

Waterdrop Removal from Hot-Rolled Steel Strip Surfaces based on Progressive Recurrent Generative Adversarial Networks

Qiwu Luo, *Member, IEEE*, Kexin Liu, Jiaojiao Su, Chunhua Yang, *Senior Member, IEEE*, Weihua Gui, *Member, IEEE*, Li Liu, *Senior Member, IEEE*, and Olli Silven, *Senior Member, IEEE*

Abstract—Automated visual inspection (AVI) instrument of surface defects for hot-rolled steel strips is conventionally installed closely before the terminal crimping machine, where the adjacent upstream process is laminar spray cooling. Waterdrops, spreading more or less over the steel strip surface, often trigger false-alarm, which is a quite common problem in AVI. Stimulated by the idea of image rain removal in visual enhancement field, this paper considers the surface waterdrops, pseudo defects in essence, as a conceptual “rain-like layer”. A targeted method, namely progressive recurrent generative adversarial network (PreGAN), is designed for active waterdrop tracking and fine-grained image inpainting. Meanwhile, a steel surface database (2400 raw images with the resolution of 1000×1000) captured from actual hot-rolling line is constructed for the first time for open evaluation of waterdrop removal. The experimental results indicate that images enhanced by the PreGAN perform the most informative and spotless, with 52.2073 peak signal-to-noise ratio (PSNR) and 0.9502 structural similarity index (SSIM), when compared with four prestigious networks. Assisted by the PreGAN, the false alarms are proved to be reduced at least a half during the application tests using four traditional simple detection methods.

Index Terms—Automated visual inspection (AVI), hot-rolled steel strip, waterdrop removal, image deraining, generative adversarial network (GAN).

I. INTRODUCTION

Automated visual inspection (AVI) instrument plays an essential role for hot-rolled steel mills to ensure product quality [1] [2] [3]. Regularly, this kind of AVI instrument is arranged at the end of the hot-rolling line as far as possible, while the adjacent upstream process is the spray cooling. Consequently, great quantities of waterdrops, water clothes and rain lines widely distribute on the steel surface in random

manner, which are captured by the AVI system and erroneously recognized as defects, essentially, pseudo defects [4]. Further, a tremendous amount of false-alarms will be triggered especially when suffered with heterogeneous pseudo defects. What is worse, some actual defects are entirely or partially masked, which will seriously decrease the detection accuracy of AVI instrument.

In essential, building a robust image enhancement method that can precisely track and remove waterdrop masks, and to finalize the image lost area inpainting is the key to solve the aforementioned problem. For surface quality inspection of hot-rolled steel strips, last five years have witnessed a great progress in deep learning methodologies [5] [6] [7] [8]. However, nearly all of them are focused on defect detection [9] [10] [11] [12] [13] or classification [4] [5] [14]. Such an important research issue of waterdrop removal seems to have been ignored by researchers. Fortunately, the topic of rain removal has been extensively studied in the field of film and television enhancement [15] [16] [17] [18]. Notably in [15], Qian *et al.* proposed Attentive GAN and realized outstanding raindrop removal from a single image based on a self-made database, where a typical application scene characteristic can be investigated that the raindrops are near the camera lens but the background scene is far away from the camera lens. In most cases, some raindrops are kind of blurred because cameras always focus on the background in scenery photograph. It should be noted that, for the surface images of hot-rolled steel strips, the imaging distance of waterdrops and steel plate is almost identical, resulting that the waterdrops are clear and the information of corresponding part occluded by the waterdrops are completely lost. As narrated above, those outstanding methods for raindrop removal popularized in film and television enhancement cannot be directly applied on the waterdrop removal for hot-rolled steel strips.

Another challenge is how to set up an image database to serve the task of surface waterdrop removal on the actual high-speed hot-rolling production line. Assuming that the image with waterdrops was taken from one position on the steel strip, it would be almost impossible to obtain a clean image in the same position. Because the hot-rolled steel strip has already left at a high speed. That means, 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 paper makes attempt to develop a targeted algorithm to effectively remove waterdrops in the surface images of hot-rolled steel strip, as well as to explore easy-to-use solution

This work was supported in part by the National Natural Science Foundation of China under Grant 61973323 and Grant 6201101509, in part by the Hunan Provincial Natural Science Foundation of China under Grant 2020JJ4747, and in part by the Innovation and Development Project of Ministry of Industry and Information Technology of the People's Republic of China under Grant TC19084DY. (Corresponding Author: Chunhua Yang, Email: ychh@csu.edu.cn)

Q. Luo, K. Liu, J. Su, C. Yang, and W. Gui are with the School of Automation, Central South University, Changsha 410083, China.

Li Liu is with the College of System Engineering, National University of Defense Technology, Changsha 410073, China, and also with the Center for Machine Vision and Signal Analysis (CMVS), University of Oulu, 90014 Oulu, Finland.

O. Silven is with the Center for Machine Vision and Signal Analysis (CMVS), University of Oulu, 90014 Oulu, Finland.

for constructing image database for surface waterdrop removal. The key contributions of this paper are highlighted as follows.

(1) *We creatively defined the pseudo defect of waterdrops on the surface of hot-rolled steel strips as rain-like layer*, and proposed a novel waterdrop removal method of progressive recurrent generative adversarial network (PReGAN). Feeding with clean images enhanced by the PReGAN, the detection accuracy of AVI instrument could be boosted significantly.

(2) A steel surface image database containing 1200 pairs of images has been constructed, half of which are raw clean images with the resolution of 1000×1000 captured from actual hot-rolling line, corresponding the other half are same background with waterdrops made artificially. *To our best knowledge, this is the first open high-resolution database in the field of waterdrop removal for automated surface defect inspection of hot-rolled steel strips*, which provides solid support for surface quality automatic inspection of industrial planar materials.

The rest of this paper is organized as follows. Section II discusses the related works about waterdrops detection and removal. Section III elaborates the proposed PReGAN in detail. Afterwards, Section IV describes how to obtain the image database of waterdrops on hot-rolled steel strips for training. Further, our experiments are evaluated quantitatively and qualitatively, and the effect of the experimental results on defect detection is assessed in Section V. Finally, Section VI concludes this paper.

II. RELATED WORKS

Some methods have been proposed to detect raindrops. Kurihata *et al.* [19] used principal component analysis (PCA) to learn the template of raindrops to achieve detection, which tried to match the regions in the test image with the regions learned from raindrops. Because the raindrops were transparent and of different shapes, the PCA could not guarantee the adequate modeling of each type of raindrops, so it has insufficient ability to prevent the regions similar to the raindrops from being detected by error. To overcome this problem, Roser and Geiger [20] [21] compared the generated raindrops to patches that might have raindrops, where raindrop shapes are assumed, but the actual shape and size of raindrops always vary, which can't be treated in the same way.

Some popular depth model-based methods are also used. Eigen *et al.* [22] focused on raindrop removal in a single image. Their basic idea was to train the convolutional neural network (CNN) with a pair of degenerated images with raindrops and corresponding raindrop-free images, which only take effects on sparse small raindrops and dust, while has little effect on larger, relatively dense raindrops. Ramensh *et al.* [23] proposed a method based on K-means clustering and median filtering to estimate clean images, but the final output result was fuzzy. Generative adversarial network (GAN) is often used for image generation. Innovatively, Qian *et al.* [15] fused attention mechanism into the traditional GAN to achieve the learning and generation from images stained with raindrops to clean images. Their method not only learns the mapping from the input images to the output, but includes a variety of loss functions. Nevertheless, due to the extremely complex and complete network structure, compared with the simple network structure,

it is not conducive to further expand the network on this basis. A novel style of residual connection named dual residual connection was put forward by Liu *et al.* [24], and 5 image restoration tasks were experimentally evaluated in 9 databases. Among which the removal of raindrops uses the database made and published by Qian *et al.* [15]. It uses DuRN-S-P network which has dual residual connection to get better results in terms of indicators than Attentive GAN. Although Pix2Pix [25] is not specifically designed to deal with waterdrops, it provides a condition GAN (CGAN) that can transform one image into another, which is a universal mapping and can also be used for the removal of waterdrops.

Previously, several outstanding publications have been reported on raindrop removal, but these methods are hard to be directly applied on the waterdrop removal of hot-rolled steel strip, due to the formation mechanism and morphology of the waterdrops to be removed are completely different. In general, there is still a blank in the removal of waterdrops on the surface of hot-rolled steel strip and it is becoming increasingly difficult to ignore effect of waterdrops in practical application. Based on these research background, our method was designed to eliminate the negative effect of waterdrops, which would provide a precedent reference in this area. In our method, GAN [26] is served as the backbone of the network. Ren *et al.* [27] proposed a progressive image rainline removal network, that is PreNet, which belongs to a basic network with a very simple structure and was mainly used for rainline removal. Compared with the networks mentioned before, especially Attentive GAN, PReNet has simpler structure and stronger expansibility, which makes it easier to create new networks based on it. Due to these merits, its baseline network was adopted as the generator of GAN in this paper. In Section V, we will show some experimental results between our PReGAN with [15], [24], [25], [27] for comparison.

III. WATERDROP REMOVAL USING PROGRESSIVE RECURRENT GENERATIVE ADVERSARIAL NETWORK

We hereby present the overall structure of the network which we built to track and remove waterdrops precisely and restore lost area in Fig. 1. Like all GANs, the network mainly consists of two parts: generator and discriminator. The reasons for building the network structure as shown in Fig. 1 are as follows:

PReNet has a certain ability of removing waterdrops, but the effect is not satisfactory. In order to improve the ability of waterdrop removal, the PReNet was extended to form our special network for waterdrop removal, and of which the baseline network was used as the generator of GAN. Moreover, to improve image recovery ability after removing waterdrops, we added a **Markovian discriminator**, known as PatchGAN [25], which is a discriminator network that helps generate clearer and more realistic clean images.

The overall flow of operations is: after import degenerated hot-rolled steel strip images with waterdrops, our generator will generate the images as realistically as possible without waterdrops. The generated images will be delivered into the discriminator to distinguish whether the generated images are realistic or not. The objective function of GAN can be described as follows:

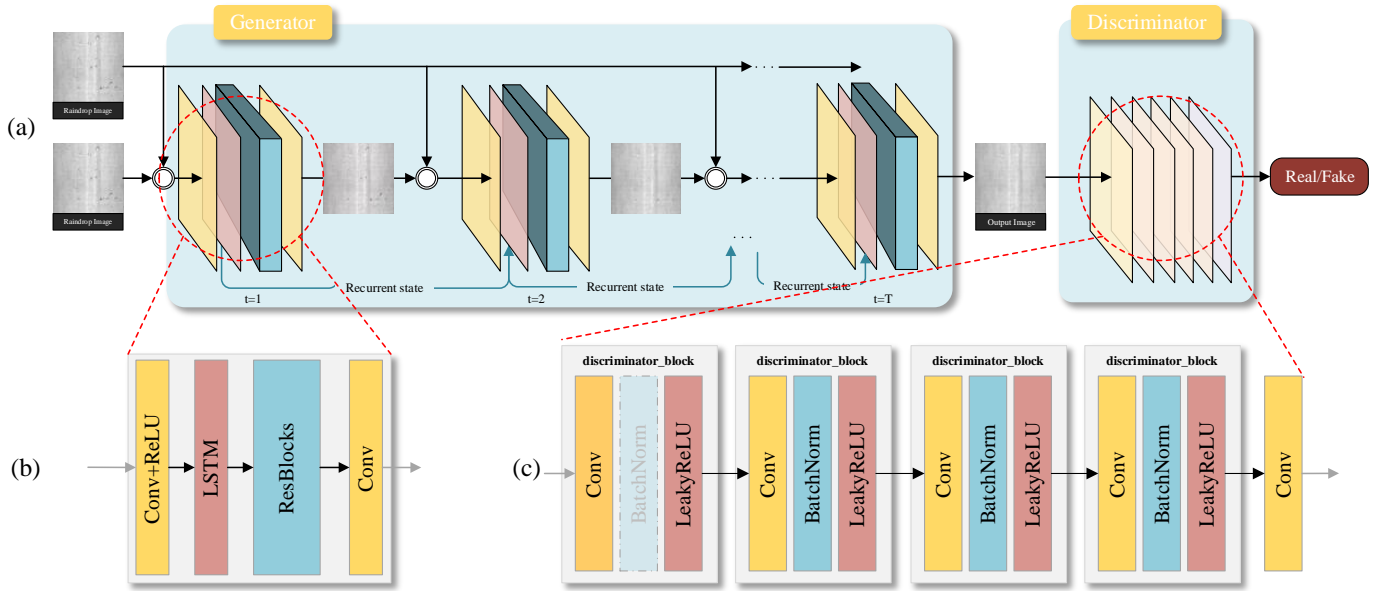


Fig.1 (a) is the overall structure of progressive recurrent generative adversarial network (PreGAN). The generator is a progressive recurrent network[27] and (b) is the network structure of each stage. (c) is the discriminator which is a PatchGAN[25].

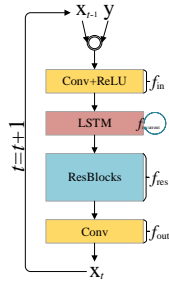


Fig.2 The network structure of each stage, where f_{in} is a convolution layer with ReLU, f_{res} is ResBlocks, f_{out} is a convolution layer, $f_{recurrent}$ is a LSTM structure.

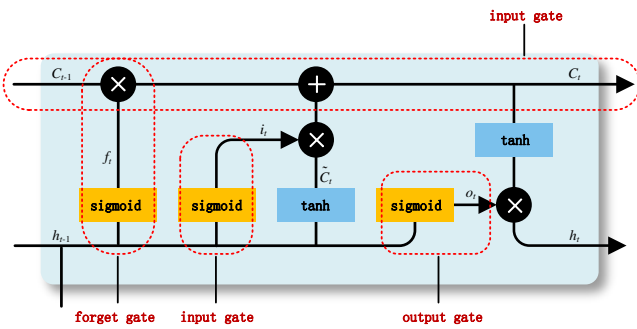


Fig.3 The network structure of LSTM, which consists of an input gate i_t , a forget gate f_t , an output gate o_t and a cell state C_t .

$$V(G, D) = \mathbf{E}_{\mathbf{R} \sim p_{clean}} [\log(D(\mathbf{R}))] + \mathbf{E}_{\mathbf{I} \sim p_{waterdrop}} [\log(1 - D(G(\mathbf{I})))] \quad (1)$$

where G and D represent the generator and discriminator, respectively. \mathbf{I} is a sample with waterdrops extracted from our database of hot-rolled steel strips which is served as the input of generator. \mathbf{R} refers to a waterdrop-free sample. Where G tries to minimize the goal while D tries to maximize to against each other:

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

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

A. Generator

The generator as a progressive recurrent network is shown in Fig. 1(a). Progressive recurrent network is divided into several stages, each of which contains a shallow ResNet (see Fig. 1(b)). For an image with waterdrops, multistage means these waterdrops cannot be removed only in one step, which is a step-by-step process. As shown in Fig. 2, the network in each stage consists of four parts:

- (1) A convolutional layer receives the input of the network, ReLU is chosen as activation function.
- (2) Long and short term memory layer (LSTM) [28] [29] is served as the recurrent layer.
- (3) Several ResBlocks are used for depth representation.
- (4) A convolutional layer is the output of the network, which is applied to export the image after waterdrop removal.

The size of all filters is 3×3 and the filling size is 1×1 . In each stage of the network, the input layer f_{in} is a one-layer convolution with ReLU non-linearity activation function [30]. $f_{recurrent}$, which is a recurrent recursion layer, is a LSTM structure. The structure of our LSTM unit is shown in Fig. 3. The interaction equations between states and gates is expressed as:

$$\begin{aligned} i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\ \tilde{C}_t &= \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \\ o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\ C_t &= f_t * C_{t-1} + i_t * \tilde{C}_t \\ h_t &= o_t * \tanh(C_t) \end{aligned} \quad (4)$$

where i_t, f_t, o_t represents input gate, forget gate and output gate, respectively. C_t refers to a cell state which will be transferred to the LSTM of next recurrent state. h_t represents the output of the LSTM unit. Operator $*$ denotes the convolution operation. The long-term memory of LSTM is derived from forgetting and preservation mechanisms. In the process of removing waterdrops, LSTM can transfer the information of the previous stage to the next stage, with transmitting the important information of the previous stage without being lost, meanwhile forgetting the worthless information. Therefore, the waterdrop removal effectiveness of our network can be guaranteed.

Five ResBlocks compose the f_{res} . f_{out} is a one-layer convolution. The input of each stage of the generator consists of two parts: the output of the previous stage x_{t-1} and the original image with waterdrops y . They are both 3-channel RGB images and have cascade connection. Therefore, f_{in} has 6 channels of input and 32 channels of output in each stage. By parity of reasoning, f_{out} takes 32-channel output from $f_{recurrent}$ as input and outputs a 3-channel RGB image. All convolutional layers used in $f_{recurrent}$ and f_{res} have 32 input channels and 32 output channels. And f_{res} is the critical components of extracting depth representation to remove waterdrops. The diagram of the network structure of each stage showed in Fig. 2 can be expressed as:

$$\begin{aligned} x_{t-0.5} &= f_{in}(x_{t-1}, y) \\ s_t &= f_{recurrent}(s_{t-1}, x_{t-0.5}) \\ x_t &= f_{out}(f_{res}(s_t)) \end{aligned} \quad (5)$$

where f_{in} , f_{res} and f_{out} are stage-invariant, that is, network parameters can be reused in different stages, which can greatly reduce the memory and computing consumption. The recurrent layer $f_{recurrent}$ uses $x_{t-0.5}$ and the recurrent state s_{t-1} as input in stage $t-1$.

f_{in} passes the currently obtained x_{t-1} into the next stage along with the original image y . Compared with only transferring x_{t-1} in [31], the waterdrop removal performance of the generator can be further improved by including the original image y as input in each stage, which is mainly due to the introduction of more auxiliary prior information. The LSTM has advantages in image elimination to achieve recurrent waterdrop removal.

Mixing the objective function of GAN with other losses can enhance operation performance, because the prediction results of the model can be comprehensively measured. In this paper, $L1$ and SSIM [32] loss functions are combined. As the discriminator of GAN, the function remains the same, but the task of the generator is increased. It not only needs to generate realistic image deceiving discriminator, but also needs to make it closer to $L1$ and SSIM of ground truth. When training generator, the same background image pairs with and without waterdrops from database of hot-rolled steel strips were used.

In the generator of PReGAN, the SSIM loss function is expressed as:

$$L_{SSIM} = -SSIM(G(\mathbf{I}), \mathbf{R}) \quad (6)$$

SSIM refers to the operation of structural similarity, which will be specifically introduced in Section V (A). The negative value

is taken when the loss is expressed, and the better when the SSIM is bigger.

The $L1$ loss function is also known as the LeastAbsolute Deviation (LAD) or the LeastAbsolute Error (LAE). In general, it minimizes the sum of the absolute difference between the target value and the estimated value. The $L1$ loss can be denoted as:

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

where $y^{(i)}$ indicates the value of each pixel on the target image, and $\hat{y}^{(i)}$ refers to the pixel value on the estimation image. Then, in PReGAN generator, $L1$ loss function is expressed as:

$$L_{L1}(G) = \mathbf{E}_{\mathbf{R} \sim p_{clean}, \mathbf{I} \sim p_{waterdrop}} [\|\mathbf{R} - G(\mathbf{I})\|_1] \quad (8)$$

Therefore, based on the mixed strategy, the loss function of the final generator is:

$$G^* = \arg \min_G \max_D V(G, D) + (\alpha \cdot L_{L1} + \beta \cdot L_{SSIM}) \cdot \gamma \quad (9)$$

where α is set to 0.75. β is set to 1.1. γ is set to 50.

B. Discriminator

$L1$ loss term only achieves the correctness of the low frequency such as color blocks in the images, which will produce fuzzy results in image generation. In this case, the discriminator of GAN is badly needed to model the high frequency structures such as edges in the images. In addition, modeling for high frequency, it is sufficient to focus our attention only on the structures in the local image patches. Based on these above theories, PatchGAN [25] is employed in our discriminator. The architecture of the discriminator is shown as Fig. 1(c).

This discriminator uses the form of Convolution-BatchNorm-Relu as modules [33], and the network determines whether each $\mathbf{N} \times \mathbf{N}$ color block in the image is positive. The input of ordinary GAN discriminator is mapped into a real number, and the final output is a vector of real (1) or false (0), representing the evaluation to judge whether the whole image is true or false. While PatchGAN maps the input to patch matrix \mathbf{X} of $\mathbf{N} \times \mathbf{N}$. In the original image, each patch corresponds to a vector of real (1) or False (0), and its label is also set as an $\mathbf{N} \times \mathbf{N}$ matrix with all real (1) or false (0). In other words, each output in the output matrix corresponds to a patch of the original image. Intuitively, it is made up entirely of convolutional layers and the results of all blocks are averaged to get the final output of the discriminator. It is worth noting that \mathbf{N} can be much smaller than the entire size of the image, and each \mathbf{X} is the feature map output by the convolutional layer. We can trace back to a patch in the original image from this feature map, and see the influence of this position on the final output result.

The merits of PatchGAN are fewer parameters, faster running speed, and high resolution. Besides, it can achieve the high detail maintenance of generated images. In essential, PatchGAN can be interpreted as a form of a texture or style loss, which can help the generator to better recover image texture features and make the trained model pay more attention to details. The final objective function of GAN remains unchanged, which is still expressed as same as Eq. (3).

IV. WATERDROP DATABASE OF HOT-ROLLED STEEL STRIPS SURFACE

Deep learning networks usually require a large number of samples for models and parameter learnings. Similar to the general deep network used for image restoration, GANs also need more image data for training. By generator, the required images are learned and generated, and then the discriminator further guides the generator to produce more realistic images. Therefore, for GAN, the image database is extremely critical which is directly related to the quality of the final generated images.

However, the scene environment of image acquisition of hot-rolled steel strip suffers from complex, high temperature, mechanical vibration, frequently occurred cooling water dispersion, etc. The poor production environment has the high requirement of data collection and the acquisition is often costly. In addition, the strip moves so fast on a production line that it is extremely difficult to obtain effective images, as a result, only high-speed cameras can be selected to capture the strip images. In this paper, the network needs to feed a group of images for training. The background of two images in each pair is basically the same, and the only difference between the pair of images is that whether there are waterdrops or not. Yet now, it is impossible to directly capture a pair of images with and without waterdrops required by the training network in the same rolling steel background using high-speed cameras, which cannot meet the requirements of training image database. Unfortunately, the available open hot-rolled steel strip database is quite scarce, which greatly limits the development of the algorithms of pseudo defect removal of hot-rolled steel strips.

Faced with the above bottlenecks, we have self-made a database of surface images of hot-rolled steel strip to serve the task of surface waterdrop removal from the actual production line. Because there is no actual industrial production line can get the image pairs we need, and the various waterdrops on the steel strip surfaces are dynamic and vibratory, which cannot be fully simulated. Besides, the field of rain removal often uses artificial database for training networks. Therefore, we have to make the image database artificially. Photoshop is artificially used to strictly create a database containing pairs of images with exactly the same background, with or without waterdrops. The original images captured by high-speed cameras have the resolution of 4096×1024 containing images with and without waterdrops. We truncated the clean surface images of 1000×1000 from the original images as the background images, then extracted the real waterdrop parts of

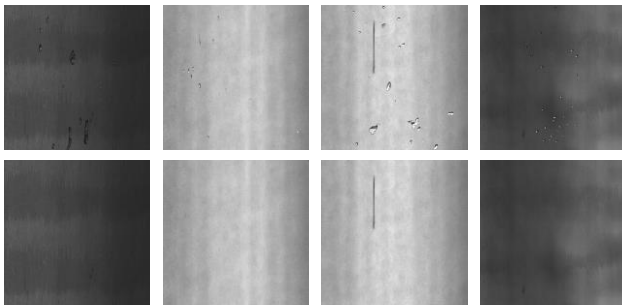


Fig.4 Samples in our hot-rolled steel surface image database. Top: Images degraded with waterdrops. Bottom: The corresponding clean background images.

the original images and pasted it on the background images, and finally 1200 pairs of images with a rigorous production process were produced. Experiments were tested with real waterdrop images to demonstrate that the database can meet the requirements of training. Some of the database samples are shown in Fig. 4. The top four images are images with waterdrops while the bottom four are corresponding clean background images.

V. EXPERIMENTS

A. Evaluation Criteria

It was considered that quantitative measures would usefully supplement and extend the qualitative analysis. Our experiments were evaluated both quantitatively and qualitatively. Finally, to assess the influence of waterdrop removal to the subsequent algorithm, defect detection algorithm is used to test the result before and after waterdrop removal.

(1) Peak signal-to-noise ratio (PSNR) [34] and structural similarity index (SSIM) [32], which are two commonly used indicators, are used to quantitatively evaluate the network's performance in removing waterdrops on the surfaces of hot-rolled steel strips. PSNR is the most common and widely used objective evaluation index. It is built on the error between corresponding pixel points, which is the image quality evaluation method depended on the error sensitivity. However, many experiments have shown that the PSNR score is not exactly consistent with the human visual quality, it is possible that a higher PSNR actually looks worse than a lower, for the intrinsic reason is that PSNR is a linear transformation to decompose image signals does not involve correlation. PSNR is defined as follows:

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_R^2}{MSE} \right) = 20 \cdot \log_{10} \left(\frac{MAX_R}{MSE} \right) \quad (10)$$

Where MAX_R is the maximum value which represents the color of an image point. If each sample point is denoted by 8 bits, it equals 255. MSE stands for mean square error. The smaller the MSE , the larger the $PSNR$; which can explain why the better image quality when the larger $PSNR$.

Natural images are highly structured and there is a strong correlation between pixels, but most error sensitivity based assessment methods always ignore correlation. SSIM is a measure of how similar two images are. From the view of image composition, structural similarity index defines structural information as independent of brightness and contrast, reflecting the property of object structure in the scene, and modeling distortion as the combination of luminance, contrast and structure. The expression of the three and SSIM are:

$$l(\mathbf{R}, \mathbf{G}(\mathbf{I})) = \frac{2\mu_R\mu_{G(\mathbf{I})} + c_1}{\mu_R^2 + \mu_{G(\mathbf{I})}^2 + c_1} \quad (11)$$

$$c(\mathbf{R}, \mathbf{G}(\mathbf{I})) = \frac{2\sigma_R\sigma_{G(\mathbf{I})} + c_2}{\sigma_R^2 + \sigma_{G(\mathbf{I})}^2 + c_2} \quad (12)$$

$$s(\mathbf{R}, \mathbf{G}(\mathbf{I})) = \frac{\sigma_{\mathbf{R} \cdot \mathbf{G}(\mathbf{I})} + c_3}{\sigma_R\sigma_{G(\mathbf{I})} + c_3} \quad (13)$$

TABLE I
COMPARISON OF PReGAN MODELS (T = 6) AT STAGE T = 1, 2, ..., 6, RESPECTIVELY

(T = 6) Stage t		1	2	3	4	5	6
Metric	PSNR	22.0901	25.3139	39.5474	44.7469	42.1231	49.5242
	SSIM	0.0349	0.0843	0.6020	0.7872	0.7023	0.8641

TABLE II
COMPARISON OF PReGAN MODELS WITH DIFFERENT T STAGES

Stages (T)		2	3	4	5	6	7*
Metric	PSNR	49.0149	47.1919	52.2073	47.0660	49.5242	54.8458/-
	SSIM	0.8567	0.9449	0.9502	0.9490	0.8641	0.9518/-

*When T=7, overfitting occurs in the model

TABLE III
QUANTITATIVE EVALUATION RESULTS

Method	Metric	
	PSNR	SSIM
Attentive GAN [15]	15.8075	0.8138
PReNet [27]	47.2200	0.9423
Pix2Pix [25]	48.1232	0.8007
DuRN-S-P [24]	48.6845	0.9558
Ours (PReGAN)	52.2073	0.9502

$$SSIM(\mathbf{R}, G(\mathbf{I})) = f(l(\mathbf{R}, G(\mathbf{I})), c(\mathbf{R}, G(\mathbf{I})), s(\mathbf{R}, G(\mathbf{I})))$$

$$= \frac{(2\mu_{\mathbf{R}}\mu_{G(\mathbf{I})} + c_1)(2\sigma_{\mathbf{R}, G(\mathbf{I})} + c_2)}{(\mu_{\mathbf{R}}^2 + \mu_{G(\mathbf{I})}^2 + c_1)(\sigma_{\mathbf{R}}^2 + \sigma_{G(\mathbf{I})}^2 + c_2)} \quad (14)$$

where $\mu_{\mathbf{R}}, \mu_{G(\mathbf{I})}$ are the mean values of the \mathbf{R} and $G(\mathbf{I})$. $\sigma_{\mathbf{R}}^2, \sigma_{G(\mathbf{I})}^2$ denote the variance of the images \mathbf{R} and $G(\mathbf{I})$, respectively. $\sigma_{\mathbf{R}, G(\mathbf{I})}$ indicates the covariance of the two. c_1, c_2 , and c_3 are constants to avoid the instability caused when the denominator approaches 0. c_3 is usually equal to $c_2/2$. When two images are identical, the value of $SSIM$ is equal to 1. Under normal circumstance, a larger PSNR or SSIM usually suggests better results, indicates that the image distortion is less.

(2) Qualitatively, human visual perception was used to observe and evaluate the removal effect of waterdrops on the image surfaces. We focus on the degree of image recovery after waterdrop removal by various methods and whether a cleaner image can be obtained, which is the people's subjective feeling and different feelings may lead to evaluation varies.

(3) Various defect detection algorithms were used before and after waterdrop removal for images respectively. The purpose of waterdrop removal on steel surface is to reduce the false alarm during defect detection and improve the final detection accuracy. Therefore, it is necessary to determine the influence of our waterdrop removal on the subsequent algorithm.

In this paper, The experiment was performed on PyTorch framework and trained on 12GB Nvidia Titan XP GPU, 2.2GHz Intel Xeon E5-2630 CPU and 64GB RAM. 1100 pairs of image data were randomly extracted as the training set of the network, while the total data pair set is 1200, and the remaining 100 pairs were used as the testing set to obtain quantitative indicators in Section V (B). The qualitative index of the image

is mainly obtained by importing the real hot-rolled steel strip images with waterdrops into the network after training in Section V (C), so as to prove that the database produced can indeed remove waterdrops in the real images in Section V (D).

B. Quantitative Evaluation

The comparison between PSNR and SSIM of PReGAN model (T=6) during stages from 1 to 6 is presented in Table I. From the table, we can see the PSNR and SSIM are gradually increase, indicating that with the iterations through the stages, the waterdrops on the image are gradually removed stage by stage, that is because more precise waterdrops characteristics are learned through these stages.

The complexity of the model greatly affects the processing result. Simple structure will lead to insufficient image learning and failure to obtain accurate features of waterdrops, so that it is inadequate to remove the waterdrops and repair the area after removing the waterdrops. Corresponding, too complex model will also contribute to degenerated results due to the over-fitting. To find an optimal structure, Table II shows the PSNR and SSIM comparison of PReGAN models with different T stages, and different T is related to models with different levels of complexity. During the experiment, when T=7, a relatively good result can occasionally happen, but it often lost in overfit and may collapse. Therefore, the iteration from 2 to 6 only compares, numerous experiments have shown that a better result can be obtained at T=4.

Our method and other existing methods, including Attentive GAN [15], PReNet [27], Pix2Pix [25], DuRN-S-P [24] are also compared in Table III. From the table, our method is significantly superior to other methods in PSNR. Meanwhile, SSIM is higher than other methods except Durn-S-P, but the difference between the two is very small. Undoubtedly, indicated by these two indicators, the results obtained by our method are closer to the image without pseudo defects, and the main reason is that our network is more sensitive to the removal of waterdrops.

C. Qualitative Evaluation

Qualitative Evaluation is applied to more intuitively feel the effects of the waterdrop removal. Fig. 5 shows the image output effect of each stage of the generator (T=6). As can be seen, the waterdrops in the image are gradually removed as the stages progress. When t=4, the waterdrops have been removed relatively cleanly. When the last layer is reached, a basically clean image is finally obtained.

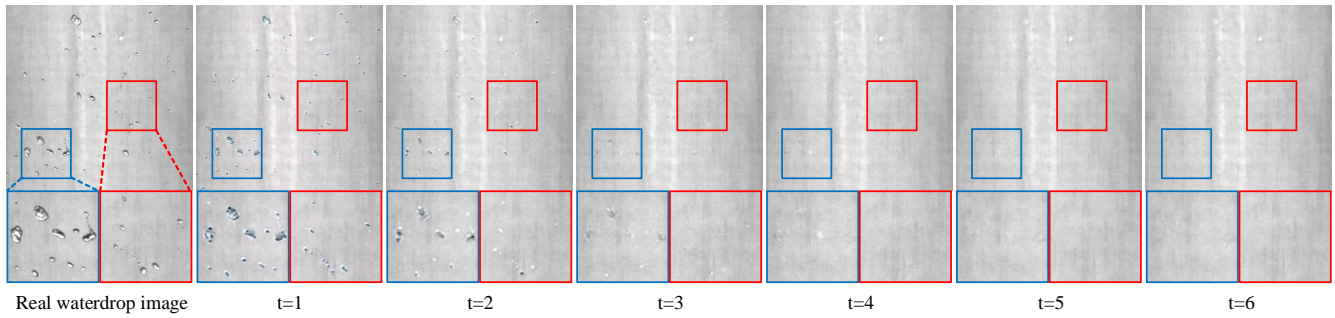


Fig.5 The image output at each stage. From left to right: real waterdrop image, the output of PReGAN (T=6) from the stage 1 to 6.

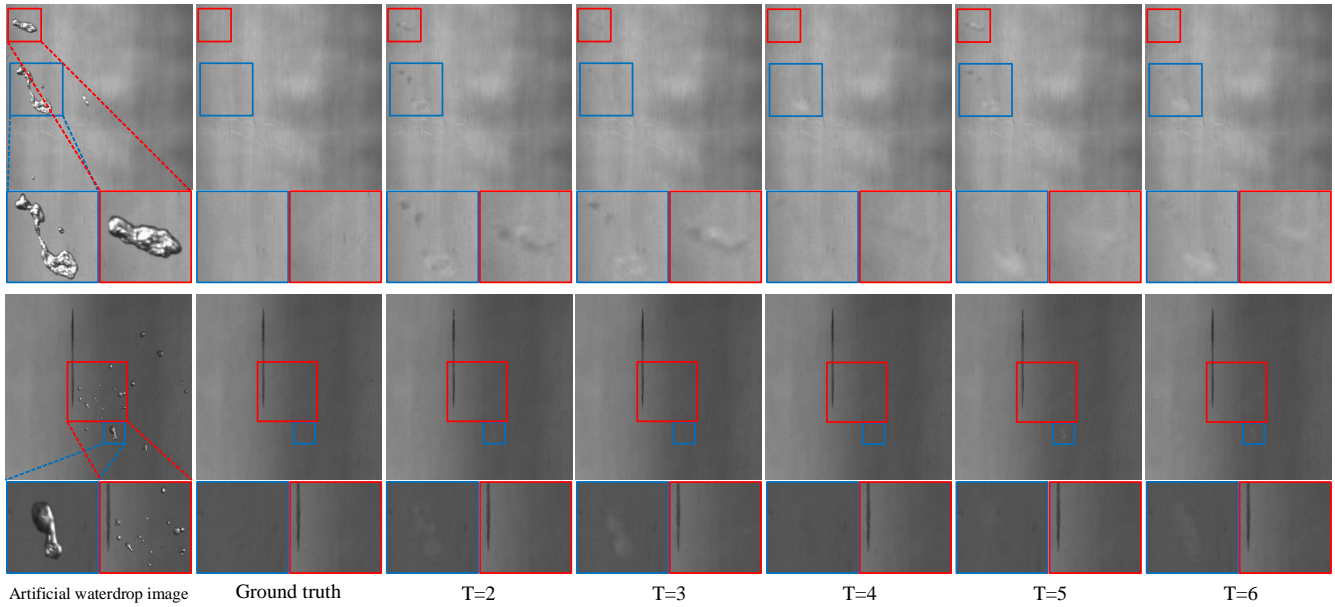


Fig.6 The images of PReGAN models with different T stages. From left to right: artificial waterdrop image (input image), ground truth, the output of PReGAN with T from 2 to 6.

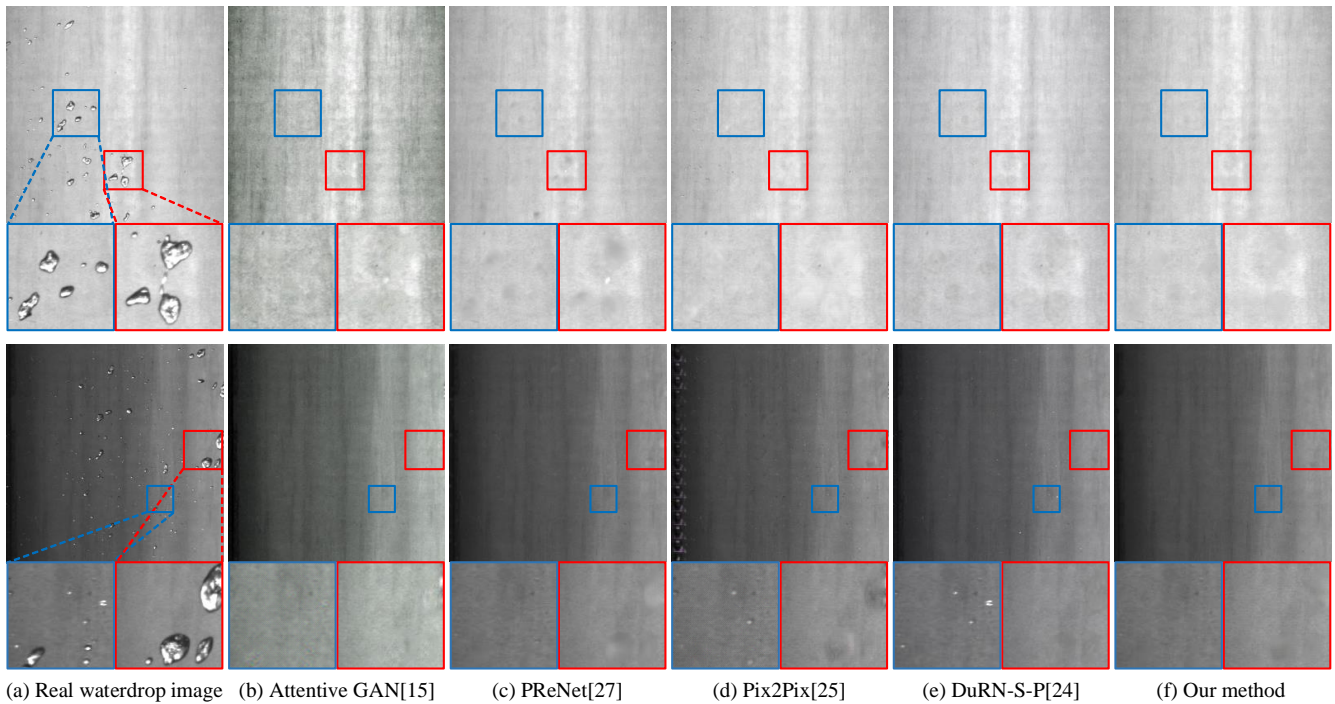


Fig.7 Waterdrop removal effects for different methods. From left to right: Real waterdrop image (input image), Attentive GAN [15], PReNet [27], Pix2Pix [25], DuRN-S-P [24] and our method. Nearly all waterdrops are removed by our method despite the diversity of their sizes and shapes.

TABLE IV
THE AVERAGE NUMBER OF FALSE ALARMS DETECTED IN EACH IMAGE OF 45 REAL WATERDROP IMAGES BEFORE AND AFTER WATERDROP REMOVAL

Method	Real waterdrop image	Attentive GAN	PReNet	Pix2Pix	DuRN-S-P	Our method	our method real raindrop image
Canny	26	13	16	20	17	12	0.46
LBP	34	45	18	21	19	17	0.50
Gabor filtering	43	30	16	43	27	15	0.35
Thresholding	53	19	10	17	15	9	0.17

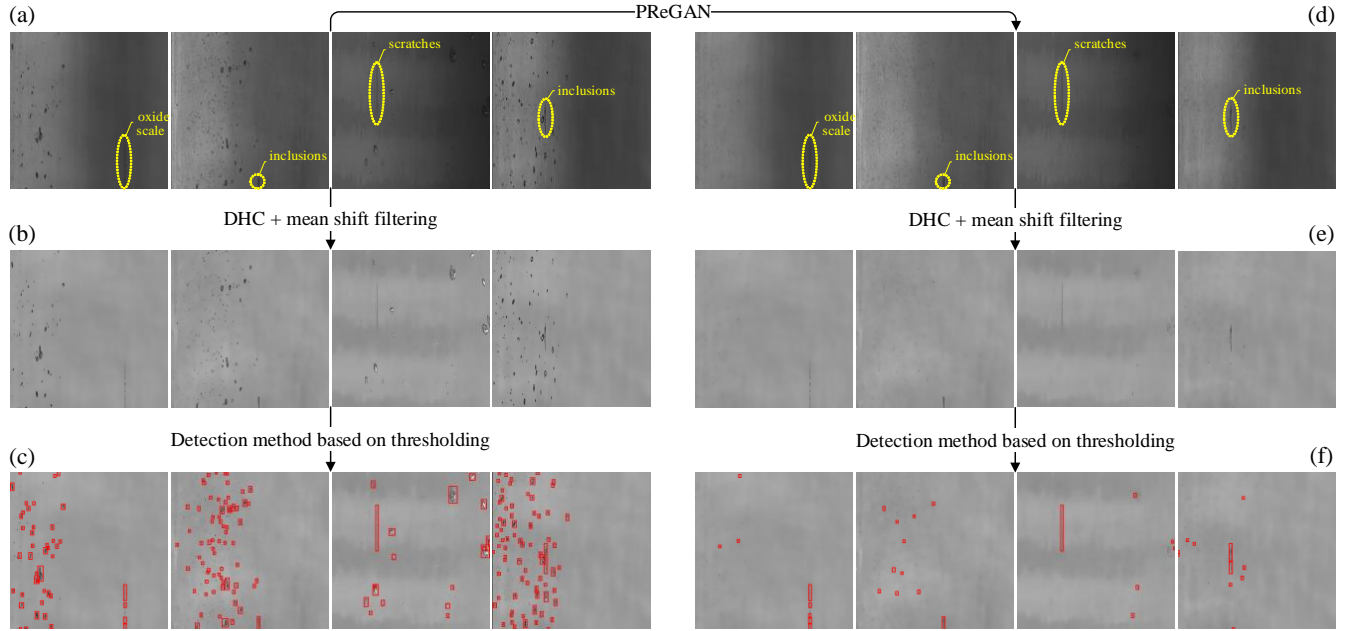


Fig.8 Variance thresholding-based defect detection results before and after waterdrop removal by using PReGAN. (a), (b) and (c) are respectively the raw images, enhanced images and the defect detection results, (d), (e), (f) are respectively the waterdrop removed images, enhanced images and the defect detection results.

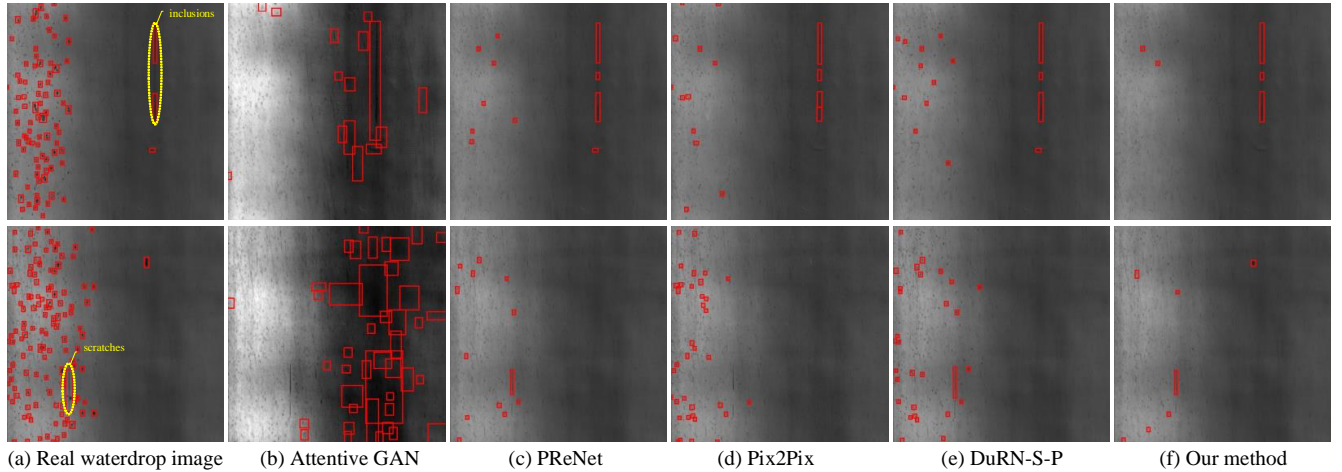


Fig.9 Comparative variance thresholding-based defect detection results after waterdrop removal by using PReGAN and its competitors. (a) without waterdrop removal method, (b) ~ (f) show the defect detection results after using Attentive GAN, PReNet, Pix2Pix, DuRN-S-P and our method.

Qualitative assessment of PReGAN models with different T is showed in Fig. 6. When $T=4$, PReGAN has a better effect of removing waterdrops. When $T=2$ or 3, due to the simple structure of the model, there are still obvious traces of waterdrops after removal. Some waterdrops are also not removed completely. When $T=5$ or 6, for the reason that the model is difficult to train due to the complex structure of the model. When $T=4$, the removal of waterdrops is the cleanest, and the surface texture is also the best repaired after the

removal of waterdrops, mainly because the structure is optimal at this time.

In Fig. 7, we compare the results of Attentive GAN [15], PReNet [27], Pix2Pix [25], DuRN-S-P [24] with our results.

Our database is artificial. Although it is very close to the real image data, there is still some slight differences compared with the real images. What the model ultimately needs to remove is the images obtained on the real industrial production line, so we used the real hot-rolled steel strip images to conduct the test.

Fig. 7 (a) is the real waterdrop image under test, and the resolution is also 1000×1000 , the same as the training image. From Fig. 7 (b) we can see, although all waterdrops are removed from Attentive GAN, the steel strip surface texture was changed and the whole image had color difference with the original image. Fig. 7 (c) is obtained by PReNet, there is no color difference. However, it does not remove waterdrops completely, and the image repair ability of this network is insufficient compared with our method. In Fig. 7(d), the Pix2Pix gains good de-watering results in the image with higher brightness, but in the dark image, the waterdrops part cannot be totally repaired and tiny drops still residual. In addition, there will be some markings on the black part of the image edge, which will affect the later defect detection. Although the DURN-S-P in Fig. 7(e) does not produce color aberration and has higher repair ability, it cannot remove fine waterdrops. Despite the SSIM of our method is slightly lower than Durn-S-P, it can be seen from Fig. 7(f) that tiny waterdrops can be removed thoroughly. Why our methods work well is originated from that the GAN can enhance the image recovery ability compared with the simple PReNet. Therefore, images obtained by our method are closer to the ground truth.

D. Application

The ultimate goal of waterdrop removal is to improve the accuracy of surface defect detection of hot-rolled steel strips. We used four defect detection algorithms to detect the real images before and after the removal of waterdrops. They are edge detection algorithm based on Canny operator, textual analysis algorithm based on Local Binary Patterns (LBP), filtering algorithm based on Gabor transform and thresholding detection algorithm based on block variance. For fair comparison, before using the above *four* detection algorithms, the raw images are preprocessed by using dynamic homogenizing compensation (DHC) [35] and mean shift filtering to suppress the illumination changes and unexpected noises.

For overall evaluation, 45 real hot-rolled steel strip images with 1000×1000 resolution including defects and waterdrops are randomly selected for test. The first step is to estimate the triggered false alarms before and after rain removal. TABLE IV shows the average false alarms per image among the 45 samples after defect detection. In most cases, no matter which detection algorithm is used, the number of false alarms is decreased after rain removal. Notably, the situation using our PReGAN has lowest false alarms, the total average number is reduced even more than a half before that without waterdrop removal. Interestingly, the PReGAN-based waterdrop removal algorithm has cut 83% false alarms when using the simplest variance-thresholding detection method. Nevertheless, there are still a few false alarms after waterdrop removal, some of them are triggered by the tiny waterlines which are out of the repairing scope of the PReGAN, the others are even triggered by the un-repaired water marks, that is why we emphasize the waterdrop repairing effect. Not surprisingly, better quantitative figures in TABLE III brought less false alarm in TABLE IV.

Again, from TABLE IV, an open issue can be discussed here, for surface defect inspection, do we need a better preprocessing algorithm or a better defect detection algorithm? Our proposal is, you can use dirty and messy images to challenge your

detection algorithm, as do this could make your algorithm more intelligent and robust, but do not do the same to your AVI instrument installed on the actual production line. For industrial instruments, stability is everything. Back to the theme, to make the raw images clearer and more informative is the design goal of the proposed PReGAN, which assists detection algorithms to trace surface defects more easily and accurately. In this case, the detection results are quite different even for the simplest variance thresholding-like methods.

For more insights, we select the above variance thresholding method for some visual investigation. As shown in Fig. 8, we randomly choose four samples for illustration. In particular, three kinds of defects, oxide scales, inclusions and scratches are involved. From the left half side, although the image quality has been greatly improved by using the image enhancement algorithm, massive false alarms are still triggered for the massive waterdrops spreading on the steel surface. After using PReGAN in the right half figure, the false alarms are dramatically decreased and the surface defect can be inspected easier. It is worth mentioning that, In the case of the image representation effect of Fig. 8(e), the remaining few false alarms can be easily addressed by improve or replace the thresholding-based detection method, because this paper mainly focuses on waterdrop removal rather than defect detection, we did not further expand this part.

Further, Fig. 9 performs a comparison test between the proposed PReGAN and its competitors, Attentive GAN, PReNet, Pix2Pix, DuRN-S-P. Similarly, we randomly select two image samples from the test in TALBE IV. As shown in Fig. 9(a), for the raw images, not surprisingly, a large quantities of false alarm are triggered especially in the left part of the image samples. Then in Fig. 9(b), the Attentive GAN addresses this problem to a large extent but another side effect arises, bringing some other false alarms for the texture morphology is changed unexpectedly. From Fig. 9(c)~(f), all of these methods can identify defects except for some small and non-obvious roll marks, the false alarms are decreased to a large extent at the same time. As expect, our method obtained the least false detection. The above preliminary experiments show good evidance that our proposed waterdrop removal algorithm could be considered by the AVI instrument of steel surface as a efficient image preprocessing method.

VI. CONCLUSION

A fundamental issue that the pseudo-defects like waterdrops spreading over the hot-rolled steel strip surface can be considered as a “rain-like layer” has been investigated in this paper. Aiming at the tough false-alarm problem triggered these waterdrops, a progressive recurrent generation adversarial network (PReGAN) is designed for active waterdrop tracking and fine-grained image inpainting. Extensive experiments prove that the images enhanced by our method perform the most informative and spotless, with **52.2073** peak signal-to-noise ratio (PSNR) and **0.9502** structural similarity index (SSIM), when compared with three prestigious networks, which indicate that our method can robustly suppress the interference of waterdrops while preserve other image details reliably. In addition, a high-resolution image database for steel surface waterdrop removal is opened for the first time in this paper. As far as we known, this is the first attempt to build a

learning method to promote the defect inspection performance of the current AVI instrument from a fresh perspective of pseudo defect removal for hot-rolled steel strips. This progress shows a positive sign that it is possible to carry out the visual inspection after the laminar spray cooling process, moving one step for the traditional layout that the AVI instrument is placed closely after the final hot-rolling mill and before the spray cooling process, where the AVI components always survive hard to anti the ultra-high temperature.

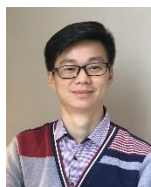
However, the above preliminary results are obtained in a single production line and single spray cooling system, Valin Lian Steel. The proposed PReGAN waterdrop removal method might not universally adapt to other steel mills, as the surface water droplet distribution rule in different steel mill varies a lot. Updating the image database and retraining the algorithm are highly recommended. The generalizability of the proposed method should be verified urgently in the near future.

REFERENCES

- [1] 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.
- [2] 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.
- [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, J. Su, J. Zhou, B. Zhou, C. Yang, *et al.*, "Automated Visual Defect Classification for Flat Steel Surface: A Survey," *IEEE Trans. Instrum. Meas.*, Oct. 2020, vol.69, no. 12, pp.9329-9349.
- [6] Q. Luo, X. Fang, L. Liu, C. Yang, *et al.*, "Automated visual defect detection for flat steel surface: A survey," *IEEE Trans. Instrum. Meas.*, Feb. 2020, vol. 69, no. 3, pp. 626-644.
- [7] X. Wen, K. Song, L. Huang, M. Niu, and Y. Yan, "Complex surface ROI detection for steel plate fusing the gray image and 3D depth information," *Optik*, vol.198:163313, Dec. 2019.
- [8] X. Wen, K. Song, M. Niu, Z. Dong, and Y. Yan, "A three-dimensional inspection system for high temperature steel product surface sample height using stereo vision and blue encoded patterns," *Optik*, vol. 130, pp.131-148, Feb. 2017.
- [9] 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.
- [10] Y. Wang, K. Song, J. Liu, H. Dong, Y. Yan, and P. Jiang, "RENet: Rectangular convolution pyramid and edge Enhancement network for salient object detection of pavement cracks," *Measurement*, vol.170:108698, Jan. 2021.
- [11] 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.*, doi: 10.1109/TII.2020.3045196, 2021.
- [12] G. Song, K. Song, and Y. Yan, "Saliency detection for strip steel surface defects using multiple constraints and improved texture features," *Opt. Laser. Eng.*, Vol. 128:106000, May 2020
- [13] 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.
- [14] 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.
- [15] R. Qian, R. T. Tan, W. Yang, J. Su, J. Liu, "Attentive generative adversarial network for raindrop removal from a single image," *IEEE Conf. Comput. Vision Pattern Recogn.*, pp. 2482-2491, 2018.
- [16] W. Yang, R. T. Tan, J. Feng, J. Liu, Z. Guo, and S. Yan, "Deep joint rain detection and removal from a single image," *IEEE Conf. Comput. Vision Pattern Recogn.*, pp. 1357-1366, Feb. 2017.
- [17] J. Liu, W. Yang, S. Yang, and Z. Guo, "D3R-Net: Dynamic routing residue recurrent network for video rain removal," *IEEE Trans. Image Process.*, vol. 28, no. 2, pp. 699-712, 2019.
- [18] 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, Feb. 2017.
- [19] H. Kurihata, T. Takahashi, I. Ide, Y. Mekada, H. Murase, Y. Tamatsu, and T. Miyahara, "Rainy weather recognition from in-vehicle camera images for driver assistance," *IEEE Proc. Intelligent Vehicles Symposium*, pp. 205-210, 2005.
- [20] M. Roser and A. Geiger, "Video-based raindrop detection for improved image registration," *Computer Vision Workshops, 2009 IEEE International Conference*, pp. 570-577, 2009.
- [21] M. Roser, J. Kurz, and A. Geiger, "Realistic modeling of waterdrops for monocular adherent raindrop recognition using bezier curves," *Asian Conference on Computer Vision*, pp. 235-244, 2010.
- [22] D. Eigen, D. Krishnan, and R. Fergus, "Restoring an image taken through a window covered with dirt or rain," *Computer Vision, 2013 IEEE International Conference*, pp. 633-640, 2013.
- [23] M. R. Kanthan and S. N. Sujatha, "Rain drop detection and removal using k-means clustering," *Proc. International Conference on Computational Intelligence and Computing Research*, pp. 1-5, 2015.
- [24] X. Liu, M. Suganuma, Z. Sun, and T. Okatani, "Dual residual networks leveraging the potential of paired operations for image restoration," *IEEE Conf. Comput. Vision Pattern Recogn.*, 2019.
- [25] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," *IEEE Conf. Comput. Vision Pattern Recogn.*, 2016.
- [26] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Mach. Learn.*, pp. 2672-2680, 2014.
- [27] D. Ren, W. Zuo, Q. Hu, P. Zhu, and D. Meng, "Progressive image deraining networks: A better and simpler baseline," *IEEE Conf. Comput. Vision Pattern Recogn.*, 2019.
- [28] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735-1780, 1997.
- [29] X. Shi, Z. Chen, H. Wang, D. Y. Yeung, W. K. Wong, and W. C. Woo, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," *IEEE Conf. Comput. Vision Pattern*

Recogn., pp. 802–810, 2015.

- [30] V. Nair and G. E. Hinton, “Rectified linear units improve restricted boltzmann machines,” *Proc. International Conference on Machine Learning*, pp. 807–814, 2010.
- [31] X. Li, J. Wu, Z. Lin, H. Liu, and H. Zha, “Recurrent squeeze-and-excitation context aggregation net for single image deraining,” *IEEE Conf. Comput. Vision Pattern Recogn.*, pp. 262–277, 2018.
- [32] 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, 2004.
- [33] S. Ioffe and C. Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” *Mach. Learn.*, Mar. 2015.
- [34] 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, 2008.
- [35] 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, 2016.



QIWU LUO (M’17) received 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 was a Senior Engineer of instrumentation with WASION Group Ltd. Company, Changsha, and the Deputy Technical Director with Hunan RAMON Technology Co., Ltd., Changsha. In 2016, he joined 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 current research interests include computer vision, industrial AOI, machine learning, parallel hardware architecture design, and reconfigurable computing.



KEXIN LIU received her 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.



JIAOJIAO SU received her B.S. degree in electronic information science and technology from the Hefei University of Technology in June 2020. 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 defect detection, image classification, and machine learning.



CHUNHUA YANG (M’09) 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.



WEIHUA GUI (M’09) received the B.Eng. degree in electrical engineering from Central South University, Changsha, China, in 1976, respectively.

From 1986 to 1988, he was a visiting scholar with the University of Duisburg-Essen, Duisburg, Germany. Since 1991, he has been a Professor with the School of Automation, Central South University. Since 2013, he has been an Academician with the Chinese Academy of Engineering, Beijing, China. His research interests include modeling and optimal control of complex industrial processes, distributed robust control, and fault diagnosis.



LI LIU (M’08-SM’19) received the Ph.D. degree in information and communication engineering from the National University of Defense Technology (NUDT), China, in 2012. She is currently a Professor with the College of System Engineering. 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 at the Chinese University of Hong Kong. From 2016.12 to 2018.11, she worked as a senior researcher at the Machine Vision Group at the University of Oulu, Finland. Her current research interests include computer vision, pattern recognition and machine learning. Her papers have currently over 3700 citations in Google Scholar.



Olli Silvén 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. His research interests include ultra-energy-efficient embedded signal processing and machine vision system design. 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.