IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 18, 2020, accepted October 23, 2020, date of publication October 27, 2020, date of current version November 9, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3034238

Wavelet Based Deep Recursive Pyramid Convolution Residual Network for Single Image Rain Removal

TIAN TIAN GONG[1] **AND JUN SHENG WANG**[2]
[1]School of Mathematics and Information Science, North Minzu University, Yinchuan 750021, China
[2]School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094, China

Corresponding author: Jun Sheng Wang (wangjunshengring@163.com)

ABSTRACT Image rain removal aims to separate the background image from the rainy image. During the past three years, the image rain removal with deep convolutional neural networks has achieved impressive performance. However, how to reach tradeoff between high de-raining performance and low model parameters is still a challenge. To address the issue, the paper is devoted to exploring a novel method based on wavelet deep recursive pyramid convolution residual network (WDRPRN), in which discrete wavelet transform is embedded to decompose the rainy image in different frequency domains, and the deep recursive pyramid convolution residual network (DRPRN) can well predict the residual coefficients between rainy image and clean image. In addition, compared with other neural networks, the DRPRN adopts recursive model that can cost fewer parameters. Plentiful of experiments on synthetic and real-world datasets show that the proposed method is significantly superior to the recent state-of-the-art algorithms.

INDEX TERMS Image rain removal, wavelet transform, residual coefficients, low model parameters.

I. INTRODUCTION

Image rain removal is a fundamental task in low-level computer vision. The bad rainy images with rain streaks and fog artifacts not only degrade human visual perception, but also reduce the accuracy of many high-level computer vision tasks, such as pedestrian detection [1], visual tracking [2], scene analysis [3] and saliency detection [4]. Thus, it is a vital research work to remove the rain streaks and recover the details of the background image. This paper is committed to the study of single image rain removal.

The light rainy image is mainly composed of slender rain streaks and detail background image. In addition, for heavy rainy image, this may be caused by rain accumulation [5], which can form a veil on the background image. All in all, the rain models are complex. The commonly used rain models can be summarized as the additive composite model of rain streak image **S** and clean background image **B**. Hence, the rain model can be expressed as follows:

$$\boldsymbol{O} = \boldsymbol{B} + \boldsymbol{S} \tag{1}$$

The associate editor coordinating the review of this manuscript and approving it for publication was Yiming Tang.

where the purpose of rain removal is to decompose the input rainy image **O** and obtain the rain-free image **B**. It can be observed from (1) that there are two unknown variables in the rain model, which can lead to the infinite solution space of the rain model. Therefore, it's an ill-posed problem.

In general, the existing rain removal methods can be grouped into two categories which are model-driven and data-driven methods. The model-driven methods refer to prior knowledge [6]–[12] and physical handcrafted feature [9], [13], [14] to constrain the ill-posed problem, where the physical handcrafted feature is relevant to repeatability [9], high-frequency property [13] and directionality [14], and the prior knowledge extracts rain streak discriminative characteristics based on different priors or physical handcrafted features. The model-driven methods can be summarized as follows: sparse coding-based methods [6]–[8], low-rank representation-based methods [9], [10], Gaussian mixture model-based (GMM) methods [11], [12] and guided filter methods [15], [16]. Although these rain removal methods based on prior knowledge afterwards building optimization cost have made great progress, the assumptions about rain streaks do not fully consider the complex factors in the real world, and only specific rain streaks can be removed, not all rain streaks.

Recently, the data-driven rain removal methods [5], [17]–[20] have attracted people's attention, which benefits from the powerful non-linear mapping function of deep CNN, without tedious handcrafted features or prior knowledge. Although those deep networks have achieved significant performance in single image de-raining, it is hard to strike the balance between low model parameters and advanced performance. To handle this issue, we propose a light wavelet deep recursive pyramid convolution residual network (WDRPRN), which only owns about 68k parameters, and the de-raining performance metrics overtake some advanced methods. We also use pyramid convolution with multiple filter scales to extract rain streak features of different scales, and make an objective analysis by making a diagram and explain the underlying reasons why this paper use it as follows:

FIGURE 1. Rain streaks of different scales.

The rain streaks in rainy image have obvious multi-scale characteristics. Figure 1 shows that the rain streaks in different areas have different range and direction characteristics. How to effectively extract rain streak features of different scales and design robust rain removal algorithm is critical. Because the convolution of a single filter scale is regular, it can only extract the local features of the relatively fixed size window, while the pyramid convolution of multiple filter scales can extract the features of different size windows, which corresponds to the multi-scale characteristics of rain streaks. Therefore, concatenating multi-scale features is beneficial to improve the performance of rain removal. Multi-scale features have been widely used in visual recognition (Pyconv) [21], image segmentation (Aspp) [22] and rain removal algorithms (ReMAEN [23] and LPNet [24]), etc., which have been proven to have great benefits for the improvement of computer vision algorithms.

In particular, we pay more attention to the rain streaks removal in the frequency domain rather than directedly removing the rain streaks from spatial domain. Firstly, the rainy image is decomposed into different sub-band images by utilizing wavelet transform. Secondly, the recursive pyramid convolution residual network model is conducted to extract the context information from different scales. Finally, the wavelet loss and SSIM loss are selected to update the parameters of WDRPRN until the network converges. In order to obtain better de-raining effect, we propose an enhanced WDRPRN for stage-by-stage rain removal. The main contributions of this paper are summarized as follows:

- A novel wavelet-based approach for deep CNN single image rain removal research is presented. We use wavelet transform to decompose the rainy image into four sub-bands, and then conduct rain streaks removal process. Compared with the direct use of deep learning to implement rain streaks removal, the biggest advantage of the proposed method is that it can separate the background details and rain streak structure in different frequency domains. Effective separation processing can reduce the training process of deep learning network and improve the performance of the model to rain streaks removal.
- This paper is also committed to designing a more effective and lighter recursive model to predict the residual rain streak image, which can achieve a better trade-off between better rain removal performance and fewer model parameters.
- The combination of wavelet-based loss and SSIM loss is used to optimize the network in the sub-band images, where the wavelet-based loss can generate better satisfactory result on the pixel level, and the SSIM loss can better preserve the edge details of the background image.

The remainder of this paper is arranged as follows. In Section II, we give a brief overview of the existing rain removal related work and wavelet transform. Section III introduces the detailed framework of wavelet deep recursive pyramid convolution residual network. To verify the de-raining performance of the proposed method, extensive fundamental experiments and ablation studies are conducted in Section IV. Finally, the conclusion is presented in Section V.

II. RELATED WORK

A. RAIN REMOVAL RESEARCH

The existing image rain removal methods are divided into two categories, which are video-based rain removal and single image rain removal. Compared with the single image rain removal, the video-based methods can acquire the temporal information between image frames. Thus, the single image rain removal is a troublesome problem. Kang *et al.* [13] has decomposed the rainy image into high-frequency and low-frequency layers by using bilateral filter, showing that the rain streaks can be removed via dictionary learning and sparse coding in high-frequency layer. Luo *et al.* [6] proposed a non-linear screen blend model, which can eliminate rain streaks by discriminative sparse representation and dictionary learning. Chen and Hsu [9] proposed a low-rank model from matrix to tensor structure considering the similar and repetitive patterns of rain streaks. Chen *et al.* [25] raised a two-stage error-optimized sparse representation (EOSR) model, in which the sparse representation stage uses dynamic patch error constraint to generate rain-free image, and the error-optimized sparse reconstruction stage adopts multi-objective optimization to obtain the final rain-free image with few rain streaks and texture preservation. Li *et al.* [11] used

patch-based Gaussian mixture model (GMM) prior knowledge to separate the image background layer and rain streak layer. Deng *et al.* [14] established a global sparse model utilizing the intrinsic direction and structural information of rain streaks, showing that it can efficiently remove rain streaks from rainy images.

The convolutional neural network has been proved to have promising performance in image processing, e.g., image denoising [26], [27], image deblurring [28], [29], image super resolution reconstruction [30], [31], image inpainting [32] and linguistic steganography [33]. The de-raining methods based on deep learning have also been greatly developed. For instance, Fu *et al.* [17] raised a deep negative residual network to directly learn the relationship of rain streak mapping in the high frequency layer. Yang *et al.* [5] built the recurrent multi-task rain detection and removal network, where the contextualized dilated network extracts features of rain streaks with different shapes and directions. Zhang and Patel [18] proposed a joint rain-density classifier and multi-stream densely connected de-raining network, which removes rain streaks under the guidance of the estimated rain density. Li *et al.* [19] raised a recurrent deep convolutional neural network, in which the intensity and transparency of rain streaks can be simulated by using the squeeze-and-excitation block. Hu *et al.* [34] established a rain imaging model based on scene depth effect, and then constructed depth-guided attention rain removal network to guide the main network to output residual image, which is mainly suitable for RainCityscapes images with depth information. Chen *et al.* [20] proposed the smoothed dilated convolution end-to-end gated network, which tackles the gridding artifacts by the smoothed dilated convolution and merges the features of different levels via the gated sub-network.

B. RELATED APPLICATION BASED ON WAVELET TRANSFORM

Wavelet transform uses a series of wavelet functions with different scales to decompose the original function, and the coefficients of the original function in different scales are obtained after transform. As wavelet transform can decompose image signal into multi-scale and different sub-bands, it has been widely applied in some low-level image processing tasks, such as image super-resolution [35], [36], image restoration [37], image rain removal [38] and image dehazing [39]. Guo *et al.* [35] established the deep wavelet super-resolution network (DWSR), which can predict the missing details between the wavelet coefficients of low-resolution images and high-resolution images, and greatly reduce the amount of computation. Huang *et al.* [36] developed a wavelet-based CNN network for face super-resolution, and the network directly predicted a series of HR wavelet coefficients corresponding to LR. Liu *et al.* [37] proposed multi-level wavelet CNN(MWCNN), which embeds discrete wavelet transform to CNN, as a replacement for pooling operation. Yang *et al.* [38] built a recurrent wavelet learning dilated residual dense network (RWL), which decomposes the rainy image into sub-band images of different scales, and then recurrently reconstructs the rain-free image. Yang and Fu [39] raised wavelet U-net, where the down-sampling and up-sampling can also be replaced by discrete wavelet transform (DWT) and inverse discrete wavelet transform (IDWT). Compared with recurrent wavelet dilated residual dense network, our method employs once wavelet transform then following a light recursive pyramid convolution [21] residual network, in which the grouped convolution is the key benefit for multi-scale features extraction and model parameters control.

III. PROPOSED METHOD

A. WAVELET DECOMPOSITION

The rain removal problem of single image can be regarded as the problem of restoring the detail structure of background image under the premise of inputting rainy image, and the wavelet decomposition can capture the image structural information in different directions. Therefore, the wavelet decomposition can well be embedded in the image rain removal task. Figure 2 shows the basic flow chart of $2-D$ discrete wavelet decomposition. For the input image signal, filters and down-sampling are used in the directions of column and row, and four sub-band images after $2-D$ discrete wavelet decomposition can be obtained. In this process, Firstly, the approximate representation LL of the image can be obtained by two low-pass filters G_L and down-sampling along the column and row directions respectively. Secondly, the low-pass filter G_L and down-sampling are used in the column direction, and then the high-pass filter G_H and down-sampling are utilized in the row direction to obtain the vertical sub-band image LH. Thirdly, the high-pass filter G_H and down-sampling are conducted along columns, and then the low-pass filter G_L and down-sampling are used along rows to gain the horizontal sub-band image HL. Finally, the diagonal sub-band image HH is obtained by two high-pass filters G_H and down-sampling in the column and row directions respectively. The $2-D$ discrete wavelet decomposition can be expressed as follows:

$$\begin{aligned} \psi_{LL}(x,y) &= \phi(x)\phi(y) \\ \psi_{LH}(x,y) &= \phi(x)\varphi(y) \\ \psi_{HL}(x,y) &= \varphi(x)\phi(y) \\ \psi_{HH}(x,y) &= \varphi(x)\varphi(y) \end{aligned} \tag{2}$$

where ϕ and φ indicate $1-D$ scaling function (i.e. $1-D$ low-pass filter) and $1-D$ wavelet function (i.e. $1-D$ high-pass filter). Therefore, we can get multi-resolution edge detail sub-band images of rainy images with different scales and directions through wavelet transform. Furthermore, the DRPRN is conducted to effectively remove rain streaks in four sub-band images, and final rain-free image can be reconstructed by inverse discrete wavelet transform.

B. DETAIL STRUCTURE OF THE WDRPRN

As illustrated in Figure 3, our wavelet-based deep recursive pyramid convolution residual network mainly consists of

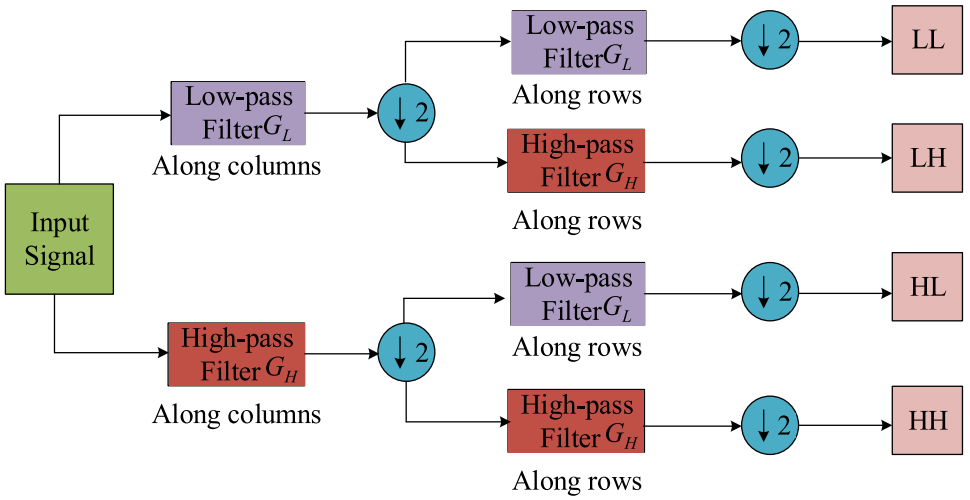


FIGURE 2. Flow diagram of the 2 – *D* discrete wavelet transform.

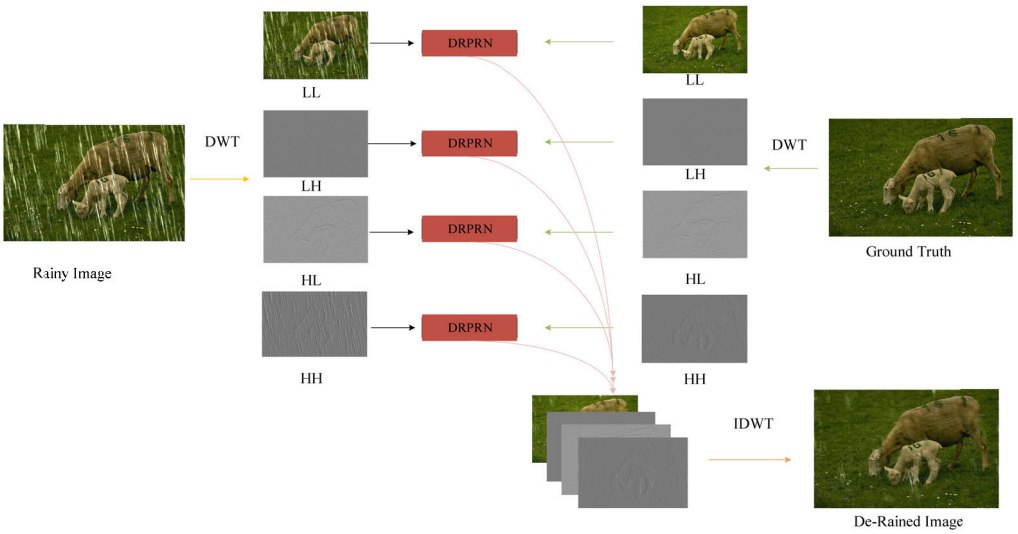


FIGURE 3. The overall architecture of wavelet based deep recursive pyramid convolution residual network.

three stages: RGB image discrete wavelet transform (DWT), wavelet coefficient prediction network (DRPRN) and inverse discrete wavelet transform (IDWT). The purpose of discrete wavelet transform is to decompose the rainy image into four sub-band images in wavelet domain. Furthermore, the wavelet coefficient prediction network outputs four rain-free wavelet coefficient images. Finally, the inverse discrete wavelet transform is used to reconstruct the de-raining image from four rain-free sub-band images.

1) THE WAVELET COEFFICIENT PREDICTION NETWORK

The input is $3 \times h \times w$ rainy image. Through Eq (2), we can use DWT on 3-channels of the rainy image, and the output is four $3 \times \frac{h}{2} \times \frac{w}{2}$ sub-band images. Then we concatenate the channels of four sub-band images. As the whole input of the wavelet coefficient prediction network shown in Figure 4, compared with the rain removal methods only based on deep learning, the biggest advantage of this network is that it makes full use of the multi-scale edge features, which is conducive to the rain streaks removal and the background details restoration. Furthermore, in order to reduce the model parameters, we also adopt the recursive learning manner similar to the DRRN [40] as shown in Figure 5(a). To enlarge the receptive filed of the network, the pyramid convolution (PyConv) [21] is used to construct the recursive block as shown in Figure 5(b). The (PyConv) process divides the input feature maps into different groups, and applies different depth kernel sizes ($3 \times 3, 5 \times 5, 7 \times 7$ and 9×9) to each input feature map group independently, which can strengthen the ability of the network extraction multi-scale features corresponding to the rain streak features, and reduce the computational cost and

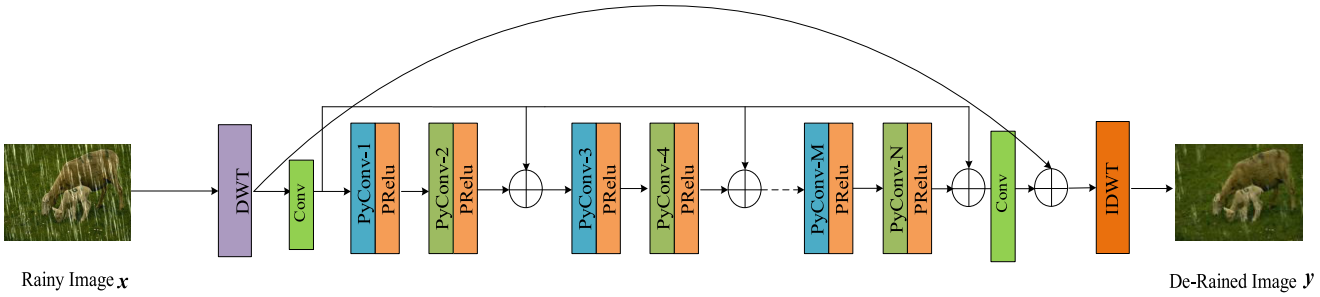


FIGURE 4. The details of the proposed wavelet coefficient prediction network.

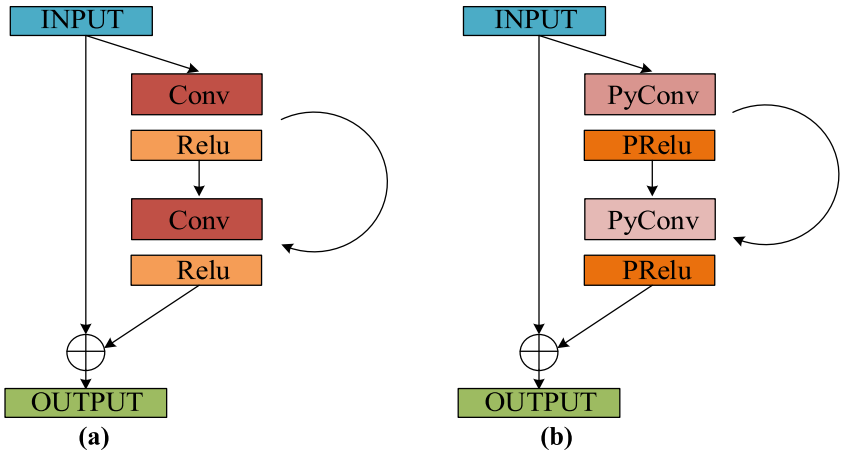


FIGURE 5. Structure of different recursive blocks: (a) recursive block of DRRN [33] and (b) recursive block of the proposed DRPRN.

parameters. The number of recursive blocks is 16, and the weight of recursive blocks is shared. We adopt zero padding to keep the feature map of layer the same size as the input tensor, so that each wavelet coefficient feature map is the same size as the input. There are 64 channels in the middle layer of our network. The inner pyramid convolution of recursive block is composed of Convolution (Conv) with different kernel sizes and Parametric Rectified Linear Unit (PReLU) [41] activation function.

After four rain-free coefficient sub-band images are obtained by using the wavelet coefficient prediction network, the final rain-free image can be reconstructed by the inverse discrete wavelet transform (IDWT).

2) ENHANCED WAVELET BASED DEEP RECURSIVE PYRAMID CONVOLUTION RESIDUAL NETWORK

It is particularly worth mentioning that considering that the heavy rain streaks are not easy to be removed in the first stage, we stack three WDRPNRs named WDRPRN-E. As shown in Figure 6, the input of the current stage is previous stage de-raining result except for the first stage. We can formulate the following procedures:

$$\begin{aligned} y_t &= f_{WDRPRN}(y_{t-1}) \\ y_1 &= f_{WDRPRN}(x_1) \\ 1 &< t \leq 3 \end{aligned} \tag{3}$$

where y_t indicates de-rained image in the $t-th$ stage, x_1 represents the input rainy image, and $\mathbf{t}$ is the range of recurrent network stage.

C. LOSS FUNCTION

Mathematically, the rain removal process aims to learn the map function $F_\theta(x)$, where θ denotes the parameter of WDRPRN. Given $\{(x_i, y_i)\}_{i=1}^{N}$ represents the rainy/ground-truth image pairs of the training dataset. The common objective loss function for rain removal is Mean Square Error loss (MSE loss) on the pixel level, which can be defined as:

$$L(\theta) = \frac{1}{N}\sum_{i=1}^{N} \left\| F_\theta(x_i) - y_i \right\|^2 \tag{4}$$

where x_i and y_i denote the $i-th$ input rainy image and corresponding ground truth rain-free image respectively. However, MSE loss may over penalize the pixel value errors, which can result blurred de-rained image. Following [24], we use L_1 loss to train our network. Different with [24], the L_1 loss is used in wavelet domain. In addition, we add the SSIM loss to better preserve the edge details. Thus, our network consists of two loss functions as:

$$\begin{aligned} L_1(\theta) = \frac{1}{N}\sum_{i=1}^{N} \Big\| & LL\left(F_\theta(x_i), y_i\right) + LH\left(F_\theta(x_i), y_i\right) \\ & + HL\left(F_\theta(x_i), y_i\right) + HH(F_\theta(x_i), y_i) \Big\|^1 \end{aligned} \tag{5}$$

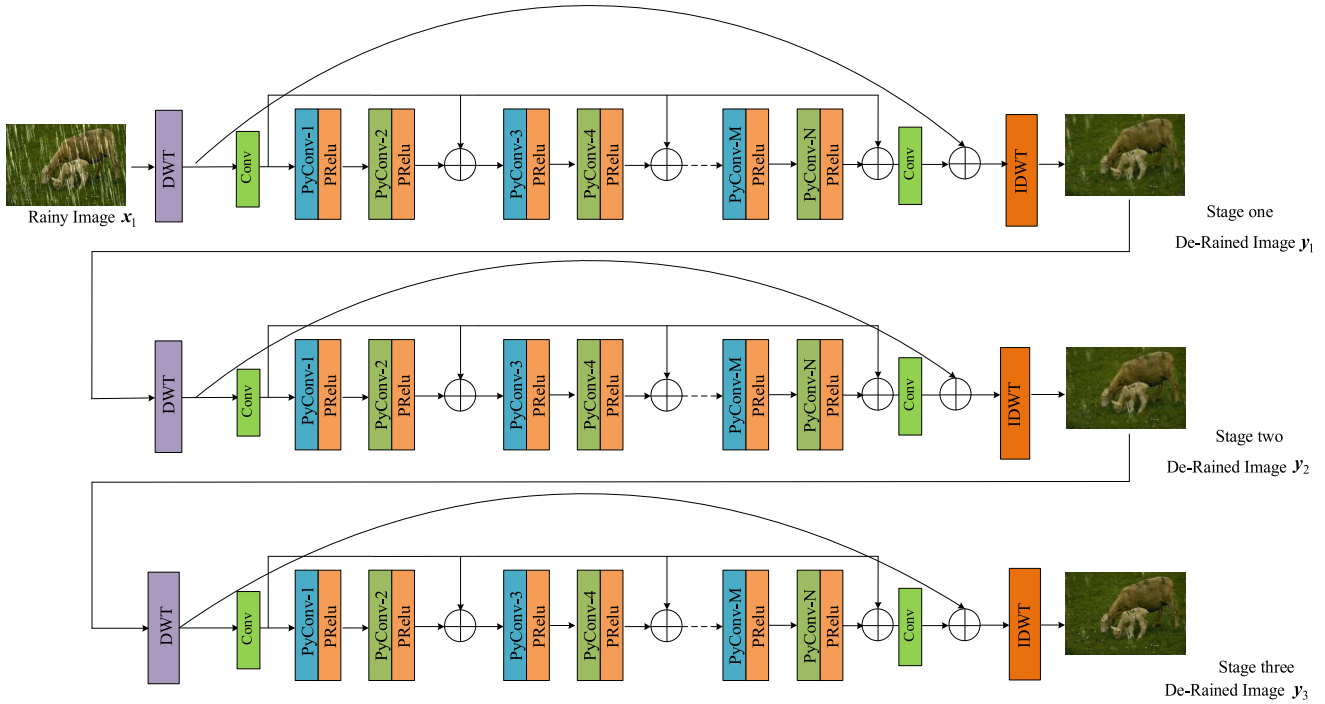


FIGURE 6. The details of the proposed WDRPRN-E.

where $LL(.)$, $LH(.)$, $HL(.)$ and HH represent the wavelet transform process at different frequencies, and

$$L_{ssim}(\theta) = 1 - SSIM(F_\theta(x_i), y_i) \tag{6}$$

where **SSIM**(.) can be calculated by structural similarity [42]. Hence, we can define the total loss as follows:

$$L_{total}(\theta) = L_1(\theta) + \lambda L_{ssim}(\theta) \tag{7}$$

where λ is trade-off parameter between $L_1(\theta)$ and $L_{ssim}(\theta)$.

IV. EXPERIMENTS

A. EXPERIMENT SETTING

This section describes the experimental details and evaluation results on four synthetic datasets and real-world datasets. The image quality metrics of Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity (SSIM) are used to evaluate the de-raining performance. Our WDRPRN is compared with six state-of-the-art methods: two model-driven methods including Gaussian Mixture Model (GMM) [11] and directional Global Sparse Model (UGSM) [14], four data-driven methods including Joint Rain Detection and Removal (JORDER) [5], Deep Detail Network (DDN) [17], Density-aware single Image De-raining using a Multi-stream Dense Network (DID-MDN) [18] and Gated Context Aggregation Network(GCANet) [20]. In the end, we make analysis in the ablation study.

1) IMPLEMENTATION DATASETS

Synthetic Data: Four synthetic datasets are chosen to train our network. The Rain100H [5] dataset has 1800 rainy/ground truth image pairs for training and 200 pairs for testing, in which each rainy image can be synthesized by six different directions and scales rain streak images and corresponding clean images. In the Rain100L [5] dataset, there are 1800 rainy/ground truth image pairs for training and 200 pairs for testing where each rainy image can be synthesized by the six different directions and scales rain streak images and corresponding clean images. Rain12 [11] has only one type of rain streaks.

Training dataset of Rain800 [43] comprises a total of 700 images, of which 500 are randomly selected from UCID dataset and 200 from BSD-500 dataset, and testing dataset of Rain800 contains 100 images, 50 images of which are randomly chosen from UCID dataset and 50 images from BSD-500 dataset.

Real-World Data: Since it is hard to collect clean images corresponding to rainy images in the real world, the real-world rainy images from Rain50 [43] real-world dataset can be tested after training on Rain800 synthetic dataset. Rain50 has 50 real-world rainy images.

2) IMPLEMENTATION DETAILS

In the experiment setup, the number of filters in our network is set to 64, and the depth of recursive module is 16 and the trade-off parameter $\lambda = 1$. In the training and testing stage, the deep learning framework Pytorch is employed to carry out our experiment on Nvidia Tian-V GPU with 12GB memory. We resize the input image to $\mathbf{320} \times \mathbf{480}$, and the size of the training batch is 4. Adam, as an optimizer with initial learning rate of 0.001, implements the decay operation by multiplying

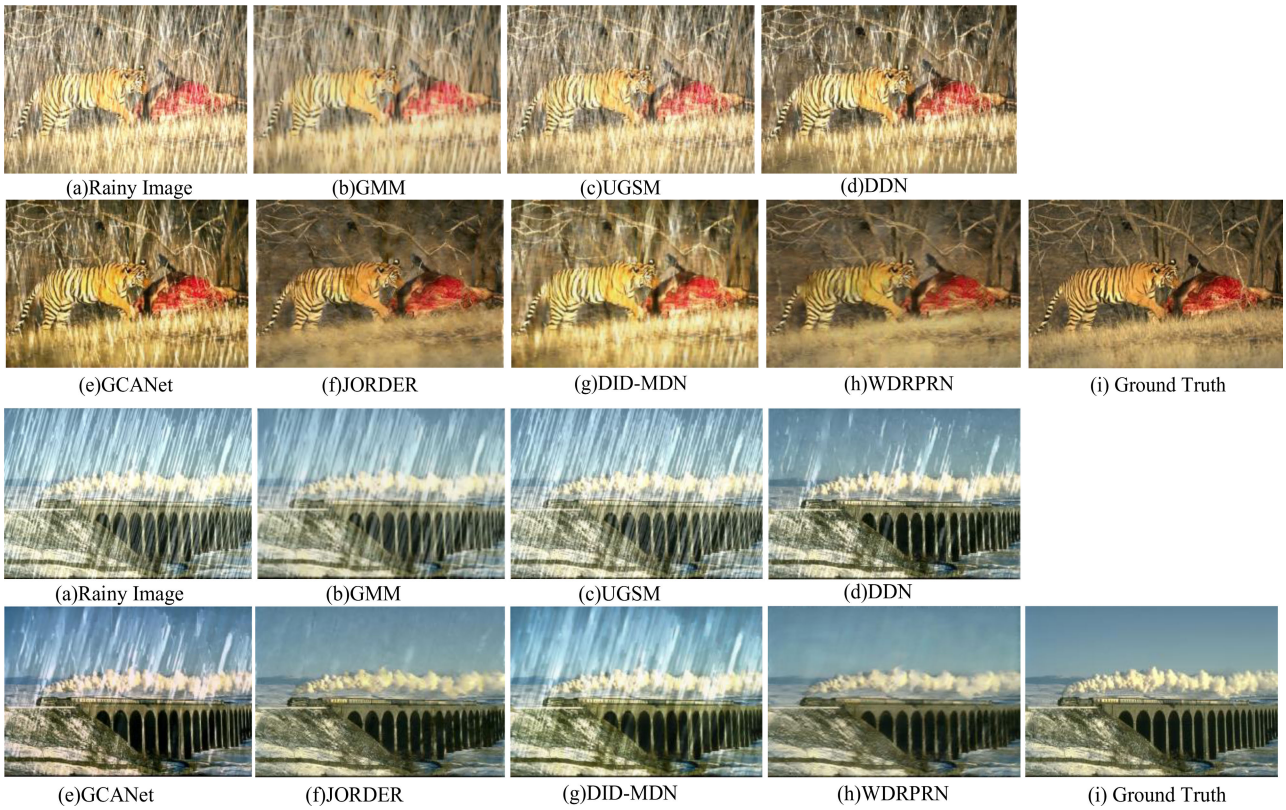


FIGURE 7. Visual comparison of the proposed WDRPRN with state-of-the-art rain removal methods in synthetic rainy images from Rain100H [5].

0.2 in 30, 50 and 80 epochs, and the number of whole training epochs is set to 100.

B. QUANTITATIVE AND QUALITATIVE EVALUATION

1) EVALUATION ON SYNTHETIC DATASETS

Figures 7 to 9 show the visual quality of de-raining performance with rain streaks of different scales and orientations (Rain100H, Rain100L and Rain12). As observed, GMM and UGSM failed to eliminate heavy rain streaks, while GMM can remove light rain streaks and meanwhile blur image details. Although DDN, GCANet and DID-MDN have the ability to remove different types of rain streaks, they also generate tiny artifacts in the de-rained images. JORDER and our WDRPRN have the best de-raining performance. In addition, our method can even restore details and remove most of the rain streaks (See the enlarged part of the red box in Figure 9) than JORDER on Rain12.

To further demonstrate the qualitative improvement on the four synthetic datasets, Table 1 and Table 2 show the quantitative evaluation on PSNR and SSIM metrics. All compared methods have been tested under the same conditions. It can be seen from Table 1 and Table 2 that JORDER is superior to other methods. In terms of image quality metrics, our method is close to JORDER. In spite of the PSNR and SSIM values of our method lower than JORDER on the Rain 100H and Rain 100L, our method is higher than JORDER on Rain800 and Rain12. Furthermore, compared with JORDER and other methods, the proposed method has fewer parameters, and can better reach tradeoff between high de-raining performance and low model parameters, which can be easily deployed in mobile devices.

2) EVALUATION ON REAL-WORLD RAINY IMAGES

To test the WDRPRN de-raining performance in real-world rainy images, our network is trained on Rain800, and then tests the actual de-raining result on Rain50 real-world dataset. Figure 10 shows the visual de-rained effects of two real-world rainy images compared with several state-of-the-art methods including the paper method, and two real-world rainy images are randomly selected from Rain50. It can be clearly seen that UGSM has the worst performance in terms of residual rain streaks, and GMM can remove rain streaks, but the disadvantage is that it makes the de-rained image blur, such as the wall edges, as shown in the second column. Although DDN, GCANet, JORDER and DID-MDN have better de-rained performance, they suffer from unexpected artifacts as shown in the corresponding columns (The details can be better observed via zooming image in the de-rained image).

The paper selects Naturalness Image Quality Evaluator (NIQE) as the no reference image quality evaluation metric of the real-world rain images as shown in Table 3, and a good score is obtained.

TABLE 1. Average PSNR and SSIM values in four synthetic datasets. Red denotes the best result and blue shows the second best result.

	GMM		UGSM		DDN		GCANet	
Metrics	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Rain100H	15.01	0.430	14.88	0.471	21.25	0.740	22.19	0.783
Rain100L	28.43	0.861	24.10	0.837	32.01	0.930	32.03	0.921
Rain800	22.03	0.711	22.97	0.747	22.02	0.742	22.28	0.749
Rain12	32.02	0.911	30.23	0.882	31.76	0.943	30.77	0.902
Parameters	–		–		57369		702818	

TABLE 2. Average PSNR and SSIM values in four synthetic datasets. Red denotes the best result and blue shows the second best result.

	JORDER		DID-MDN		WDRPRN		WDRPRN-E	
Metrics	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Rain100H	22.94	0.811	23.51	0.797	22.87	0.787	23.53	0.802
Rain100L	36.11	0.969	25.78	0.861	34.02	0.943	35.89	0.959
Rain800	22.18	0.757	21.34	0.795	22.66	0.781	23.17	0.798
Rain12	33.92	0.951	33.06	0.917	33.89	0.944	33.95	0.952
Parameters	369792		372839		68044		68044	

FIGURE 8. Visual comparison of our WDRPRN with progressive rain removal methods in synthetic rainy images from Rain100L [5].

The WDRPRN method proposed in the paper can effectively remove most of the rain streaks and restore the background image structure in the covered region.

C. ABLATION STUDY

1) RUNNING TIME AND CONVERGENCE

The running time and model parameters of several state-of-the-art methods are conducted as shown in Table 4. The GMM(CPU), UGSM(CPU), DDN(GPU) and JORDER (GPU) are implemented in MATLAB, and other methods and our method are implemented on Pytorch (GPU). Considering that the optimization process of GMM is time-consuming, we only compare the running time of other methods under 320×480 image. The model-driven method UGSM runs the slowest, while data-driven methods JORDER and DID-MDN run faster. Even though the test time of our proposed method with fewer model parameters is not the fastest, it can be easily used in mobile devices with better rain removal performance and fewer model parameters as shown in Table 1 and Table 2.

TABLE 3. Average NIQE value in real-world rainy images from Rain50 real-world dataset.

Methods	GMM	UGSM	DDN	GCANet	JORDER	DID-MDN	WDRPRN	WDRPRN-E
NIQE	5.98	5.67	5.04	5.11	4.54	4.43	4.42	4.37

TABLE 4. Running time and parameters of different methods (seconds).

320×480	UGSM	DDN	GCANet	JORDER	DID-MDN	WDRPRN
Running time	0.98	0.56	0.49	0.13	0.37	0.43
Parameters	-	57369	702818	369792	372839	68044

FIGURE 9. Visual comparison of the WDRPRN with advanced rain removal methods in synthetic rainy images from Rain12 [11].

To present the training process of WDRPRN, we show the convergence of training loss function and the improvement process of PSNR and SSIM throughout the epoch. It can be observed from Figure 11 that WDRPRN converges quickly, and meanwhile the PSNR and SSIM values reach a high level in the Rain100L dataset, which can reflect the rationality and effectiveness of the training setting.

2) ABLATION STUDY ON ENHANCED WDRPRN

In order to explore de-raining performance in different stages, we conduct an ablation study. As shown in Figure 12, with the increase of the number of single WDRPRN network, the de-raining visual result is much better. To objectively evaluate the effect of different stages on Rain100L, Table 5 shows

TABLE 5. Average PSNR and SSIM values of de-raining result at different stages on Rain100L.

Metrics	WDRPRN-1	WDRPRN-2	WDRPRN-3
PSNR	34.20	35.36	35.89
SSIM	0.943	0.950	0.959

that the multiple WDRPRNs can better remove rain streaks. Considering the limited memory, our enhanced WDRPRN is a three-stage recursive network.

3) ABLATION STUDY ON LOSS

To investigate the de-raining result of different loss strategies (MSE loss and MAE+SSIM loss), we conduct the ablation loss study. Throughout the experiment, the reason

FIGURE 10. Visual comparison of the proposed WDRPRN with state-of-the-art rain removal methods in real-world rainy images from Rain50 [43].

why we did not adopt the MSE loss function is that it may over penalize the pixel value error, which can generate the blurred image. In order to keep the structural details of local patch image the same between de-raining image and clean image, the MAE loss and SSIM loss are accepted. The quantitative results concerning on different loss functions are tabulated as observed in Table 6. Compared with the de-raining performance by MSE loss, the image quality metrics PSNR and SSIM are obviously improved by MAE+SSIM loss. Figure 13 shows that our experiment setup loss function can better remove rain streaks and preserve more details.

TABLE 6. Average PSNR and SSIM values of de-raining effect under different loss function strategies onRain100L.

Rain100L	Loss function strategies	
	MSE	MAE+SSIM
PSNR	31.18	35.89
SSIM	0.912	0.959

4) ANALYSIS OF RAIN REMOVAL ON DIFFERENT FREQUENCIES

In order to further investigate the advantage of de-raining result in the different frequency domains, we add the visualization contrast between four sub-bands of de-raining image

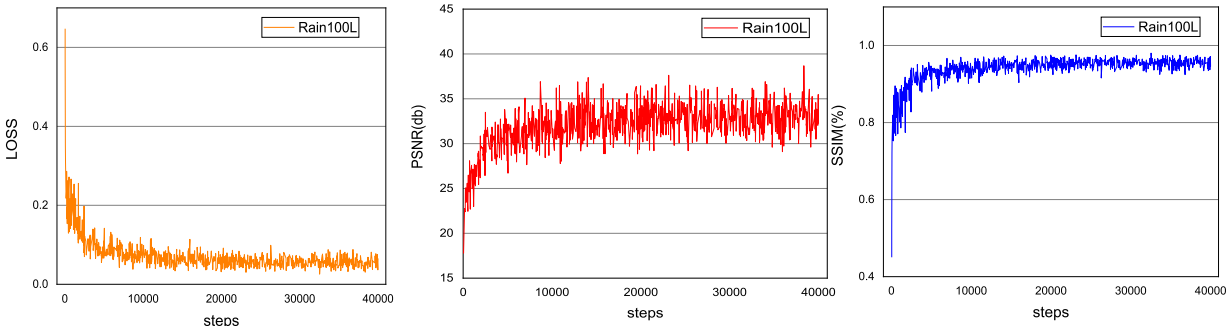


FIGURE 11. Training convergence analysis on LOSS, PSNR and SSIM of WDRPRN.

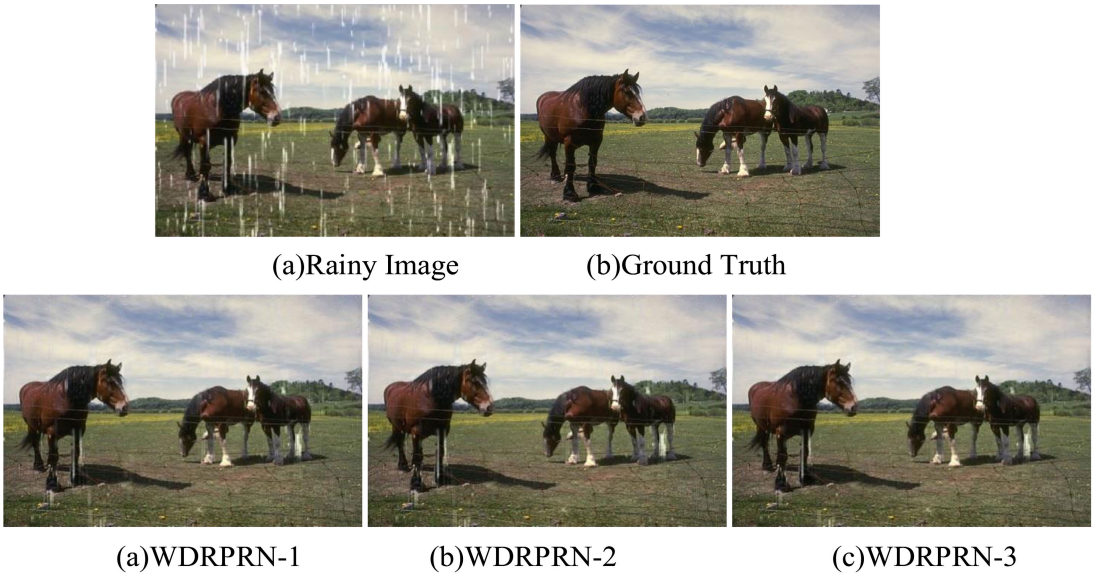


FIGURE 12. The visual effect of de-raining performance in different stages. The number E in WDRPRN-E indicates the E-th stage WDRPRN.

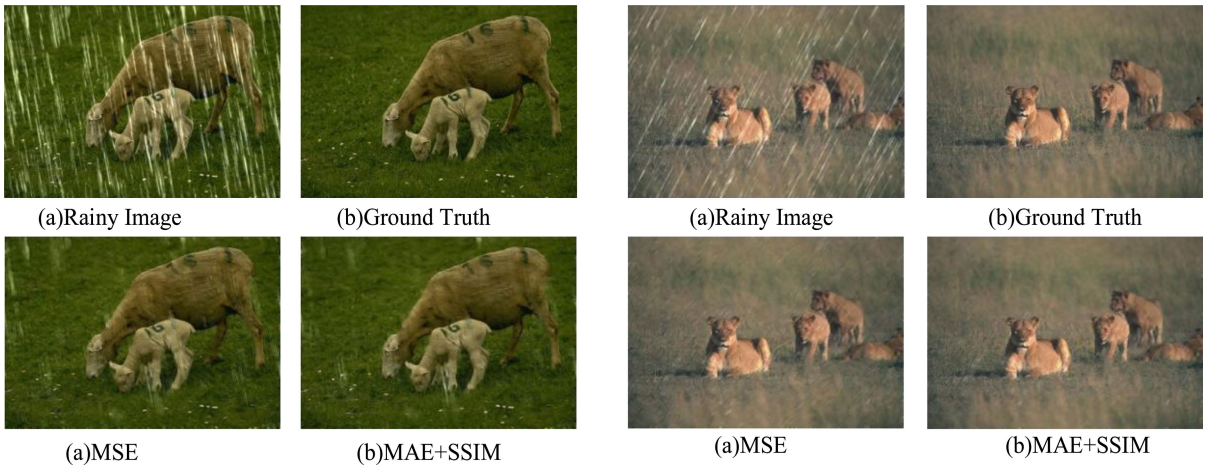
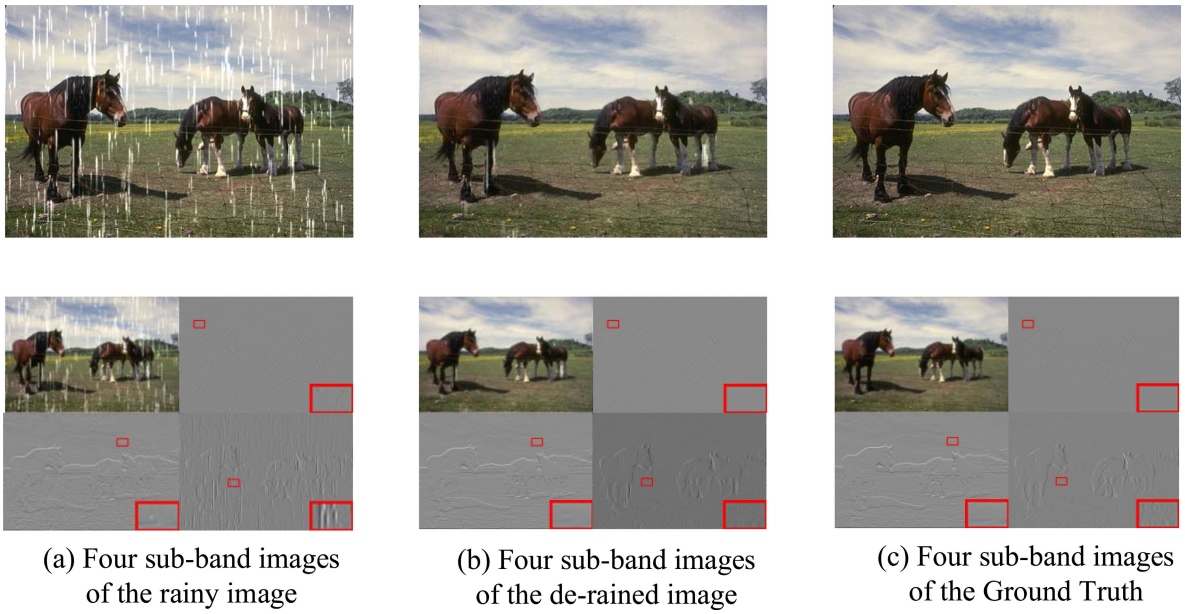


FIGURE 13. De-rained results by different losses.

and ground truth by using wavelet transform. Figure 14 illustrates the visual effect of rain removal, where the rain streaks can well be removed in different sub-band images, and the details, which are distorted, can be better restored. Moreover, the edge structure of rain-free region is maintained well, which is attributed to that wavelet transform can decompose the image signal into multi-scale sub-bands. We can restore the rainy image according to the characteristics of rain streaks in different frequencies.



(a) Four sub-band images of the rainy image

(b) Four sub-band images of the de-rained image

(c) Four sub-band images of the Ground Truth

FIGURE 14. The visual result of rain removal and detail restoration in different frequencies by wavelet transform.

V. CONCLUSION

In this paper, we have proposed a light wavelet deep recursive pyramid convolution residual network (WDRPRN). In comparison with the existing approaches, which failed to reach the tradeoff between better de-raining performance and fewer model parameters, we explored the de-raining method with wavelet transform, which can achieve satisfactory performance. The WDRPRN method can take advantage of the characteristics of rain streaks at different frequencies, and then remove different types of rain streaks by embedding deep learning approach. Comparative experiments and ablation studies were conducted to illustrate the well de-rained performance under different baseline settings.

In spite of the superior performance achieved by our method, there are still shortcomings that wavelet transform incorporated into the deep learning method is plain model, which may not maximize the de-raining performance. Therefore, we will pay more attention to establishing effective rain removal methods in future studies.

REFERENCES

[1] J. Mao, T. Xiao, Y. Jiang, and Z. Cao, "What can help pedestrian detection?" in *Proc. CVPR*, Jul. 2017, pp. 3127–3136.

[2] K. Li, Y. Kong, and Y. Fu, "Multi-stream deep similarity learning networks for visual tracking," in *Proc. IJCAI*, Aug. 2017, pp. 2166–2172.

[3] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 11, pp. 1254–1259, Nov. 1998.

[4] G. Li, Y. Xie, L. Lin, and Y. Yu, "Instance-level salient object segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 247–256.

[5] 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 Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 1685–1694.

[6] Y. Luo, Y. Xu, and H. Ji, "Removing rain from a single image via discriminative sparse coding," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 3397–3405.

[7] L. Zhu, C.-W. Fu, D. Lischinski, and P.-A. Heng, "Joint bi-layer optimization for single-image rain streak removal," in *Proc. ICCV*, Oct. 2017, pp. 2526–2534.

[8] Y. Wang, S. Liu, C. Chen, and B. Zeng, "A hierarchical approach for rain or snow removing in a single color image," *IEEE Trans. Image Process.*, vol. 26, no. 8, pp. 3936–3950, Aug. 2017.

[9] Y.-L. Chen and C.-T. Hsu, "A generalized low-rank appearance model for spatio-temporally correlated rain streaks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2013, pp. 1968–1975.

[10] H. Zhang and V. M. Patel, "Convolutional sparse and low-rank coding-based rain streak removal," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2017, pp. 1259–1267.

[11] Y. Li, R. T. Tan, X. Guo, J. Lu, and M. S. Brown, "Rain streak removal using layer priors," in *Proc. CVPR*, Jun. 2016, pp. 2736–2744.

[12] Y. Li, R. T. Tan, X. Guo, J. Lu, and M. S. Brown, "Single image rain streak decomposition using layer priors," *IEEE Trans. Image Process.*, vol. 26, no. 8, pp. 3874–3885, Aug. 2017.

[13] L.-W. Kang, C.-W. Lin, and Y.-H. Fu, "Automatic single-image-based rain streaks removal via image decomposition," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1742–1755, Apr. 2012.

[14] L.-J. Deng, T.-Z. Huang, X.-L. Zhao, and T.-X. Jiang, "A directional global sparse model for single image rain removal," *Appl. Math. Model.*, vol. 59, pp. 662–679, Jul. 2018.

[15] J. Xu, W. Zhao, P. Liu, and X. Tang, "Removing rain and snow in a single image using guided filter," in *Proc. IEEE Int. Conf. Comput. Sci. Autom. Eng. (CSAE)*, May 2012, pp. 304–307.

[16] Q. Zhu, L. Shao, P. A. Heng, and X. Li, "A novel rain detection and removal approach using guided filtering and formation modeling," in *Proc. IEEE Int. Conf. Robot. Biomimetics (ROBIO)*, Shenzhen, China, Dec. 2013, pp. 563–567.

[17] 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. CVPR*, Jul. 2017, pp. 3855–3863.

[18] 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.*, Salt Lake City, UT, USA, Jun. 2018, pp. 695–704.

[19] X. Li, J. Wu, Z. Lin, H. Liu, and H. Zha, "Recurrent squeeze-and-excitation context aggregation net for single image de-raining," in *Proc. 15th Eur. Conf. Comput. Vis. (ECCV)*, Munich, Germany, Sep. 2018, pp. 262–277.

[20] D. Chen, M. He, Q. Fan, J. Liao, L. Zhang, D. Hou, L. Yuan, and G. Hua, "Gated context aggregation network for image dehazing and deraining," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Waikoloa Village, HI, USA, Jan. 2019, pp. 1375–1383.
[21] I. Cosmin Duta, L. Liu, F. Zhu, and L. Shao, "Pyramidal convolution: Rethinking convolutional neural networks for visual recognition," 2020, *arXiv:2006.11538*. [Online]. Available: http://arxiv.org/abs/2006.11538
[22] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, Apr. 2018.
[23] Y. Yang and H. Lu, "Single image deraining using a recurrent multi-scale aggregation and enhancement network," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2019, pp. 1378–1383.
[24] X. Fu, B. Liang, Y. Huang, X. Ding, and J. Paisley, "Lightweight pyramid networks for image deraining," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 6, pp. 1794–1807, Jun. 2020.
[25] B.-H. Chen, S.-C. Huang, and S.-Y. Kuo, "Error-optimized sparse representation for single image rain removal," *IEEE Trans. Ind. Electron.*, vol. 64, no. 8, pp. 6573–6581, Aug. 2017.
[26] V. Jain and H. S. Seung, "Natural image denoising with convolutional networks," in *Proc. NIPS*, 2008, pp. 769–776.
[27] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.
[28] M. Hradiš, J. Kotera, P. Zemcík, and F. Šroubek, "Convolutional neural networks for direct text deblurring," in *Proc. BMVC*, 2015, p. 2.
[29] S. Nah, T. H. Kim, and K. M. Lee, "Deep multi-scale convolutional neural network for dynamic scene deblurring," in *Proc. CVPR*, Jul. 2017, pp. 3883–3891.
[30] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in *Proc. ECCV*, 2014, pp. 184–199.
[31] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb. 2016.
[32] W. Cai and Z. Wei, "PiiGAN: Generative adversarial networks for pluralistic image inpainting," *IEEE Access*, vol. 8, pp. 48451–48463, 2020.
[33] Z.-L. Yang, X.-Q. Guo, Z.-M. Chen, Y.-F. Huang, and Y.-J. Zhang, "RNN-stega: Linguistic steganography based on recurrent neural networks," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 5, pp. 1280–1295, May 2019.
[34] X. Hu, C.-W. Fu, L. Zhu, and P.-A. Heng, "Depth-attentional features for single-image rain removal," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 8022–8031.
[35] T. Guo, H. S. Mousavi, T. H. Vu, and V. Monga, "Deep wavelet prediction for image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 1100–1109.
[36] H. Huang, R. He, Z. Sun, and T. Tan, "Wavelet-SRNet: A wavelet-based CNN for multi-scale face super resolution," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 1698–1706.
[37] P. Liu, H. Zhang, K. Zhang, L. Lin, and W. Zuo, "Multi-level wavelet-CNN for image restoration," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2018, pp. 886–895.
[38] W. Yang, J. Liu, S. Yang, and Z. Guo, "Scale-free single image deraining via visibility-enhanced recurrent wavelet learning," *IEEE Trans. Image Process.*, vol. 28, no. 6, pp. 2948–2961, Jun. 2019.
[39] H.-H. Yang and Y. Fu, "Wavelet U-net and the chromatic adaptation transform for single image dehazing," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Taipei, Taiwan, Sep. 2019, pp. 2736–2740.
[40] Y. Tai, J. Yang, and X. Liu, "Image super-resolution via deep recursive residual network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 2790–2798.
[41] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Santiago, Chile, Dec. 2015, pp. 1026–1034.
[42] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[43] H. Zhang, V. Sindagi, and V. M. Patel, "Image de-raining using a conditional generative adversarial network," *IEEE Trans. Circuits Syst. Video Technol.*, early access, Jun. 3, 2019. [Online]. Available: http://arxiv.org/abs/1701.05957, doi: 10.1109/TCSVT.2019.2920407.

TIAN TIAN GONG received the master's degree in mathematics and information science from North Minzu University, China. She is currently researching wavelet analysis, image processing, and robust optimization.

JUN SHENG WANG received the M.S. degree from Nanchang Hangkong University, China. He is currently pursuing the Ph.D. degree with the Nanjing University of Science and Technology, China. His current research interests include computer vision, deep learning, and image processing.