obtains better global attention. However, there are still some tiny waterdrops left in the dark image areas. As discussed in TABLE IV, neither Fig. 8(g) nor Fig. 8(h) generated by DuRN-S-P and PReGAN respectively can handle tiny white droplets and water streaks well, increasing false detection rate to AVI instrument. Notably, all of these rain-like pseudo defects have been eliminated nearly completely by our ADRGAN in Fig. 8(h). These positive results prove that our algorithm indeed enable the enhanced images more identical to the ground truth. Furthermore, it also indicates that our fresh Steel_Rain dataset can simulate the morphology and distribution of pseudo defects effectively, so as to achieve the purpose of removing rain-like layer of images captured from real industrial production line.

E. Industrial verification

The rain-like layer removal belongs to the pre-processing step of hot-rolled steel strip images and its ultimate purpose is to promote the accuracy and efficiency of defect identification. We try to verify the practical industrial value of our method by comparing the false detection rate of *four* simple and typical defect detection algorithms when assisted with the rain-like layer removal or not. They are simple thresholding algorithm based on block variance, textual analysis algorithm based on local binary patterns (LBP), edge detection algorithm based on Canny operator, and filtering algorithm based on Gabor transform. For fair comparison, the dynamic homogenizing compensation (DHC) [44] and mean shift filtering were performed on the original images before using the above *four* detection algorithms, to refrain the illumination changes and occasional noises.

We selected 60 original hot-rolled steel images randomly from the Steel_Rain datasets, and further obtained quantitative and qualitative evaluation. First from Table V, our proposed method has efficiently decreased the false alarms for all four involved detection algorithms. Notably, the average number of

De-Raining using ADRGAN

DHC + mean shift filtering

Detection method based on thresholding

DHC + mean shift filtering

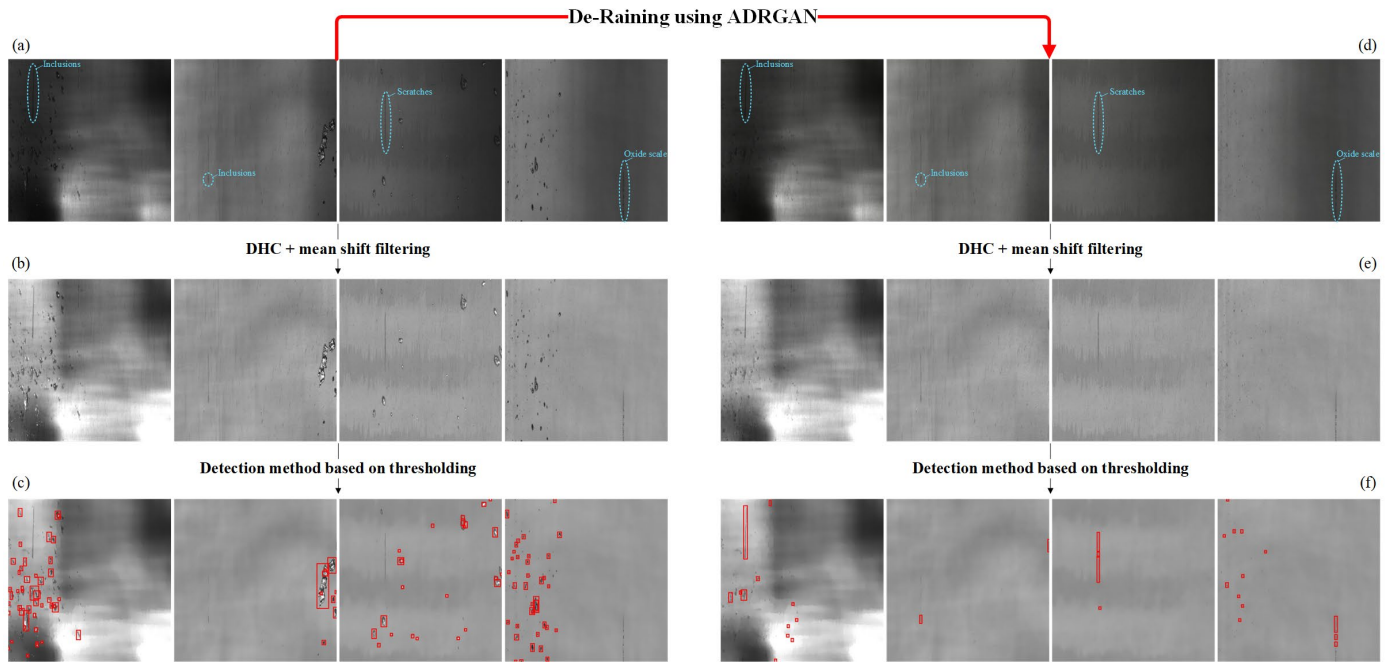**Detection method based on thresholding**

Fig.9 The comparative results of original images restored by ADRGAN using a defect detection method based on variance threshold. (a) and (d) are the original images and the rain-like layer removal images respectively; (b) and (e) are the corresponding enhanced images; (c) and (f) are the defect detection results.

(a) Original **(b) Attentive GAN** **(c) Pix2Pix** **(d) PReNet** **(e) IADN** **(f) DuRN-S-P** **(g) PReGAN** **(h) Our method**
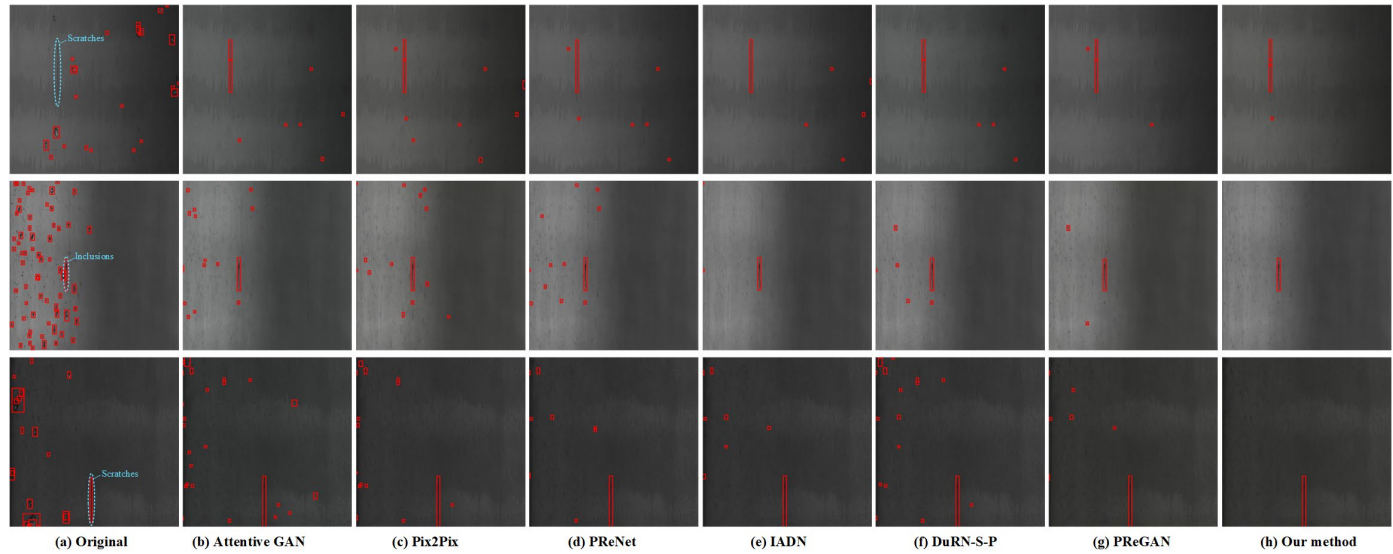
Fig.10 The comparative results of images restored by ADRGAN and its competitors using a simple defect detection method, variance thresholding.

false alarms using our ADRGAN were reduced by at least 70%, even more significantly, cutting by more than 85% on both LBP- and thresholding-based algorithms, compared with those without any rain removal measure, which is inseparable from our efforts in the removal of splashing water streaks and tiny white droplets. The least false alarms obtained in Table V benefits from the best quantitative figures in Table IV.

To be honest, we believe that an excellent target recognition method can complete the defect inspection task to challenging images in a harsh industrial environment. However, if we have an excellent pre-processing algorithm to tackle the dirty images to be sufficiently clean, why not adopt such a kind of roadmap as an optional, that is, applying extremely simple inspection algorithm on sufficiently clean surface images? All in all, stability is everything for industrial AVI equipment. It is undeniable that rain-like layer removal can indeed reduce the false detection rate of AVI instruments, providing a reasonable and efficient choice for surface defect inspection. Certainly, if conditions permit, the AVI instrument is highly recommended to be installed at the upstream site of the laminar colling devices, so as to avoid the false alarms triggered by the rain-like pseudo defects thoroughly.

For more insights, the simplest variance-based thresholding method is selected for some visual investigation. Fig. 9 takes *four* randomly chosen samples for elaboration. Particularly, *three* kinds of defects are involved, oxide scales, inclusions and scratches. As can be seen from the left half side, a large number of rain-like pseudo defects are incorrectly identified as defects, triggering false alarms on AVI instrument, which proves the necessity of rain-like layer removal, one more time.

This article has been accepted for publication in IEEE Transactions on Instrumentation and Measurement. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIM.2023.3265761

>Paper ID: TIM-22-06341R   Title: Rain-like Layer Removal from Hot-Rolled Steel Strip Based on ADRGAN<   13

As for the right half figure, after using ADRGAN, the false alarms are apparently reduced, and the surface defect is easier to be inspected. It is important to mention that, due to the limitations of the detection methods, a few numbers of false alarms also exist in the experimental results after using our ADRGAN (see Fig. 9(f)). This issue can be easily solved by improving or replacing the current simple detection algorithms. Few-shot learning [49], which might be a feasible road map to inspect the steel surface defect more robustly under the challenging environment on the hot-rolling line, which will be our nearly future work, making attempts to inspect defects more accurately and reliably among even complicated interferences.

Again, Fig. 10 continues to compare the proposed ADRGAN with its *six* opponents—Attentative GAN, Pix2Pix, PReNet, IADN, DuRN-S-P and PReGAN. In Fig. 10(a), the original image samples suffered with rain-like pseudo defects, triggers a huge amount of false alarms. In Fig. 10(b), Attentive GAN tackles the above-mention challenge to the maximum, but water streaks and tiny white droplets are stubbornly survived. From Fig. 10(c)~(g), all methods are difficult to overcome the challenges brought by splashing water streaks and tiny white droplets, simultaneously, and our ADRGAN obtains the least false alarms due to its robustest global attention and image restoration ability in Fig. 10(h). The above experiments show that the rain-like layer removal concept possesses believable stability, which can be applied to steel surface AVI instrument.

VI. CONCLUSION

Aiming at solving the serious false-alarm problem caused by the hybrid rain-like pseudo defects, in this paper, a global search algorithm called attentive dual residual generative adversarial network (ADRGAN) is developed to remove the "rain-like layer" in the steel surface image under the premise of preserving edge and texture details. In the extensive comparison experiments with *eleven* state-of-the-art rain removal methods on *six* public dataset and our newly-opened industrial dataset, our method has reliably maintained the most excellent capabilities of rain-like layer identification and detail restoration from both aspects of quantitative and qualitative evaluation. The verification tests on *four* simple and typicle defect inspection methods prove that steel surface images enhanced by ADRGAN suffered with the least false alarms, which proves that the rain-like layer removal method in this paper provides a feasible optional for addressing the remaining problem that is not yet solved due to the limitations of practical scenarios.

Our previously-opened high-resolution dataset [18] for the rain-like layer removal in automatic steel surface inspection are expanded in this paper. In addition to cooling water droplets, we also consider splashing water streaks and tiny water droplets. This continuously opened datasets are expected to stimulate more petential rain-like layer removal algorithms for surface inspection of industrial sheet meterials, especially in the real-world scenario.

REFERENCES

[1]   Y. He, K. Song, Q. Meng, and Y. Yan, "An end-to-end steel surface defect detection approach via fusing multiple hierarchical features," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 4, pp. 1493-1504, Apr. 2020.

[2]   H. Dong, K. Song, Y. He, J. Xu, Y. Yan, and Q. Meng, "PGA-Net: pyramid feature fusion and global context attention network for automated surface defect detection," *IEEE Trans. Ind. Informat.*, vol. 16, no. 12, pp. 7448-7458, Dec. 2020.

[3]   S. Ghorai, A. Mukherjee, M. Gangadaran, and P. K. Dutta, "Automatic defect detection on hot-rolled flat steel products," *IEEE Trans. Instrum. Meas.*, vol. 62, no. 3, pp. 612–621, Mar. 2013.

[4]   Q. Luo, Y. Sun, P. Li, O. Simpson, L. Tian, and Y. He, "Generalized completed local binary patterns for time-efficient steel surface defect classification," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 3, pp. 667–679, Mar. 2019.

[5]   Q. Luo, X. Fang, L. Liu, C. Yang, et al., "Automated visual defect detection for flat steel surface: A survey," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 3, pp. 626-644, Mar. 2020.

[6]   G. Song, K. Song and Y. Yan, "Saliency detection for strip steel surface defects using multiple constraints and improved texture features," *Opt. Lasers Eng.*, vol. 128, May 2020.

[7]   G. Song, K. Song, and Y. Yan, "EDRNet: Encoder–decoder residual network for salient object detection of strip steel surface defects," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 12, pp. 9709-9719, Dec. 2020.

[8]   J. Liu, K. Song, M. Feng, Y. Yan, Z. Tu, and L. Zhu, "Semi-supervised anomaly detection with dual prototypes autoencoder for industrial surface inspection," *Opt. Laser. Eng.*, vol. 136:106324, Jan. 2021.

[9]   D. Zhang, K. Song, Q. Wang, Y. He, X. Wen, and Y. Yan, "Two deep learning networks for rail surface defect inspection of limited samples with line-level label," *IEEE Trans Ind. Informat.*, vol. 17, no. 10, pp. 6731-6741, Oct. 2021.

[10]   X. Zhou, H. Fang, Z. Liu, et al. "Dense attention-guided cascaded network for salient object detection of strip steel surface defects," *IEEE Trans. Instrum. Meas.*, vol. 71, Dec. 2021.

[11]   Q. Luo, X. Fang, J. Su, J. Zhou, B. Zhou, C. Yang, et al., "Automated Visual Defect Classification for Flat Steel Surface: A Survey," *IEEE Trans. Instrum. Meas.*, vol.69, no. 12, pp.9329-9349, Oct. 2020.

[12]   Q. Luo, X. Fang, Y. Sun, et al., "Surface Defect Classification for Hot-Rolled Steel Strips by Selectively Dominant Local Binary Patterns," *IEEE Access*, vol. 7, pp. 23488-23499, Feb. 2019.

[13]   Y. He, K. Song, H. Dong, and Y. Yan, "Semi-supervised defect classification of steel surface based on multi-training and generative adversarial network," *Opt. Laser. Eng.*, vol. 122, pp. 294-302, Nov. 2019.

[14]   F. Huang, J. Qi, H. Lu, L. Zhang and X. Ruan, "Salient object detection via multiple instance learning," *IEEE Trans. Image Process.*, vol. 26, no. 4, pp. 1911-1922, Apr. 2017.

[15]   W. Zhu, S. Liang, Y. Wei and J. Sun, "Saliency optimization from robust background detection," in Proc. *IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 2814-2821, Jun. 2014.

[16]   H. Song, Z. Liu, H. Du, G. Sun, et al., "Depth-aware salient object detection and segmentation via multiscale discriminative saliency fusion and bootstrap learning," *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4204-4216, Sep. 2017.

[17]   L. Zhou, Z. Yang, Z. Zhou and D. Hu, "Salient region detection

using diffusion process on a two-layer sparse graph," *IEEE Trans. Image Process.*, vol. 26, no. 12, pp. 5882-5894, Dec. 2017.

[18] Q. Luo, K. Liu, J. Su, C. Yang, W. Gui, L. Liu, et al., "Waterdrop removal from hot-rolled steel strip surfaces based on progressive recurrent generative adversarial networks," *IEEE Trans. Instrum. Meas.*, vol. 70: 5017011, pp. 1-11, Art. 2021.

[19] Ahn, N., Jo, S. Y., Kang, S. J., "EAGNet: elementwise attentive gating network-based single image de-raining with rain simplification," *IEEE Trans. Circuits Syst. Video Technol.*, 2021.

[20] L. W. Kang, C. W. Lin, and Y. H. Fu., "Automatic single-image-based rain streaks removal via image decomposition," *IEEE Trans. Image Process.*, 21(4):1742-1755, Apr. 2012.

[21] Y. Luo, Y. Xu, and H. Ji., "Removing rain from a single image via discriminative sparse coding," In Proc. *IEEE Int'l Conf. Computer Vision*, pages 3397–3405, 2015.

[22] X. Li, J. Wu, Z. Lin, H. Liu, and H. Zha, "Recurrent squeeze-and-excitation context aggregation net for single image deraining," in *Proc. Eur. Conf. Comput. Vis*, pp.262-277, Springer, 2018.

[23] R. Qian, R. T. Tan, W. Yang, J. Su, J. Liu, "Attentive generative adversarial network for raindrop removal from a single image," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 2482-2491, Jun. 2018.

[24] W. Luo, J. Lai, X. Xie, "Weakly supervised learning for raindrop removal on a single image," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 5, pp. 1673-1683, May 2021.

[25] J. Hu, L. Shen, S. Albanie, G. Sun, E. Wu, "Squeeze-and-excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2018.

[26] Q. Wu, L. Wang, et al., "Subjective and objective de-raining quality assessment towards authentic rain image," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 11, pp. 3883-3897, Nov. 2020.

[27] D. Chen, C. Chen, L. Kang. "Visual depth guided color image rain streaks removal using sparse coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 8, pp. 1430-1455, Aug. 2014.

[28] S. Deng, M. Wei, J. Wang, *et al.*, "Detail-recovery image deraining via context aggregation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 14560-14569, Jun. 2020.

[29] W. Yang, R. T. Tan, J. Feng, Z. Guo, S. Yan, and J. Liu, "Joint rain detection and removal from a single image with contextualized deep networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 6, pp. 1377–1393, 2019.

[30] Y. Du, J. Xu, X. Zhen, M. M. Cheng, and L. Shao, "Conditional variational image deraining," *IEEE Trans. Image Process.*, vol. 29, pp. 6288–6301, 2020.

[31] K. Jiang, Z. Wang, P. Yi, et al., "Decomposition makes better rain removal: an improved attention-guided deraining network," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 10, pp. 3981-3995, Oct. 2021.

[32] M. Roser, J. Kurz, and A. Geiger, "Realistic modeling of water droplets for monocular adherent raindrop recognition using bezier curves," *LNCS*, vol. 6469, pp. 235-244, 2011.

[33] D. Eigen, D. Krishnan, and R. Fergus, "Restoring an image taken through a window covered with dirt or rain," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, pp. 633–640, Mar. 2014.

[34] M. R. Kanthan and S. N. Sujatha, "Rain drop detection and removal using k-means clustering," in *Proc. IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, pp. 811-815, 2015.

[35] X. Liu, M. Suganuma, Z. Sun, and T. Okatani, "Dual residual networks leveraging the potential of paired operations for image restoration," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 7000-7009, Jun. 2019.

[36] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Nov. 2016.

[37] I. Goodfellow et al., "Generative adversarial networks," *Advance in NIPS*, vol. 27, pp. 2672–2680, 2014.

[38] D. Ren, W. Zuo, Q. Hu, P. Zhu, and D. Meng, "Progressive image deraining networks: a better and simpler baseline," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 3932-3941, 2019.

[39] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proc. Int. Conf. Mach. Learn.*, pp. 807-814, 2010.

[40] S. Gao, M. Cheng, K. Zhao, et al., "Res2Net: a new multi-scale backbone architecture," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 2, pp. 652–662, Feb. 2021.

[41] Z. WANG, A. C. BOVIK, H. R. SHEIKH, et al., "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600-612, Apr. 2004.

[42] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," *Mach. Learn.*, vol. 37, pp. 448-456, Mar. 2015.

[43] Q. Huynh-Thu and M. Ghanbari, "Scope of validity of PSNR in image/video quality assessment," *Electron. Lett.*, vol. 44, no 13, pp. 800–U35, Jun. 2008.

[44] Q. Luo and Y. He, "A cost-effective and automatic surface defect inspection system for hot-rolled flat steel," *Robot. CIM-INT Manuf.*, vol. 38, pp. 16-30, Apr. 2016.

[45] H. Zhang and V. M. Patel, "Density-aware single image de-raining using a multi-stream dense network," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 695-704, 2018.

[46] W. Yang, R. T. Tan, J. Feng, J. Liu, Z. Guo, and S. Yan, "Deep joint rain detection and removal from a single image," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 1357–1366, 2017.

[47] X. Fu, J. Huang, X. Ding, Y. Liao, and J. Paisley, "Clearing the skies: A deep network architecture for single-image rain removal," *IEEE Trans. Image Process.*, vol. 26, no. 6, pp. 2944–2956, 2017.

[48] X. Fu, J. Huang, D. Zeng, Y. Huang, X. Ding, and J. Paisley, "Removing rain from single images via a deep detail network," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 1715-1723, 2017.

[49] H.H. Wang, Z.L. Li, et al. "Few-shot steel surface defect detection," *IEEE Trans. Instrum. Meas.*, vol. 71, no. 5003912, Mar. 2022.

[50] Zhang R, Isola P, Efros A. A, et al. "The unreasonable effectiveness of deep features as a perceptual metric". in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jan. 2018.

[51] Vincent P, Larochelle H, et al. "Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion," *Journal of Machine Learning Research*, vol. 11, no 12, pp. 3371-3408, 2010.

QIWU LUO (Senior Member, IEEE) was awarded the B.S. degree in communication engineering from the National University of Defense Technology, Changsha, China, in 2008, and the M.Sc. degree in electronic science and technology and the Ph.D. degree in electrical engineering from Hunan University, Changsha, in 2011 and 2016, respectively.

He served as a Senior Engineer of instrumentation with WASION Group Ltd. Company, Changsha, and the Deputy Technical Director with Hunan RAMON Technology Company, Ltd., Changsha. In 2016, he worked at the School of Electrical Engineering and Automation, Hefei University of Technology, Hefei, China, where he also completed his postdoctoral research on automatic optic inspection (AOI). Since 2019, he has been an Associate Professor with the School of Automation, Central South University, Changsha. His currently researches on computer vision, industrial AOI, machine learning, parallel hardware architecture design, and reconfigurable computing.

HANDONG HE received his B.S. degree in intelligent science and technology from Central South University in June 2021. He is currently pursuing a M.Sc. degree in advance in control science and engineering with the School of Automation, Central South University, Changsha, China, under the guidance of Dr. Luo. His currently researches on image/video processing and computer vision.

KEXIN LIU was conferred the B.S. degree in intelligent science and technology from Central South University in June 2019. She is currently pursuing a M.Sc. degree in advance in control science and engineering with the School of Automation, Central South University, Changsha, China, under the supervision of Dr. Luo. Her current research interests include waterdrop removal and defect detection.

CHUNHUA YANG (Fellow, IEEE) received the M.S. degree in automatic control engineering and the Ph.D. degree in control science and engineering from Central South University, Changsha, China, in 1988 and 2002, respectively.

From 1999 to 2001, she was a Visiting Professor with the University of Leuven, Leuven, Belgium. Since 1999, she has been a Full Professor with the School of Information Science and Engineering, Central South University, Changsha, China. From 2009 to 2010, she was a Senior Visiting Scholar with the University of Western Ontario, London, ON, Canada. She is currently the HoD of the School of Automation, Central South University. Her current research interests include modeling and optimal control of complex industrial processes, and intelligent control systems.

Olli Silvén (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees in electrical and computer engineering from the University of Oulu, Finland, in 1982 and 1988, respectively. Since 1996, he has been a professor of signal processing engineering with the University of Oulu. He has contributed to the development of numerous solutions from real-time 3-D imaging in reverse vending machines to IP blocks for mobile video coding. His research focuses on ultra-energy-efficient-embedded signal processing and machine vision system design.

LI LIU (Senior Member, IEEE) received the Ph.D. degree in information and communication engineering from the National University of Defense Technology (NUDT), Changsha, China, in 2012. She is currently a Professor with the College of System Engineering, National University of Defense Technology. She is also an assistant professor with the Center for Machine Vision and Signal Analysis (CMVS), University of Oulu, Finland. During her PhD study, she spent more than two years as a Visiting Student at the University of Waterloo, Canada, from 2008 to 2010. From 2015 to 2016, she spent ten months visiting the Multimedia Laboratory, The Chinese University of Hong Kong. From 2016 to 2018, she worked as a Senior Researcher at the Machine Vision Group, University of Oulu, Finland. Her papers have currently over 3700 citations in Google Scholar. Her current research interests include computer vision, pattern recognition and machine learning.