IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 29 January 2024, accepted 24 February 2024, date of publication 29 February 2024, date of current version 14 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3371484

RESEARCH ARTICLE

LPN-IDD: A Lightweight Pyramid Network for Image Deraining and Detection

KAINAT BABAR[1], MUHAMMAD USMAN YASEEN[2], AHMAD SAMI AL-SHAMAYLEH[3], MUHAMMAD IMRAN[2], ABDULLAH HUSSEIN AL-GHUSHAMI[4], (Senior Member, IEEE), AND ADNAN AKHUNZADA[5], (Senior Member, IEEE)

[1]Department of Computer Science, Shifa Tameer-e-Millat University, Islamabad 45550, Pakistan
[2]Department of Computer Science, COMSATS University Islamabad (CUI), Islamabad 45550, Pakistan
[3]Department of Data Science and Artificial Intelligence, Faculty of Information Technology, Al-Ahliyya Amman University, Amman 19328, Jordan
[4]Department of Information Technology, Community College of Qatar, Doha, Qatar
[5]College of Computing and Information Technology, University of Doha for Science and Technology, Doha, Qatar

Corresponding authors: Abdullah Hussein Al-Ghushami (abdullah.alghushami@ccq.edu.qa) and Muhammad Imran (mimran@comsats.edu.pk)

The open access funding is provided by Qatar National Library (QNL), Doha, Qatar. We appreciate QNL for supporting this research work.

ABSTRACT One of the challenges in image processing involves single image deraining (SID). Existing methods do not exploit images at multi-scales and often overlook spatial and channel information. These methods fail to account for different rain conditions, leading to challenges in effectively removing a wide range of rain streak patterns, such as diverse directional or dense streaks. Moreover, they often result in the loss of fine texture details or a blurred background in the process of eliminating rain streaks from images captured in heavy rain. Mostly state-of-the-art (SOTA) deraining models achieve higher performance in removing rain from rainy images but at the expense of a high number of parameters, which results in computational complexity and memory requirements. Furthermore, they also do not consider high-level visioned evaluation metrics to perform deep evaluations of the proposed models. In this paper, a simple lightweight network with few parameters and relatively shallow depth is proposed by fusing the traditional Gaussian-Laplacian pyramid technique with the attention module. We propose a novel attention-based lightweight pyramid network for image deraining and detection (LPN-IDD) to achieve better deraining performance. The proposed model includes a dual attention module integrated with the Gaussian-Laplacian pyramids network. In LPN-IDD, different levels of Laplacian pyramids can extract multi-scale features to adapt to different shapes and types of rain streaks. Residual and recursive blocks are used in each subnetwork along with dual attention blocks to resist the occlusion or texture feature while suppressing unnecessary features. Extensive experimentation performed on the SID synthetic and real-world datasets demonstrate the effectiveness of the proposed model in image deraining tasks. For deep evaluation of the proposed methodology, object detection models including faster-RCNN, YOLO-V3, YOLO-V7 are used along with full reference evaluation metrics.

INDEX TERMS Channel attention, dual attention, feature enhancement, Laplacian pyramid network, object detection, recursive learning, residual learning, single image deraining, spatial attention.

I. INTRODUCTION

Rain is a frequent weather occurrence that has a significant impact on our everyday experiences. When it rains, photographs may suffer from decreased visibility caused by rain-obscured views or streaks. Rain has a complex atmospheric process, causing rain droplets to stick to camera lenses or windscreens, resulting in blurred and distorted images. Rain masks can take on different forms, such as sparse rain streaks, rain mist, or randomly dispersed droplets, depending on the circumstances. The removal of rain streaks, droplets, or mist from rainy images has gained

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

FIGURE 1. An example of single Image deraining (SID): (a) synthetic rainy image (b) original ground truth image.

importance because these degradations significantly impact the effectiveness of vision systems that rely on clear, high-resolution images. The complexity and unpredictability of rain, which can vary in size, density, and direction, makes this a difficult computer vision task. To address this issue in image processing and computer vision, single image rain removal methods or algorithms, also known as single image deraining (SID), are used and have become increasingly significant in various fields.

Single image deraining aims to recover a clear background image that has been marred by raindrops or rain streaks. There are several SID techniques available for removing rain from a single image. These methods can be broadly classified into two categories: model-driven methods and deep learning-based methods. Model-driven methods use physical models to simulate the effects of rain and remove it from the images. On the other hand, deep learning-based data-driven methods rely on sophisticated algorithms based on deep learning to learn and understand rain patterns from a large dataset of images. They then use this knowledge to remove rain from new images. Recent studies revealed that the image deraining problem is still facing dissatisfying results due to the requirement of an enormous number of parameters for handling complex heavy rainy conditions, Although existing SID methods achieve remarkable results their performance can be further improved.

Detecting rain streaks' size and shape during dynamic weather conditions, like rainy days, is challenging due to limitations in traditional convolution methods, particularly their convolution kernel size. Existing deraining networks improve performance by increasing convolutional neural network (CNN) width and depth but often overlook spatial and channel information, limiting feature utilization. These methods also tend to employ complex structural changes to remove rain streaks. These methods are computationally intensive, and they require substantial memory resources during the execution of the algorithms. These factors can make the rain removal process relatively slow and resource-consuming. This limitation hampers their efficiency in effectively deraining images in real-time scenarios, which is particularly disadvantageous for applications requiring quick processing, such as outdoor surveillance systems or real-time image enhancement.

The most extensively utilized image-deraining methods are data-driven-based pyramid methods. These methods do not consider various rain conditions and types and face trouble while eliminating diverse directional or dense streaks ultimately producing over-smoothed texture details or blurry background images after extracting rain streaks. They give equal importance to all extracted features, whereas few features are more important for effective restoration. While attention-based deraining methods instead of emphasizing the removal of background images or texture details concentrate on learning rain streaks. In addition, most of the existing methods only consider full reference evaluation metrics for comparison (structural similarity index (SSIM), peak signal-to-noise ratio (PSNR)), and some consider no-reference metrics (natural image quality evaluator (NIQE), subjective structural image quality (SSIQ)) but they do not consider high-level visioned evaluation metrics to perform deep evaluations of the proposed model. Researchers are continuously working to improve the efficiency of these methods to make them more practical for real-world applications, because single image deraining may improve the quality of images as seen by humans as well as numerous computer vision applications, including intelligent vehicles [1], object recognition [2], video processing tasks [3], outdoor monitoring systems [4], and auto-driving [5], etc.

To address these issues we suggest a Laplacian pyramid-based lightweight network for image deraining and detection (LPN-IDD) with fewer parameters. Instead of creating a complicated network structure, we leverage domain-specific knowledge to streamline the learning procedure. The proposed system model is composed of Laplacian, and Gaussian pyramids to decompose images at multiple scales to extract image features at every scale. After getting these features, they are passed to independent subnetworks for clean Gaussian pyramid reconstruction. Based on the unique physical characteristics of each subnetwork, a customized loss function is selected for training. The multi-task supervision is used during the whole training period. The base of the reconstructed Gaussian pyramid represents the final derained image. The main contributions of the proposed network are summarized below.

- A simple, lightweight network structure with few parameters and relatively shallow depth is proposed by fusing the traditional Gaussian-Laplacian pyramid technique with the attention module. Multi-scale local and global feature fusion is performed.
- To extract the key features that may be hidden due to the complex rainy conditions from images, a dual attention module integrated with channel and spatial attention is used.
- Residual and recursive blocks are employed in each subnetwork along with dual attention blocks to resist the occlusion or texture feature while suppressing unnecessary features. They focus on the most important

features of rainy images and at last result in a non-blurry and smooth background.
- Extensive experimentation is performed on the SID training set to evaluate the deraining performance of the proposed model. For deep evaluation of the proposed methodology, object detection models are used along with the full reference evaluation metrics.

The rest of the paper is organised as follows: Section II provides a literature review for image deraining approaches. Section III presents the proposed model. Results and analysis are discussed in Section IV. Section V presents the conclusion and Section VI provides future work extensions of the proposed work.

II. RELATED WORK

Within the context of addressing rain removal, there are two distinct categories: rain removal from videos and rain removal from a single image. In this case, our focus will be specifically on single image deraining methods, which aim to remove rain artifacts from individual images rather than entire video sequences. Currently, two distinct approaches are being used to deal with single image deraining problem: one is model-driven approaches and the second one is data-driven approaches.

A. MODEL-DRIVEN IMAGE DERAINING METHODS

The model-driven solutions place a special emphasis on designing effective algorithms that address optimization challenges by skillfully utilizing and encoding the relevant physical attributes of rain, along with a deep understanding of the surrounding environment. Model-based rain removal Approaches include filter-based [6], [7], [8], [9], [10], [11], [12] traditional prior-based solutions [13], [14], [15], [16], GMM based [17], [18], sparce coding based methods [19], [20], [21], [22], [23]. In [6] and [7], a guided filter is proposed for snow-rain removal, involving dark channel prediction and guided filter refinement. Furthermore, [8] introduced a guided L0 smoothing filter for rain removal in images. In order to remove line pattern noise from images using Fourier transformation and separate noise and image estimation, [9] offers a changed low-rank model. For removing rain streaks from video cycles [10] models spatio-temporally coupled rain streaks. in [11] a texture-based approach used to identify and eliminate rain or snow patches from images. In [12] a SID network with guided learning is proposed via three sub-networks and uses multi-scale residual blocks (MSRBs) for better deraining by exploiting multi-scale data.

In [13], a layer decomposition technique is used to address how rain streaks affect image visibility and computer vision algorithms. By dividing images into rain-free backdrop (B) and rain-containing (R) layers, [14] introduces a novel single image streak extraction technique. This technique is made possible by an iterative optimization approach [15]. An unrolling technique is employed by [16] to integrate data-dependent network topology within existing iterations and introduce a learning bi-level layer priors method for effective rain streak removal assessment and efficiency. In [17], a multi-detail unit-based priori-guided model (PGRAN) is proposed for single image deraining. By utilizing Gaussian mixture models learned from small patches, [18] introduces a structural residue phase to enhance image quality through background noise removal. In [19], a joint convolutional analysis and synthesis sparse representation model JCAS is suggested. Reference [20] proposes a global sparse approach to model and estimate rain streaks across diverse directions. Sparse coding-based algorithms [21], [22], [23] utilize clear images for rain removal while overcoming traditional batch-model learning limitations by exploiting structural similarity.

After an extensive review of model-driven deraining approaches, it is evident that model-driven methods for single image deraining struggle with precise rain layer separation from backgrounds and complex rain streaks due to fixed hand-crafted attributes, which are inadequate for the diverse array of rain streak types.

B. DATA-DRIVEN IMAGE DERAINING METHODS

Due to recent rapid advancements in this field, deep learning has substantially surpassed conventional deraining techniques in single image rain removal and other image restoration tasks. Data-driven based rain removal approaches include CNN based [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], Attention based [39], [40], [41], [42] GANs or Lightweight pyramid based [43], [44], [45], [46], [47], [48], [49], [50], [51], [52], [53], [54], [55], [56], [57], [58], [59]. Reference [24] introduces a stereo-based rainy dataset and the paired rain removal network (PRRNet) for monocular deraining, addressing semantic segmentation and scene deraining challenges. Reference [25] introduces the conditional variational image deraining (CVID) network, which employs spatial density estimation (SDE) and channel-wise (CW) deraining methods to enhance deraining performance. Reference [26] proposes ResGuideNet, a novel single image deraining architecture employing cascaded blocks guided by residuals for enhanced rain removal performance and aesthetically pleasing results. Reference [27] utilizes phase congruence features in rainy videos to detect rain streaks, leveraging inter-frame feature variations for removal and refining results through minimization operations and proposed filters.

In [28], a deep detail network (DDN) is introduced by merging deep learning and image processing with detail enhancement (DES) and rain removal sub-networks (RRS). Reference [29] presents a progressive image deraining network (PReNet) for single image deraining, including feature extraction and deraining modules. Reference [30] utilizes SNet, an encoder-decoder CNN, for rain streak transmission learning with VNet for multi-scale predictions

and ANet for atmospheric light estimation. In [31], a DRD-Net was proposed to address lost details in derained images, employing rain residual and detail restoration networks for effective detail recovery. The [32] restructures network layers into residual functions, demonstrating improved optimization and accuracy with deeper residual networks. In [33], DerainNet exploits image detail (high-pass) layers instead of the entire image domain, demonstrating strong real-world performance despite artificial rain training.

For rain-streak removal, [34] presents DID-MDN, utilizing densely linked convolutional neural networks to process multiple data streams. A deep detail network inspired by ResNet [28] and an automatic rain density-aware network [35] is proposed for rain streak removal. Scale variation in vision tasks is a common issue, affecting both low and high-level processes. Existing CNN-based methods often overlook crucial details due to single-scale analysis. They frequently ignore spatial and channel information, which limits the use of features, and instead often prioritize expanding CNN dimensions to boost deraining performance. To address this issue, multi-scale techniques or feature pyramids are used, extracting local and global information to enhance image restoration.

As attention models focus on important image areas, commonly applied in high-level vision tasks such as super-resolution [36], object detection [37], image captioning [38], and also in single image deraining due to their remarkable feature selection capability. The [39] employs depth-guided attention for SID through residual map regression. In [40], a spatial attention network is introduced for rain removal, and [41] proposes DARGNet, a dual attention model for dense rain streak removal. Furthermore, in [42], a robust attention deraining network (RadNet) introduces a universal SID model with robustness and enhanced real scenario data effectiveness. Attention models display promising results in rain streak removal, yet struggle with dense or large droplets. Moreover, their computational complexity hampers real-time applicability.

Lightweight pyramid networks are vital for efficient low-level image processing and computer vision tasks. Despite deep-CNNs' success in deraining, their parameter complexity is addressed by lightweight pyramids, utilized in various tasks like image deblurring [43] and lightweight super-resolution [44]. In [45], a lightweight pyramid network LPNet is introduced for SID using Gaussian-Laplacian decomposition and recursive-residual networks. The [46] proposes a rain streak removal strategy through a residual multi-scale pyramid approach, augmented by a multi-scale kernel selection network (MSKSN). In [47], a multi-sub-network structure is introduced with cross-scale fusion using a gated recurrent unit and inner-scale dense block. The [45] suggests a semi-supervised Gaussian Process architecture, trained on synthetic and unlabeled real-world data. In [48] and [49], a multi-scale progressive fusion network (MSPFN) and a multi-receptive field aggregation network (MRFAN) based on self-supervised, aggregation and multi-receptive field feature extraction blocks are proposed for image deraining.

In [50], a cross-domain collaborative learning model based on a multi-scale attention residual block (MSARB) is proposed. In [51] and [52], a recurrent wavelet structure-preserving residual network (RWSRNet) and recurrent context-aware multi-stage network (ReCMN) were proposed for image deraining. In [53] and [54], an end-to-end recurrent multilevel residual and global attention network and multi-scale context information and attention network are proposed. In [55], a multi-scale aggregation residual channel attention fusion network (MARCAFNet) is proposed. A lightweight semi-supervised network (LSNet) for single image deraining is proposed by [56]. Based on the lightweight pyramid deraining (LPD) block and recursive mechanism, [57] proposes a lightweight recursive pyramid network (LRP-Net). Furthermore, [58], [59] proposes multi-scale learning and attention-based densely connected mechanisms (CMADNet) for image deraining and shows when networks expand and become more intricate, the demand for additional storage capacity grows. However, lightweight pyramid models underperform in heavy rain, affecting complex single image deraining. Also, many deraining techniques excel in rain streak removal, they struggle with diverse rain types and heavy rainy conditions due to reliance on synthetic datasets.

III. PROPOSED MODEL

This section describes a comprehensive explanation of the proposed approach, a lightweight pyramid network for image derain and detection (LPN-IDD), which is designed for effectively removing rain artifacts from individual images. The overall architecture of the proposed network is presented in Fig. 2 and its parameters are provided in Table 1. Subsequently, the intricacies of the subnetwork structure including Laplacian and Gaussian pyramids, attention and feature enhancement block as well as recursive and residual learning are explored. We also elaborate on the various loss functions employed within the LPN-IDD framework. Algorithm 1 provides an overview of the entire process of the proposed lightweight pyramid deraining network.

A. THE FRAMEWORK OF LPN-IDD

Fig. 1 demonstrates a restored rain-free image from a rainy one using our model in two main stages: feature extraction and reconstruction. Shallow features are extracted initially, followed by deeper feature extraction.

1) LAPLACIAN PYRAMIDS

A rainy image X is decomposed into its Laplacian pyramids. The decomposed image comprises of a set of images L with

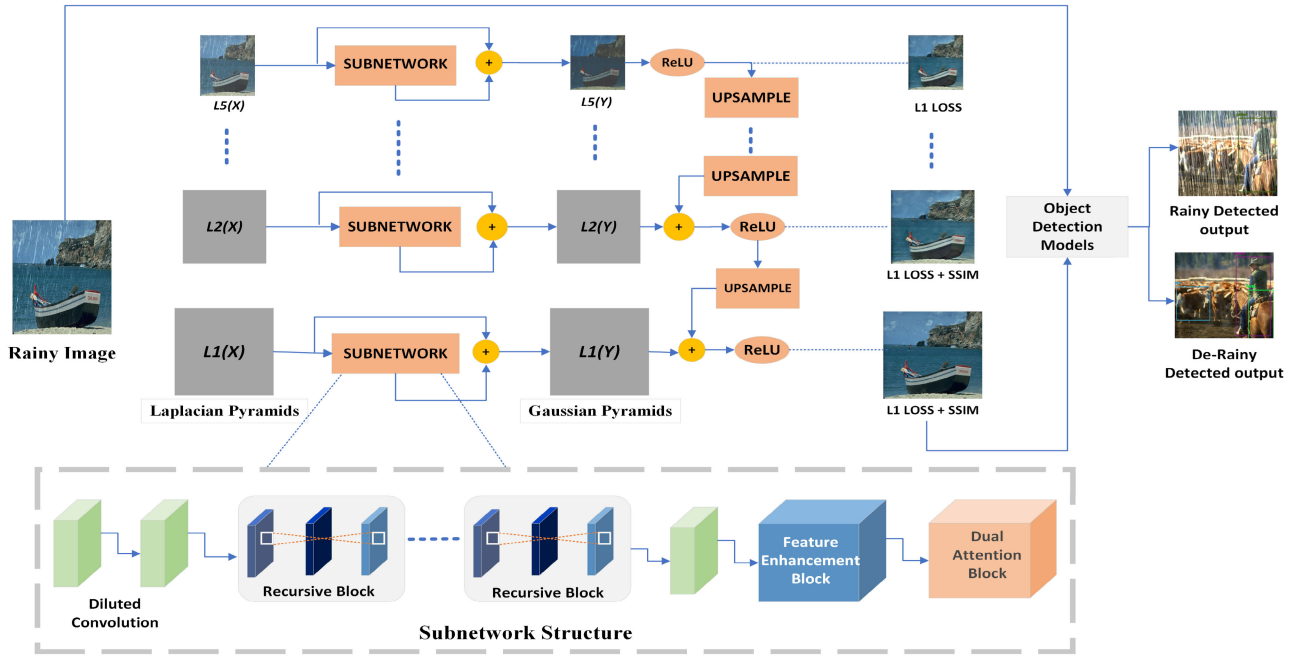


FIGURE 2. Proposed lightweight pyramid network for image draining and detection (LPN-IDD) architecture.

Algorithm 1 LPN-IDD: Lightweight Pyramid Network for Image Deraining and Detection

Input: Training data $TD = [R^i, G^i]$.
Output: Derained image Y.
Initialize: Network structure, hyper-parameters.
Deraining Pipeline:
for each training epoch **do**
  for each iteration **do**
Step 1: Forward pass to calculate predicted images.
Step 2: Compute Loss function

$$L = \frac{1}{2}\sum_{i=1}^{M}\left(\sum_{i=1}^{N} L^{l_1}(G_n(Y^i), G_n(Y^i_{GT})) + \sum_{n=1}^{2} L^{l_{SSIM}}(G_n(Y^i), G_n(Y^i_{GT}))\right)$$

Step 3: Update parameter θ_d using Adam optimizer

$$\theta_d = \text{Adam}(\nabla\theta_d(B), \theta_d)$$

Detection Pipeline:
  for each image i in X **do**
Create annotation files for X and derained images Y.
Apply pre-trained object detection models YOLO-V3, YOLO-V7, and Faster-RCNN on X and Y.
Calculate mAP for rainy and derained images.

N levels, which can be defined as:

$$LP_n(X) = GP_n(X) - upsample(GP_n + 1(X)) \quad (1)$$

In the Equation 1, the function GP_n represents the Gaussian pyramid where $n = 1,\ldots N\text{-}1$, and it can be computed by downsampling $GP_{n-1}(X)$. Unlike previous image restoration methods, which decompose images using a single scale, the Laplacian pyramid adopted a multi-scale decomposition strategy using fixed smoothed kernels. The top level contains detailed background information of a given rainy image, while other levels contain spatial information of an image at multiple scales. To simplify the problem and take advantage of sparsity, the multi-scale decomposition of an image is performed. To improve information flow, we incorporate global skip connections for transmitting shallow-level details.

Three factors led us to use the traditional Laplacian pyramid to decompose the rainy image.

1) The Laplacian pyramid is a computationally efficient algorithm, primarily utilizing convolutions (Gaussian filtering), which can be seamlessly incorporated into GPU-accelerated systems.
2) Top-level of L_n extracts the background; other levels contain rain streaks and details at varying scales. This eliminates rain interference, allowing subnetworks to handle single-scale high-frequency components.
3) Unlike earlier deraining techniques that employ single-scale decomposition, LPN-IDD utilizes a multi-scale approach with Laplacian pyramids. This strategy leverages sparsity at each level, a principle that has driven various deraining methods to simplify the learning task.

2) SUBNETWORK STRUCTURE

A set of independent subnetworks is built for each pyramid level to generate a clean Gaussian pyramid against each

TABLE 1. Architecture parameters for proposed lightweight pyramid network for image draining and detection (LPN-IDD).

Name	Setting	Input	Output
Input	-	$H \times W \times 3$	$H \times W \times 3$
	Shallow features extraction		
Dilated Conv-0	3×3, dil:1 padding:1, stride:1	$H \times W \times 3$	$H \times W \times 3$
Dilated Conv-1	3×3, dil:2 padding:2, stride:1	$H \times W \times 3$	$H \times W \times 3$
	Deep feature extraction (Subnetwork structure with Residual block ×5)		
Depthwise Conv	3×3, padding:1, stride:1	$H \times W \times 3$	$H \times W \times 32$
Depthwise Conv	3×3, padding:1, stride:1	$H \times W \times 32$	$H \times W \times 32$
Pointwise Conv	1×1, padding:0, stride:1	$H \times W \times 32$	$H \times W \times 32$
Dilated Conv-2	3×3, dil:1 padding:1, stride:1	$H \times W \times 32$	$H \times W \times 32$
	Feature Enhancement Block ×1		
Conv-1	3×3, padding:1, stride:1	$H \times W \times 3$	$H \times W \times 32$
	Dual Attention Module ×1	-	-
	Spatial attention module	-	-
Dilated Conv-3	3×3, dil:1, padding:1, stride:1	$H \times W \times 32$	$H \times W \times 32$
Dilated Conv-4	3×3, dil:2, padding:2, stride:1	$H \times W \times 32$	$H \times W \times 32$
	Channel attention module	-	-
Conv-2	3×3, padding:1, stride:1	$H \times W \times 32$	$H \times W \times 3$

input Laplacian pyramid level. The proposed sub-network structure (presented in Fig. 3) comprises of dual attention block (DAB) comprised of channel attention, spatial attention and residual blocks to accelerate the speed of training.

The input image is divided into five pyramid levels and each pyramid level is reconstructed using five independent subnetworks. After the decomposition of images shallow features are extracted by using two dilated convolution layers and deep features are extracted using dual attention blocks. Model deraining performance is enhanced using a feature enhanced block (FEB) module. This is done recursively to form a clean Laplacian pyramid. These clean Laplacian pyramids are then up-sampled with the below level and pass from activation function ReLu to bring non-linearity. At the end, the desired output is clean Gaussian pyramids where the bottom level is our derained image. The following sections provide a detailed breakdown of each subnetwork component and its functioning.

1) Shallow Feature Extraction
The initial layer captures features from the *nth* input level as follows:

$$H_{Conv} = \sigma(W_n^0(Input) + b_n^0) \tag{2}$$

$H_{conv}(.)$ represents basic convolution operation. W represents weights, b denotes biases, and σ is the activation function used to add non-linearity.

$$F^{n,0} = H_{dil-Conv1} * L_n(X) \tag{3}$$

$F^{n,0}$ denotes features utilized for global residual learning, which are then leveraged for further feature extraction.

$$F^{n,1} = H_{dil-Conv2} * F^{n,0} \tag{4}$$

$F^{n,1}$ is the outcome after performing convolution operation on $F^{n,0}$ and serves as the input for deep feature extraction.

2) Recursive Blocks
After extracting shallow features by using two dilated convolution operations, the resultant $F^{n,1}$ feature map is input to the n recursive blocks. The residual and recursive blocks enhance training speed, address vanishing gradient with skip connections, and decrease parameter count through parameter-sharing. To overcome the constraints of conventional convolution for inter-channel relationships within local receiving areas, as depicted, we introduce two depthwise separable convolutions and one point-wise convolutional operation within each recursive block based on the results of our experiments. Although mapping problems are already resolved due to the Laplacian pyramid, however, during feed-forward convolution operation image information may be lost. To tackle this issue, we use skip

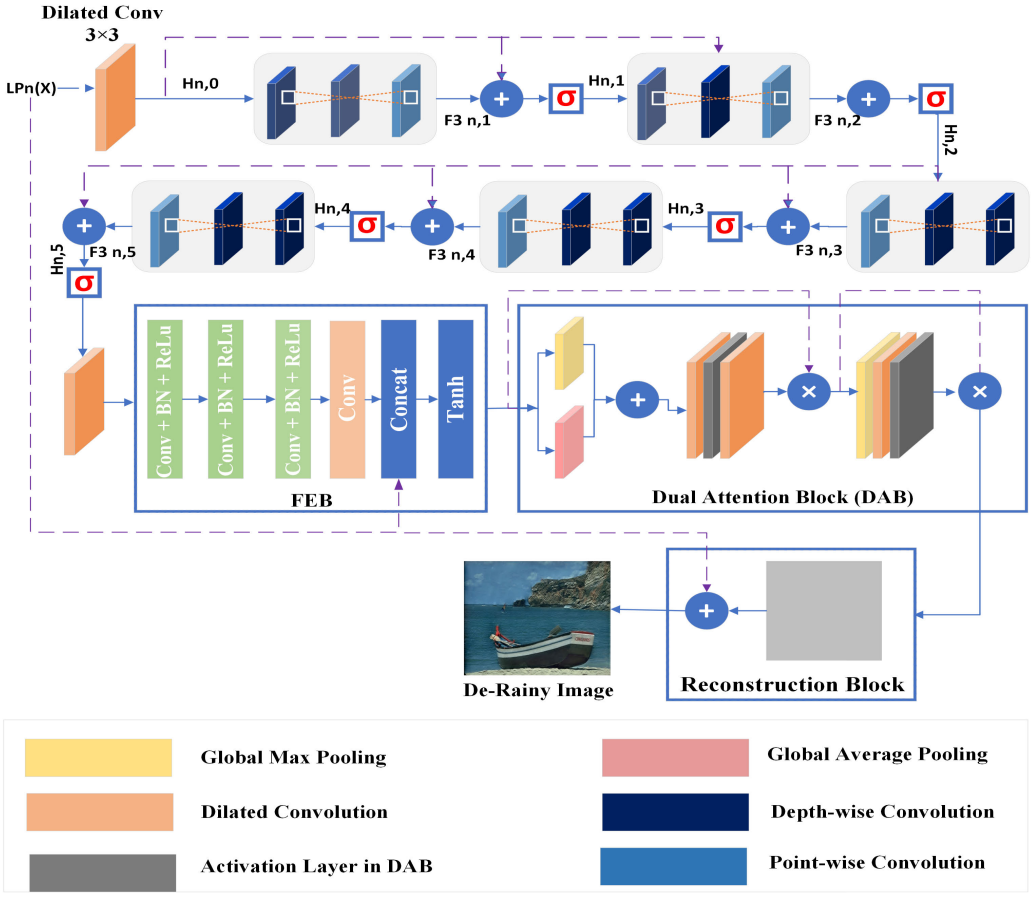


FIGURE 3. The detailed subnetwork structure of the proposed lightweight pyramid network for image draining and detection (LPN-IDD) architecture.

connections in each recursive block. The calculations in the t_{th} recursive block are as shown in Equation 5 and 6.

$$F_t^{n,1} = H_{depth-Conv} * F^{n,1} \tag{5}$$

$$F_t^{n,2} = H_{depth-Conv} * F_t^{n,1} \tag{6}$$

The $F_t^{n,2}$ is the output feature map of two depth-wise convolutions and becomes the input to the point-wise convolution.

$$F_t^{n,3} = H_{point-Conv} * F_t^{n,2} \tag{7}$$

$H_{depth-Conv}$ and $H_{point-Conv}$ represent depth-wise and point-wise convolution(.) operations.

$$F_t^n = F_t^{n,3} + F^{n,0} \tag{8}$$

For information propagation and gradient back-propagation, the output feature map F_t^n of the t_{th} recursive block is obtained through the summation with $F^{n,0}$ (Equation 8).

3) Feature Enhancement Block (FEB)

As network depth grows, the influence of shallow layers on deeper ones weakens. To address this issue, a streamlined four-layer network structure is introduced. Drawing inspiration from this insight, each subnetwork incorporates a feature enhancement block (FEB) to enhance deraining performance. The FEB gathers input from the *nth* image level, combining extracted features with the original image to extract both local and global features effectively. FEB block includes four layers: three dilated convolution layers with batch normalization (BN) and activation function ReLU, plus one convolution layer. Equation 9 and 10 describes the operations.

$$FEB^n = ReLU(BN(H_{Conv} * F_t^n)) \tag{9}$$

$$Y = H_{Conv} * FEB^n \tag{10}$$

Our aim is to combine local and global features. Hence, Y represents the feature map obtained after applying convolution operation on the output of three layered Conv+BN+ReLU. To enhance the expressive ability of our proposed LPN-IDD network, the input rainy image X and the output of convolution layers Y are fused via concatenation operation that helped in getting more diversified features (O).

$$O = TanH(Concat(Y, X)) \tag{11}$$

4) Dual Attention Block (DAB)
The FEB output serves as the input for the proposed DAB module, featuring both channel and spatial attention mechanisms to extract spatial features and maintain image quality. This integration ensures the preservation of sharp edges and texture details.

$$F_{DAB} = F_{SA}(O) * F_{CA}(O) \tag{12}$$

Channel attention involves a fusion of average and max pooling on the input feature map. The pooled features undergo concatenation, followed by a dilated convolution. The outcome of channel attention F_{SA} feeds into spatial attention pooling operation, connected to a convolutional layer. This map identifies crucial spatial regions for deraining. Sigmoid activation and a skip connection address the gradient issue. The obtained weights F_{CA} are element-wise multiplied with the FEB block's output, accentuating robust features. Both FEB and DAB integrated jointly to extract complex streaks information in the existence of the complex background.

5) Gaussian Pyramid
After reconstructing the clean Laplacian pyramid, the goal is to recreate the clean Gaussian pyramid levels. The reconstruction layer is expressed as follows to get the output level of the pyramid.

$$LP_n(Y) = H_{Conv}(F_{DAB}) + LP_n(X) \tag{13}$$

As we have five levels of Laplacian pyramids, the output of $LP_5(Y)$ is fed via a ReLU non-linear activation function to forecast clean Gaussian pyramid level GP_5 starting at the top level (Y).

$$LP_n(Y) = GP_n(X) - upsample(GP_n + 1(X)) \tag{14}$$

The $LP_5(Y)$ output is upsampled and then concatenated with the matching Laplacian pyramid level to produce the bottom-level GP_5 of the clean Gaussian pyramid. So, the derained image's Gaussian pyramid $GP_n(Y)$ can be reconstructed from the obtained Laplacian pyramid output $LP_n(Y)$ as follows.

$$GP_n(Y) = max(0, LP_n(Y) \tag{15}$$

$$GP_n(Y) = max(0, LP_n(Y)) + upsample(GP_{n+1} + (Y)) \tag{16}$$

where $n = 1$ to $N - 1$, the levels of a Gaussian pyramid are ensured to be non-negative using the operation $x = max(0, x)$, akin to rectified linear units (ReLUs). Derained Gaussian pyramid level $GP_4(Y)$ is projected after the ReLU activation function. This process is repeated until $GP_1(Y)$ is reached, which is the resulting derained image.

6) Detection Pipeline
After applying the lightweight pyramid network to derive derained images from original rainy images, we conducted object detection evaluations using prevalent pre-trained models: Faster-RCNN, YOLO-V3, and YOLO-V7. Our evaluation employed a rainy dataset suited for rain-related tasks but lacking annotations necessary for object detection assessments. For object detection on derained images, annotated files were required for both rainy data (model inputs) and the proposed model's output (derained data). These annotations encompass image coordinates, class labels, x-y coordinates, height, and width.

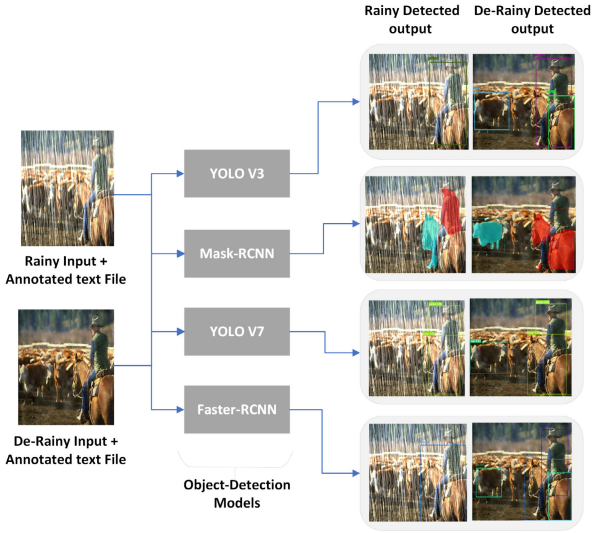


FIGURE 4. Object detection pipeline architecture for proposed LPN-IDD model assessment.

We computed annotations for derained and rainy images alike, enabling the calculation of mean average precision (mAP) for the detection model. The annotation files covered 15 classes, mirroring the 90 classes of the MS COCO dataset. Upon inputting derained images into pre-trained object detection models, the models identified objects and delineated them with bounding boxes, accompanied by confidence scores. For comparative analysis, we also conducted detection on rainy or ground truth images.

Our assessment employed mean average precision (mAP) as the evaluation metric for object detection models. To compute derained image mAP, we utilized annotation files from ground truth images (comprising annotated text files for derained images) alongside corresponding detected annotations yielded by the detection model. Similarly, for rainy image mAP, annotation files from ground truth images (including annotated text files for rainy images) were paired with detected annotation files outputted by the detection model. For both rainy and derained images, we conducted detection using the aforementioned object detection models, calculating average precision (AP), recall, and precision for each class.

B. LOSS FUNCTION

The widely used mean square error (MSE) loss [60], [61] in image restoration often emphasizes squared penalties and pixel-wise discrepancies [45], leading to excessively smoothed outcomes. To address this concern and capture semantic interdependencies among pixels, [45] introduces structural similarity index (SSIM) and L1 losses, especially relevant due to our multi-scale decomposition via Laplacian pyramids. Following this approach, we integrate the same loss functions into our image deraining network across each scale. To retain high-frequency details, all pyramid levels undergo training using a combination of SSIM and L1 losses [68], as denoted in Equation 17. Furthermore, for the initial pyramid level, the focus shifts to L1 loss to uphold higher-frequency nuances. The overall L1 loss is defined for all levels, whereas levels two, three, and four used SSIM loss, which can be defined as:

$$L = \frac{1}{2}\sum_{i=1}^{M}\begin{pmatrix}\sum_{i=1}^{N} L^{l_1}(G_n(Y^i), (G_n(Y_{GT}^i))+ \\ \sum_{n=1}^{2} L^{l_{SSIM}}(G_n(Y^i), (G_n(Y_{GT}^i))\end{pmatrix} \quad (17)$$

Each subnetwork integrates shallow and deep feature extraction blocks, incorporating dual attention mechanisms to enrich and extract local and global features. Furthermore, during back-propagation, gradients from lower pyramid levels inform higher levels, facilitating effective parameter updates. Thus, our streamlined LPN-IDD network, utilizing SSIM and L1 losses, outperforms more complex deep models with extensive parameters employing MSE loss.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

We compare our LPN-IDD model with eight state-of-the-art deraining models like JORDER [60], DID-MDN [34], RESCAN [35], ReHEN [61], DGCN [62], LPNet [45], DGNL [63], DRD-Net [31], MPRNet [64], NLEDN [65], SDDRNet [66], PreNet [29]. All CNN-based comparative methods are retrained on the Augmented-RainTrainH/L dataset and their performance is evaluated on standard evaluation metrics like PSNR and SSIM [68].

A. TRAINING SETTINGS

The proposed LPN-IDD model is trained using TensorFlow version 2.12.0, CUDA Version 11.4 (NVIDIA Tesla P100-PCIE) GPU of Kaggle. The code for the proposed model is written in Python 3.10.11.

1) TRAINING DATASET

Due to the challenge of acquiring extensive real-world clean/rain image pairs, we employ three synthetic rainy datasets, namely Rain100H [62], Rain100L [62], and Rain12 [63], for both training and testing our proposed network. Rain100H [62] comprises 1800 pairs of heavily rainy images for training. For Rain100L [62], 200 pairs of light rain training images and 100 test image pairs are available. Rain12 offers 12 pairs of clean/rain images. Given the limited size of Rain12, the training outcomes from Rain100L are also utilized for Rain12 testing purposes.

To enhance the training dataset, augmentation methods like hflip, vflip, vhflip, and hvflip were used to modify rain streak directions. This broadens the model's ability to recognize diverse directional patterns, enabling effective learning from varied angles and orientations. So, after applying augmentation techniques on datasets, the training dataset includes 8000 images (321 × 481 pixels) with variable Gaussian noise and rain streaks for enhanced training. To capture detailed information, images are divided into 100 × 100 patches, bolstering feature robustness and deraining efficiency.

TABLE 2. Quantitative performance comparison of LPN-IDD with different state-of-the-art image deraining methods on synthetic datasets.

Datasets	Rain100H	Rain100L	Rain12	Parameters
Methods	SSIM/PSNR	SSIM/PSNR	SSIM/PSNR	
DGNL	0.874/29.59	0.959/35.11	0.95/36.08	4,036,586
ReHEN	0.86/27.97	0.98/37.41	0.95/35.84	499,668
DGCN	0.90/30.48	0.98/37.59	0.95/35.9	2,731,071
JORDER	0.83/26.6	0.96/36.6	0.95/35.9	369,792
DRD-Net	0.88/28.09	0.97/37.11	0.95/36.59	5,230,214
DID-MDN	0.61/17.39	0.85/25.79	0.88/29.77	372,839
RESCAN	0.84/26.75	0.98/37.23	0.93/35.51	499,668
LPN-IDD	**0.868/27.2**	0.97/33.6	**0.967/34.7**	**64,554**

2) TESTING DATASET

For Rain100H [62], Rain1400 [28] and Rain100L [62] 200 and 100 image pairs are used for testing. Reference [45] released 300 real images of rain taken from the Internet as a fresh data collection. The effectiveness of our LPN-IDD model in real-world conditions is evaluated using this real-world dataset and RainDS-real [69] and IVIPC-DQA [70] real rainy datasets. The BSD500 dataset [31], known as the Berkeley Segmentation Dataset, is a prominent benchmark for CV and image processing tasks like segmentation, edge detection, and texture analysis. It features 500 diverse, human-analyzed images from the natural world. We examine 500 clean or non-rainy images from the BSD500 data set to show that our LPN-IDD output is nearly identical to the clean image.

B. EVALUATION ON SYNTHETIC DATASETS

According to Table 2, a quantitative analysis of our LPN-IDD model shows the proposed method outperforms other state-of-the-art deraining techniques on three different synthetic rainy datasets affirming the efficacy of our approach in producing clean outcomes. The LPN-IDD model surpasses the comparative methods, mainly in SSIM metrics. The computational analysis highlights LPN-IDD's efficiency with fewer parameters and lower cost while maintaining competitive SSIM scores against larger-parameter alternatives.

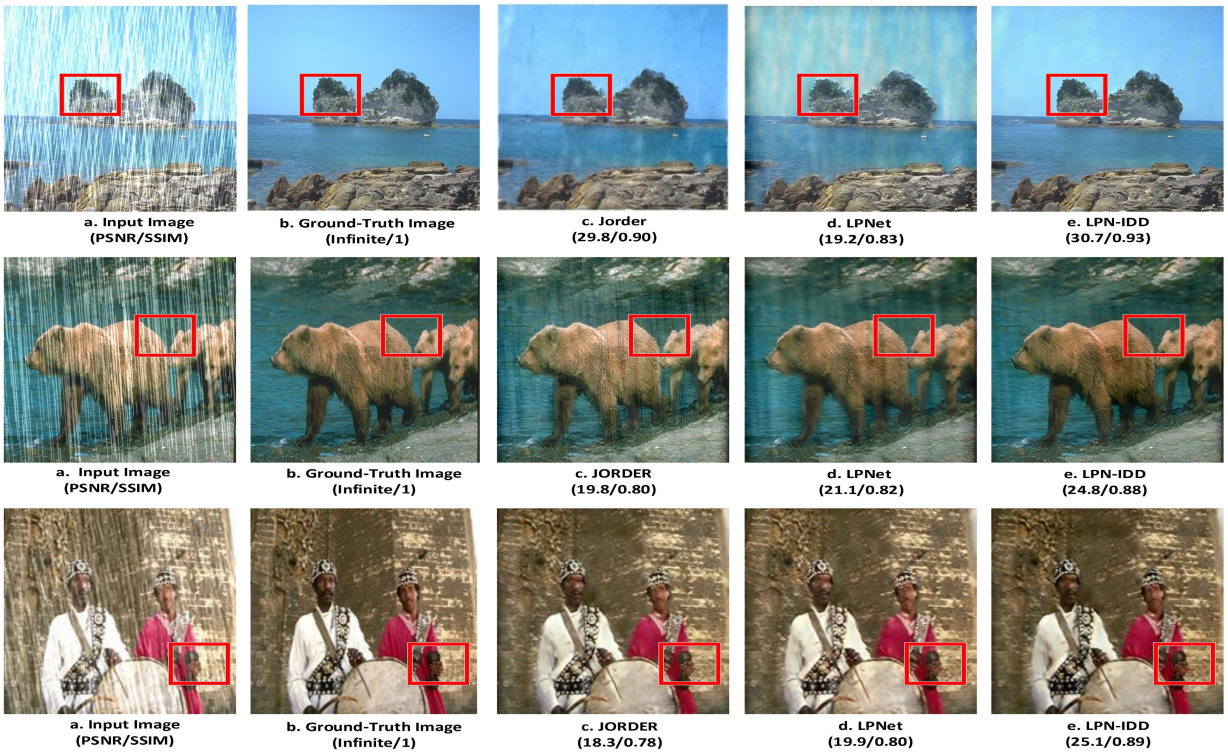


FIGURE 5. Rain-removal results of all comparing approaches on synthetic images selected from Rain 100H dataset with a static and complex background and dynamic rain.

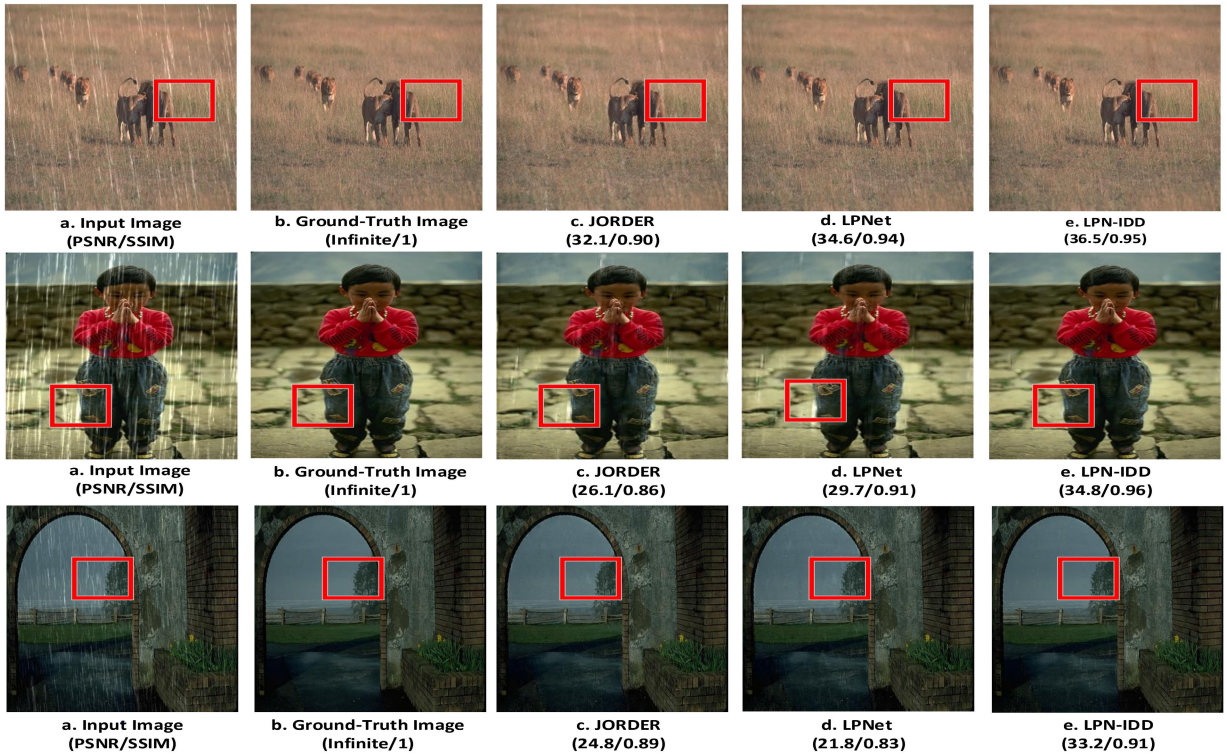


FIGURE 6. Rain-removal results of all comparing approaches on synthetic images selected from Rain 100L dataset with a static and complex background and dynamic rain.

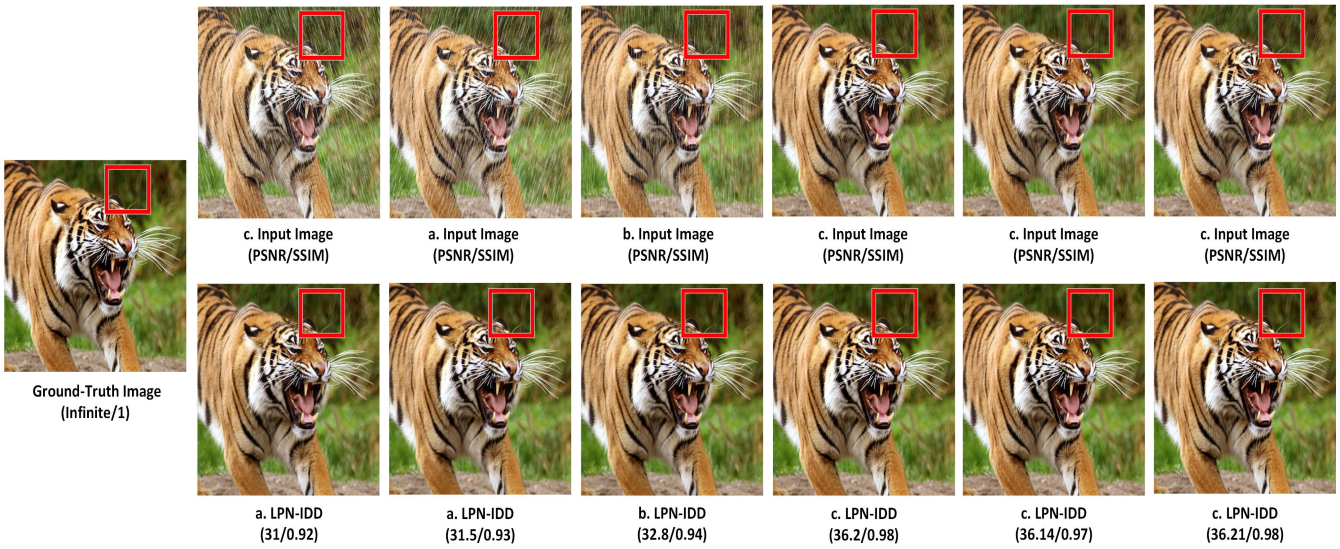


FIGURE 7. Rain-removal results of all comparing approaches on synthetic images selected from the Rain1400 dataset with a static and complex background and dynamic rain.

TABLE 3. Computational complexity of LPN-IDD and other state-of-the-art models.

Method	Avg run-time(s)	Parameters
DID-MDN	0.539	372,839
NLEDN	0.891	1,005,379
RESCAN	0.940	499,668
ReHEN	0.576	298,263
PreNet	0.255	168,963
DRD-Net	13.019	5,230,214
SDDRNet	0.454	244,975
DGNL	0.037	4,036,586
DGCN	2.511	2,731,071
MPRNet	1.133	3,637,303
JORDER	0.18	369,792
LPN-IDD	0.18	64,554

While CNN-based methods excel in synthetic light rain images, they often struggle with heavy rain, leaving behind artifacts and face difficulty in handling diverse directional rain streaks. Our qualitative evaluation compares LPN-IDD with diverse deraining algorithms across various synthetic rainy datasets and real-world data, illustrating robust generalization and capturing variations in rain streak lengths and directions.

In Fig. 5, we qualitatively compare our proposed LPN-IDD model with various methods on the synthetic rainy dataset Rain100H. The first image (a) depicts the rainy input, the second (b) shows the clear ground truth and subsequent images exhibit outcomes from different methods. Notably, (c) and (d) models exhibit inadequate rain removal, introducing blurriness and artifacts, and obscuring key details. In contrast, our proposed model (e) visually outperforms, closely resembling the ground-truth image while maintaining edge and crucial details.

In Fig. 6, the proposed model LPN-IDD is qualitatively compared with various other comparison methods on the

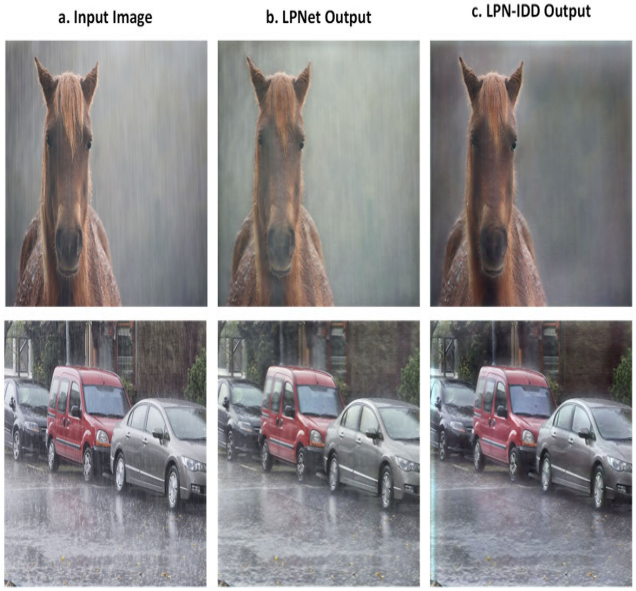


FIGURE 8. Qualitative evaluation of LPNet and proposed rain removal model LPN-IDD on a rainy image of real-world dataset [45].

synthetic rainy datasets Rain100L and Rain12. The first image (a) represents the rainy input image and the second image (b) is the ground truth image or clear image, other images represent various comparison scheme results. It can be clearly seen that (c) and (d) models do not remove rain properly and it also creates blurry artifacts in the images. The last image (e) which is our proposed model's output image visually outperforms the other comparative models.

The proposed model LPN-IDD is qualitatively evaluated on the synthetic rainy dataset Rain1400 [28] which contains

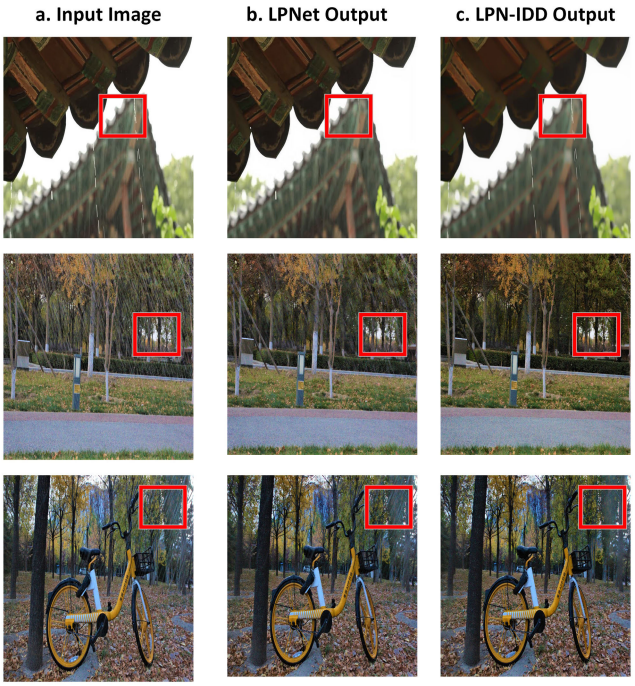


FIGURE 9. Qualitative evaluation of LPNet and proposed rain removal model LPN-IDD on a rainy image of real-world RainDS-real [69] and IVIPC-DQA [70] datasets.

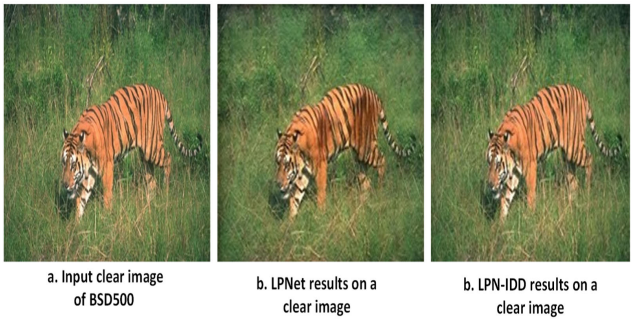


FIGURE 10. Qualitative evaluation of baseline and proposed rain removal model on a non-rainy image of BSD500 dataset [31].

diverse rainy scenarios and streak directions (rainy images contain 14 different directional streaks with diverse intensities). In Fig. 7, the first three images in the first row represent the heavy rainy input images of Rain1400 with different directional streaks, and the next three images represent the light rainy input images with different directional streaks. The left first image is the ground truth image or clear image. The images in the second row are our proposed model LPN-IDD output images. Our method excels in various rainy scenarios, producing results that closely resemble clear, non-rainy images.

Table 3 demonstrates that our technique has a lower computational cost as compared to all the state-of-the-art methods as it has the least number of parameters when compared to alternative strategies which have a huge number

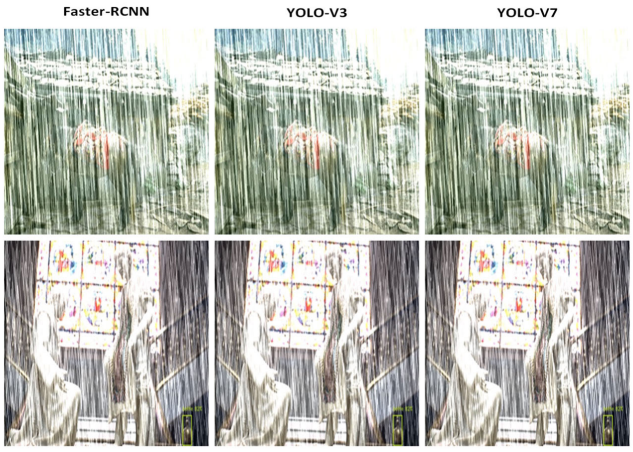


FIGURE 11. Task-driven evaluation of different object detection models on rainy input.

of parameters. From all comparison methods, only two achieve higher SSIM values but these models have a huge number of parameters.

C. EVALUATION ON REAL-WORLD AND NON-RAINY DATASETS

To further verify the robust adaptability of our model, the proposed model LPN-IDD is qualitatively compared with different rain removal strategies like LPNet and JORDER on a real-world dataset. We conducted evaluations using rainy images from the real-world dataset [45], RainDS-real [69] and IVIPC-DQA [70] datasets which contain real rain images of various lighting conditions and different scenes. The outcomes, as illustrated in Fig. 8 and 9, indicate that while LPNet does not entirely eliminate rain streaks and some rainy streaks are still visible since rain patterns frequently have complex distributions and diverse shapes, but our approach LPN-IDD demonstrates significant improvements. It more effectively reduces the visibility of rain streaks, yielding results that more closely resemble the appearance of images without rain. This evidence suggests that our method not only stands out in synthetic rainy evaluations but also shows promising capabilities in processing images from real-world rain scenarios.

In Fig. 10, the proposed model LPN-IDD is qualitatively compared with the baseline approach LPNet on a non-rainy dataset BSD500 dataset [31]. LPN-IDD produces more accurate results, while LPNet produces over-smoothness in the image. Both quantitative and visual results clearly depict that the proposed LPN-IDD approach is better than conventional techniques in terms of PSNR and SSIM scores [68], keeping a balance between edge and texture details and also in terms of storage space due to its lightweight nature.

D. TASK DRIVEN EVALUATION

For evaluating our proposed model's object detection performance on both rainy and derained images, we employed

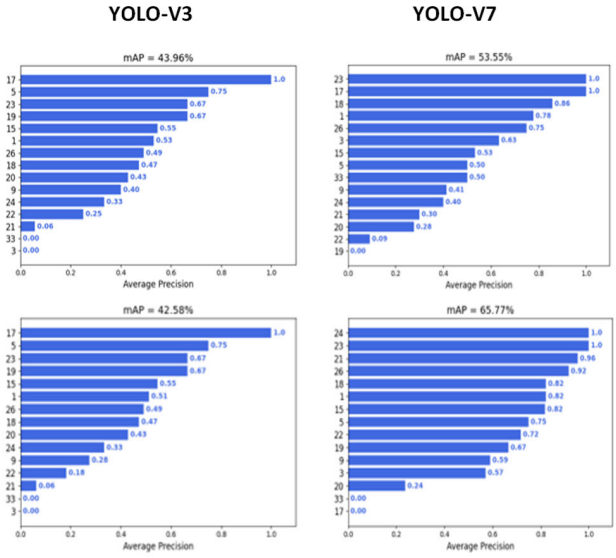


FIGURE 12. mAP of different object detection models on rainy images (first row) and derained results (second row) of our proposed LPN-IDD model.

established state-of-the-art algorithms. We integrated these algorithms by leveraging pre-trained object detection models, readily available in prominent deep learning frameworks. These models were originally trained on the COCO dataset [67].

In essence, our objective isn't solely enhancing rain detection accuracy. Instead, we utilize a robust object detection model as a consistent benchmark to gauge diverse deraining algorithms. This impartial approach ensures fairness, as we refrain from fine-tuning object detectors specifically for rainy or derained images. The object detection models, pre-trained on the MS COCO dataset [67], provide an objective assessment of the efficacy of distinct deraining techniques from a complementary standpoint.

Fig. 11 illustrates the challenges encountered by object detection models when applied to rainy images. The presence of intricate directional rain streaks introduces complexity, leading to diminished performance. Objects and backgrounds often lack clarity due to rain-induced distortions, making even human interpretation challenging. Fig. 12 shows that by using the baseline LPNet deraining algorithm to remove rain from images, the resulting images tend to have worsened performance as compared to using the original rainy images without applying any deraining mechanism in detecting objects. Because LPNet produces blurry artifacts that are not very clear and might remove important and meaningful information that is why object detection models do not perform well even after applying the deraining model on rainy images.

Fig. 13 shows that by using the proposed LPN-IDD deraining algorithm to remove rain from images, the resulting images tend to have better performance as compared to using the original rainy images without applying any deraining model in detecting objects by using various object detection models.

TABLE 4. Quantitative analysis in terms of the number of kernel numbers and the number of parameters.

Kernels	Datasets			Parameters
	SSIM/PSNR			
	Rain100H	Rain100L	Rain12	
16	0.82/23.4	0.96/33.4	0.95/33.7	7454
32	0.87/27.2	0.97/33.6	0.967/34.7	64554
64	0.88/29.2	0.98/33.9	0.98/35.6	273560

TABLE 5. Quantitative analysis in terms of number of pyramid levels.

Pyramids	Datasets			Parameters
	SSIM/PSNR			
	Rain100H	Rain100L	Rain12	
3	0.82/23.4	0.96/33.4	0.95/33.7	64554
5	0.87/27.2	0.97/33.6	0.967/34.7	64554
7	0.87/29.2	0.98/33.9	0.97/35.6	64554

E. ABLATION STUDY

In this section, different configuration settings are discussed and their impact on the performance of the proposed methodology is evaluated.

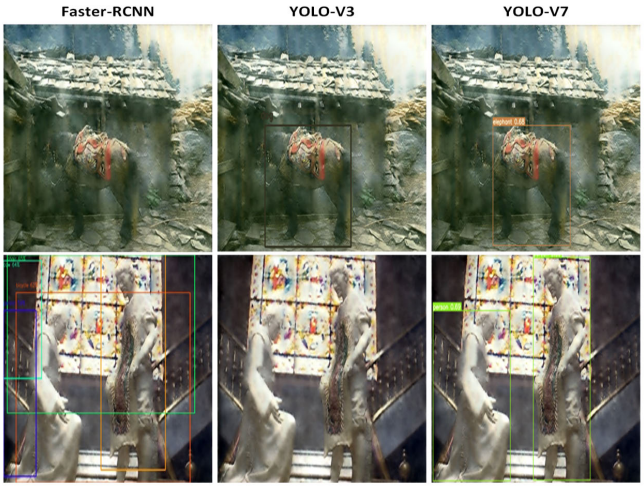
1) INCREASING KERNAL NUMBER

To accommodate diverse image details or complex background details we validate our model on a varied number of kernels for every pyramid level. We use higher kernel numbers (32, 16) for lower levels and (4, 4, 16) for higher levels on default. Table 4 demonstrates that our method outperforms comparative approaches in both SSIM and PSNR values across all datasets using the default kernel configuration.

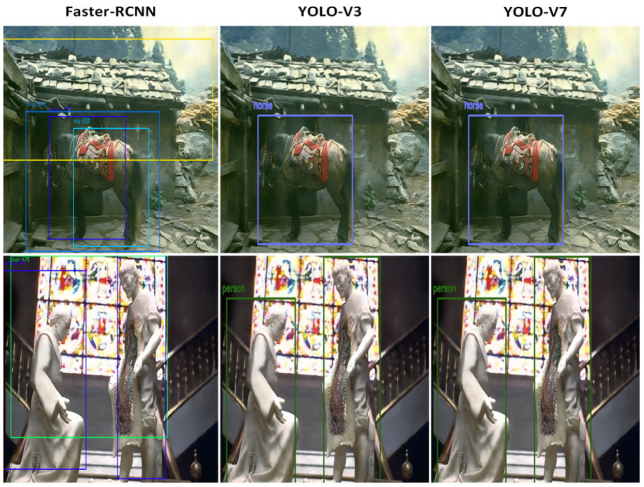
We have conducted experiments with the increased kernel parameters for different pyramid levels. Using more kernels, in our opinion, will further enhance performance. But when the kernel count rises, more storage and processing power are needed. We use 16 feature maps for layers where we use 8 convolution layers on default. We also perform experiments by increasing kernel count along different pyramid levels i.e. 8,16,32. As can be seen, increased kernel numbers marginally enhance PSNR and SSIM metrics but don't significantly alter visual quality. So, to attain a balance between effectiveness and efficiency, we set our kernel parameter to decrease.

2) NUMBER OF PYRAMID LEVELS

The default setting for the number of pyramids is set to 5, although experiments are conducted for other pyramid numbers as well. In Table 5, average PSNR and SSIM comparison is performed based on a different number of



(a) Task-Driven evaluation of different object detection models on LPNet-generated derained images.



(b) Task-Driven evaluation of different object detection models on LPN-IDD's generated derained images.

FIGURE 13. Task-driven evaluation of different object detection models on (a) LPNet and (b) LPN-IDD's generated derained images.

TABLE 6. Quantitative comparison of LPN-IDD with different state-of-the-art image deraining methods based on performance.

Datasets	Rain100H	Rain100L	Parameters
LPN-IDD	SSIM/PSNR	SSIM/PSNR	
Spatial Attention	0.83/23.0	0.93/29.4	64494
Dual Attention	0.87/27.2	0.97/33.6	64554

pyramids for different datasets. Results clearly state that increasing levels from 3 to 5 gives better performance. But when we further increase the levels, results show there is a minor increase in PSNR and SSIM values. This is because, with five pyramid levels, the image is already split well, making the learning task simpler. Adding more levels doesn't significantly help solve the problem better. Hence, we stick to the default setting of using five pyramid levels for image splitting.

3) SUBNETWORK STRUCTURE WITH DIFFERENT ATTENTION BLOCKS

In this section, we analyze the performance of our model by using different combinations of attention blocks used in the subnetworks structure.

We first concatenate only the spatial attention block with the proposed lightweight pyramid network, then we perform experiments by concatenating a combination of spatial and channel attention networks known as dual attention blocks with the proposed LPN-IDD model. In Table 6 average SSIM and PSNR comparison is performed based on different combinations of attention blocks for different datasets. Results clearly state that there is a huge difference in image quality as the dual attention block shows remarkable results in removing rain streaks from rainy images, however, there is a minor increase in the number of parameters.

V. CONCLUSION

A lightweight multi-scale LPN-IDD network based on a dual attention network and Gaussian-Laplacian pyramid for image deraining is proposed. The proposed method is comprised of multiple independent subnetworks. Each subnetwork takes the input from the Laplacian pyramid, enhances the features of a decomposed image using a feature extraction block, extracts key features using a multi-scale dual attention module and predicts corresponding clean Gaussian pyramid at each scale using recursive blocks. To accelerate the training speed residual blocks are adopted and to reduce parameters using a parameter-sharing strategy, recursive blocks are implemented. Proposed LPN-IDD with approximately 60k parameters showed outstanding performance as compared to various other state-of-the-art image deraining methods consisting of parameters in millions that have been used for comparison in terms of visual and quantitative analysis. To get a better generalization capability of the proposed method, extensive experiments were performed on various image deraining methods and also tested on real-world rainy images for rain removal. We also used multiple object detection methods to perform task-driven evaluations of our proposed and comparative deraining models. The proposed method in comparison with other approaches showed a PSNR score of 23.755114 (dB) and an SSIM score of 0.87.

VI. FUTURE WORK

The proposed model LPN-IDD shows great performance on the heavy and light rain streaks datasets. Our future work includes considering more rainy patterns (Rain-drops, Rain-mist, Rain-drops-streaks) and various rainy conditions (cloudy, night rain), etc which will help to improve the model's ability to generalize well with real-world scenarios. Increasing the number of training samples using a generative

adversarial network can be a good choice to enhance the draining performance. Furthermore, the lightweight and multi-scale decomposition nature of the proposed model LPN-IDD can be adapted to other image restoration tasks like image dehazing. In the future, real-world rainy images can also be used for training, but the lack of available clean or reference pairs makes this process challenging. We plan to shift towards semi-supervised or unsupervised learning to better utilize unlabeled data particularly and to generalize well with real-world scenarios. Additionally, our ongoing efforts will focus on integrating the latest advancements in deep learning and image processing, ensuring our model stays advanced and effective for a wide range of image processing challenges.

ACKNOWLEDGMENT

The authors also extend their appreciation for the necessary support of Al-Ahliyya University, Amman, Jordan.

REFERENCES

[1] A. Broggi, A. Zelinsky, U. Ozguner, and C. Laugier, "Intelligent vehicles," in *Springer Handbook of Robotics*. Cham, Switzerland: Springer, 2016, pp. 1627–1656.

[2] M. Riesenhuber and T. Poggio, "Models of object recognition," *Nature Neurosci.*, vol. 3, no. 11, pp. 1199–1204, 2000.

[3] X. Lu, E. Erkip, Y. Wang, and D. Goodman, "Power efficient multimedia communication over wireless channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 10, pp. 1738–1751, Dec. 2003.

[4] L.-H. Wang, Y.-M. Hsiao, X.-Q. Xie, and S.-Y. Lee, "An outdoor intelligent healthcare monitoring device for the elderly," *IEEE Trans. Consum. Electron.*, vol. 62, no. 2, pp. 128–135, May 2016.

[5] S. Yasunobu and R. Sasaki, "An auto-driving system by interactive driving knowledge acquisition," in *Proc. IEEE SICE Annu. Conf.*, Fukui, Japan, Aug. 2003, pp. 2935–2938.

[6] 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)*, vol. 2. Zhangjiajie, China, May 2012, pp. 304–307.

[7] X. Zheng, Y. Liao, W. Guo, X. Fu, and X. Ding, "Single-image-based rain and snow removal using multi-guided filter," in *Proc. Int. Conf. Neural Info. Process.*, Berlin, Germany, 2013, pp. 258–265.

[8] X. Ding, L. Chen, X. Zheng, Y. Huang, and D. Zeng, "Single image rain and snow removal via guided l0 smoothing filter," *Multimedia Tools Appl.*, vol. 75, no. 5, pp. 2697–2712, Mar. 2016.

[9] Y. Chang, L. Yan, and S. Zhong, "Transformed low-rank model for line pattern noise removal," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Venice, Italy, Oct. 2017, pp. 1735–1743.

[10] 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.*, Sydney, NSW, Australia, Dec. 2013, pp. 1968–1975.

[11] 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.

[12] H. Zhu, C. Wang, Y. Zhang, Z. Su, and G. Zhao, "Physical model guided deep image deraining," in *Proc. IEEE Int. Conf. Multimedia Expo. (ICME)*, Jul. 2020, pp. 1–6.

[13] Y. Li, R. T. Tan, X. Guo, J. Lu, and M. S. Brown, "Rain streak removal using layer priors," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 2736–2744.

[14] L. Zhu, C.-W. Fu, D. Lischinski, and P.-A. Heng, "Joint bi-layer optimization for single-image rain streak removal," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Venice, Italy, Oct. 2017, pp. 2545–2553.

[15] K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep CNN denoiser prior for image restoration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 2808–2817.

[16] P. Mu, J. Chen, R. Liu, X. Fan, and Z. Luo, "Learning bilevel layer priors for single image rain streaks removal," *IEEE Signal Process. Lett.*, vol. 26, no. 2, pp. 307–311, Feb. 2019.

[17] G. Sun, H. Shao, and C. Cattani, "*A priori*-guided multi-layer rain-aware network for single image deraining," *Knowl.-Based Syst.*, vol. 235, Jan. 2022, Art. no. 107613.

[18] 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.

[19] S. Gu, D. Meng, W. Zuo, and L. Zhang, "Joint convolutional analysis and synthesis sparse representation for single image layer separation," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Venice, Italy, Oct. 2017, pp. 1717–1725.

[20] 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.

[21] S.-H. Sun, S.-P. Fan, and Y. F. Wang, "Exploiting image structural similarity for single image rain removal," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Paris, France, Oct. 2014, pp. 4482–4486.

[22] 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)*, Santiago, Chile, Dec. 2015, pp. 3397–3405.

[23] P. C. Barnum, S. Narasimhan, and T. Kanade, "Analysis of rain and snow in frequency space," *Int. J. Comput. Vis.*, vol. 86, nos. 2–3, pp. 256–274, Jan. 2010.

[24] K. Zhang, W. Luo, W. Ren, J. Wang, F. Zhao, L. Ma, and H. Li, "Beyond monocular deraining: Stereo image deraining via semantic understanding," in *Proc. 16th Eur. Conf. Comp. Visi. (ECCV)*, Glasgow, U.K., Aug. 2020, pp. 71–89.

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

[26] Z. Fan, H. Wu, X. Fu, Y. Hunag, and X. Ding, "Residual-guide feature fusion network for single image deraining," 2018, *arXiv:1804.07493*.

[27] V. Santhaseelan and V. K. Asari, "Utilizing local phase information to remove rain from video," *Int. J. Comput. Vis.*, vol. 112, no. 1, pp. 71–89, Mar. 2015.

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

[29] D. Ren, W. Zuo, Q. Hu, P. Zhu, and D. Meng, "Progressive image deraining networks: A better and simpler baseline," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Long Beach, CA, USA, Jun. 2019, pp. 3932–3941.

[30] Y. Wang, Y. Song, C. Ma, and B. Zeng, "Rethinking image deraining via rain streaks and vapors," in *Proc. 16th Eur. Conf. Comp. Visi. (ECCV)*, Glasgow, U.K., Aug. 2328, 2020, pp. 367–382.

[31] S. Deng, M. Wei, J. Wang, Y. Feng, L. Liang, H. Xie, F. L. Wang, and M. Wang, "Detail-recovery image deraining via context aggregation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Seattle, WA, USA, Jun. 2020, pp. 14548–14557.

[32] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.

[33] 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, Jun. 2017.

[34] 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.

[35] X. Li, J. Wu, Z. Lin, H. Liu, and H. Zha, "Recurrent squeeze-and-excitation context aggregation net for single image deraining," in *Proc. Eur. Conf. Comput. Vis.*, Sep. 2018, pp. 254–269.

[36] H. Wu, Z. Zou, J. Gui, W.-J. Zeng, J. Ye, J. Zhang, H. Liu, and Z. Wei, "Multi-grained attention networks for single image super-resolution," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 2, pp. 512–522, Feb. 2021.

[37] V. Mnih, N. Heess, and A. Graves, "Recurrent models of visual attention," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 27, 2014, pp. 1–8.

[38] L. Chen, H. Zhang, J. Xiao, L. Nie, J. Shao, W. Liu, and T.-S. Chua, "SCA-CNN: Spatial and channel-wise attention in convolutional networks for image captioning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 5659–5667.

[39] 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. 8014–8023.

[40] T. Wang, X. Yang, K. Xu, S. Chen, Q. Zhang, and R. W. H. Lau, "Spatial attentive single-image deraining with a high quality real rain dataset," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Long Beach, CA, USA, Jun. 2019, pp. 12262–12271.

[41] H. Zhang, Q. Xie, B. Lu, and S. Gai, "Dual attention residual group networks for single image deraining," *Digit. Signal Process.*, vol. 116, Sep. 2021, Art. no. 103106.

[42] Y. Wei, Z. Zhang, M. Xu, R. Hong, J. Fan, and S. Yan, "Robust attention deraining network for synchronous rain streaks and raindrops removal," in *Proc. 30th ACM Int. Conf. Multi.*, Lisboa Portugal, Oct. 2022, pp. 6464–6472.

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

[44] N. Ahn, B. Kang, and K.-A. Sohn, "Fast, accurate, and lightweight superresolution with cascading residual network," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Munich, Germany, Sep. 2018, pp. 252–268.

[45] 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.

[46] Y. Zheng, X. Yu, M. Liu, and S. Zhang, "Single-image deraining via recurrent residual multiscale networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 3, pp. 1310–1323, Mar. 2022.

[47] C. Wang, X. Xing, Y. Wu, Z. Su, and J. Chen, "DCSFN: Deep cross-scale fusion network for single image rain removal," in *Proc. 28th ACM Int. Conf. Multimedia*, Seattle, WA, USA, Oct. 2020, pp. 1643–1651.

[48] K. Jiang, Z. Wang, P. Yi, C. Chen, B. Huang, Y. Luo, J. Ma, and J. Jiang, "Multi-scale progressive fusion network for single image deraining," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Seattle, WA, USA, Jun. 2020, pp. 8343–8352.

[49] S. Liang, X. Meng, Z. Su, and F. Zhou, "Multi-receptive field aggregation network for single image deraining," *J. Vis. Commun. Image Represent.*, vol. 84, Apr. 2022, Art. no. 103469.

[50] Z. Pan, J. Wang, Z. Shen, S. Han, and J. Zhu, "Cross-domain collaborative learning for single image deraining," *Expert Syst. Appl.*, vol. 211, Jan. 2023, Art. no. 118611.

[51] W.-Y. Hsu and W.-C. Chang, "Recurrent wavelet structure-preserving residual network for single image deraining," *Pattern Recognit.*, vol. 137, pp. 109–294, May 2023.

[52] Y. Liu, R. Zhang, Y. Zhang, X. Pan, X. Yao, Z. Ni, and H. Han, "Recurrent context-aware multi-stage network for single image deraining," *Comput. Vis. Image Understand.*, vol. 227, Jan. 2023, Art. no. 103612.

[53] M. Wang, C. Li, and F. Ke, "Recurrent multi-level residual and global attention network for single image deraining," *Neural Comput. Appl.*, vol. 35, no. 5, pp. 3697–3708, Feb. 2023.

[54] P. Li and S. Gai, "Single image deraining using multi-scales context information and attention network," *J. Vis. Commun. Image Represent.*, vol. 90, Feb. 2023, Art. no. 103695.

[55] J.-G. Wang and C.-S. Wu, "Multi-scale aggregation residual channel attention fusion network for single image deraining," *Appl. Sci.*, vol. 13, no. 4, p. 2709, Feb. 2023.

[56] N. Jiang, J. Luo, J. Lin, W. Chen, and T. Zhao, "Lightweight semi-supervised network for single image rain removal," *Pattern Recognit.*, vol. 137, May 2023, Art. no. 109277.

[57] B. Xiaojun, C. Zheng, Y. Jianyu, and W. Haibo, "LRP-Net: A lightweight recursive pyramid network for single image deraining," *Neurocomputing*, vol. 490, pp. 181–192, Jun. 2022.

[58] Z. Zhang, Y. Wei, H. Zhang, Y. Yang, S. Yan, and M. Wang, "Data-driven single image deraining: A comprehensive review and new perspectives," *Pattern Recognit.*, vol. 143, Nov. 2023, Art. no. 109740.

[59] S. Han, H. Mao, and W. J. Dally, "Deep compression: Compressing deep neural networks with pruning, trained quantization and Huffman coding," in *Proc. Int. Conf. Learn. Represent.*, Feb. 2016, pp. 1–14.

[60] 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.

[61] Y. Yang and H. Lu, "Single image deraining via recurrent hierarchy enhancement network," in *Proc. 27th ACM Int. Conf. Multimedia*, Nice, France, Oct. 2019, pp. 1814–1822.

[62] X. Fu, Q. Qi, Z. J. Zha, Y. Zhu, and X. Ding, "Rain streak removal via dual graph convolutional network," in *Proc. AAAI Conf. Artif. Intell.*, Feb. 2021, pp. 1–9.

[63] X. Hu, L. Zhu, T. Wang, C.-W. Fu, and P.-A. Heng, "Single-image real-time rain removal based on depth-guided non-local features," *IEEE Trans. Image Process.*, vol. 30, pp. 1759–1770, 2021.

[64] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, M.-H. Yang, and L. Shao, "Multi-stage progressive image restoration," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Nashville, TN, USA, Jun. 2021, pp. 14816–14826.

[65] G. Li, X. He, W. Zhang, H. Chang, L. Dong, and L. Lin, "Non-locally enhanced encoder–decoder network for single image de-raining," in *Proc. 26th ACM Int. Conf. Multimedia*, Oct. 2018, Art. no. 10561064.

[66] C.-Y. Lin, Z. Tao, A.-S. Xu, L.-W. Kang, and F. Akhyar, "Sequential dual attention network for rain streak removal in a single image," *IEEE Trans. Image Process.*, vol. 29, pp. 9250–9265, 2020.

[67] T. Lin, "Microsoft COCO: Common objects in context," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 740–755.

[68] 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.

[69] R. Quan, X. Yu, Y. Liang, and Y. Yang, "Removing raindrops and rain streaks in one go," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Nashville, TN, USA, Jun. 2021, pp. 9143–9152.

[70] Q. Wu, L. Wang, K. N. Ngan, H. Li, F. Meng, and L. Xu, "Subjective and objective de-raining quality assessment towards authentic rain image," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 11, pp. 3883–3897, Nov. 2020.

KAINAT BABAR received the B.S. degree in information and technology from the University of Punjab, Pakistan, in 2020, and the M.S. degree in computer science from COMSATS University Islamabad, Islamabad, Pakistan, in 2023. She is currently a Lecturer with the Department of Computer Science, Shifa Tameer-e-Millat University, Islamabad, teaching various topics in machine learning, artificial intelligence, and core computer science concepts. She has published a conference paper, in 2022. Her research interests include exploring emerging technologies, such as machine learning, natural language processing, computer vision, and data science. She is also passionate about applying AI to solve real-world problems and staying updated with the latest developments in the field.

MUHAMMAD USMAN YASEEN is currently an Assistant Professor with COMSATS University Islamabad (CUI), Islamabad, Pakistan. His current research interests include video analytics, big data analysis, machine learning, and distributed systems.

AHMAD SAMI AL-SHAMAYLEH received the master's degree in information systems from The University of Jordan, Jordan, in 2014, and the Ph.D. degree in artificial intelligence from the University of Malaya, Malaysia, in 2020. He is currently an Assistant Professor with the Faculty of Information Technology, Al-Ahliyya Amman University, Jordan. His research interests include artificial intelligence, human–computer interaction, the IoT, Arabic NLP, Arabic sign language recognition, language resources production, the design and evaluation of interactive applications for handicapped people, multimodality, and software engineering.

MUHAMMAD IMRAN received the degree in software engineering from the University of Engineering and Technology, Taxila, Pakistan, in 2006, and the master's degree in software engineering and the Ph.D. degree in computer science from the University of Southampton, U.K., in 2009 and 2015, respectively. He was a Lecturer with COMSATS University Islamabad (CUI), Islamabad, Pakistan, from 2007 to 2008. He is currently an Assistant Professor with the Department of Computer Science, CUI. His research interests include artificial intelligence, machine learning, natural language processing, and semantic web.

ABDULLAH HUSSEIN AL-GHUSHAMI (Senior Member, IEEE) received the Bachelor of Information Technology degree in network security, the Master Science of Information Technology degree in cyber and network security, and the Ph.D. degree in information technology (cloud security) from Universiti Utara Malaysia, in 2012, 2014, and 2018, respectively. He is currently an Assistant Professor of cyber and network security with the Department of Information Technology, Community College of Qatar (CCQ). He has a long experience in working with higher education as a Lecturer and an Assistant Professor, since 2009. He has published several research papers in well-known international journals and conferences. He is actively pursuing research in the area of cloud computing security, network security, computer networks, cybersecurity, mobile and wireless networks, and cybercrime investigation. He is a member of the Safecyber Systems Corporation for security solutions development. The current focus of Safecyber Corporation is developing several systems, including Safecyber. In terms of awards, he has been received many gold, silver, and bronze medals from local, national, and international exhibitions and universities. He acts as a Volunteer Reviewer in well-reputed journals, such as IEEE Systems Journal and Computer and Security and IEEE Journal of Cloud Computing.

ADNAN AKHUNZADA (Senior Member, IEEE) worked as a Lecturer, a Senior Lecturer, a Year Tutor, an Occasional Lecturer at other engineering departments; an Assistant Professor with COMSATS University Islamabad (CUI), a Senior Researcher with RISE SICs Vãsteras AB, Sweden; a Research Fellow and the Scientific Lead with the DTU Compute, Technical University of Denmark (DTU); the Course Director of ethical hacking with The Knowledge Hub Universities (TKH), Coventry University, U.K.; an Associate Professor with USM; and a Visiting Professor having the mentorship of graduate students and supervision of academic and research and development projects both at UG and PG levels. His experience as an Educator and a Researcher is diverse. He has a proven track record of high-impact published research (i.e., U.S. patents, journals, transactions, reputable magazines, book chapters, conferences, and conference proceedings) and commercial products. His current research interests include cyber security, secure future internet, artificial intelligence (i.e., machine learning, deep learning, and reinforcement learning), large-scale distributed systems (i.e., edge, fog, cloud, and SDNs), the IoT, the IIoT, QoS/QoE, affective computing, industry 4.0, and the Internet of Everything (IoE). He is a member of technical program committees of varied reputable conferences, journals, and editorial boards. He is a Professional Member ACM, with 15 years of research and development experience both in ICT industry and academia.

Open Access funding provided by 'Qatar National Library' within the CRUI CARE Agreement