Noname manuscript No.
(will be inserted by the editor)

DerainGAN: Single Image Deraining Using Wasserstein GAN

Sahil Yadav* · **Aryan Mehra*** · **Honnesh Rohmetra*** · **Rahul Ratnakumar** †‡ · **Pratik Narang** *

Received: date / Accepted: date

Abstract Rainy weather greatly affects the visibility of salient objects and scene in the captured images and videos. The object/scene visibility varies with the type of rain drops, i.e., adherent rain droplets on camera lens, streaks, rain and mist, etc. Moreover, they pose multifaceted challenges to detect and remove the rain drops to reconstruct the rain free image. Recently, both CNN and GAN based models have been designed to remove rain droplets from a single image. In this paper, we design a simple yet effective GAN framework to achieve improved deraining performance over the existing state-of-the-art methods. The learning is based on a Wasserstein GAN and perceptual loss. We empirically analyze the effect of different parameter choices to train the model for better optimization. We also identify the strengths and limitations of GAN based models for single image deraining by performing multiple ablation studies. The robustness of the proposed method is evaluated over two synthetic and one real-world rainy image datasets. The proposed DerainGAN significantly outperforms both the CNN and GAN based state-of-the-art models in all the datasets both in structural similarity and visual appearance while reducing computation time of generated images.

Keywords Image deraining · GAN · Deep learning

1 Introduction

One of the most challenging problems faced by the photographers as well as imaging experts is the degradation of the image quality due to mist, rain streaks and rain drops. The highly reflective behavior of rain to visible light and infrared radiations results in the presence of bright streaks in the image, which not only reduces its quality but also makes it difficult to comprehend [31]. This is especially true for the field of Computer vision and Digital Image Processing. Therefore, the removal of rain streaks has been a challenge for photographers, image processing experts

* Department of CSIS, BITS Pilani, Rajasthan, India
† Department of ECE, MNIT, Jaipur, Rajasthan, India
‡ Department of Electronics Engg., KNIT, Sultanpur, U.P., India

as well as forensic-scientists. In 2005, Garg and Nayar [13] threw light on an interesting fact that the intensity fluctuations caused due to rain is dependent on various factors like (i) camera parameters (time of exposure, depth of field, etc.), (ii) rain properties (drop size, velocity, density etc.) and (iii) brightness of the scene. They revealed how the visual effects of the rain can be reduced or enhanced by timely tuning the camera parameters. Hence the images degraded due to rain are quintessential to be restored for important computer vision tasks such as semantic segmentation, object detection, tracking, etc. for applications in outdoor surveillance and intelligent vehicles.

Owing to the exponentially increasing usage of smartphone camera and gadgets, many indoor and outdoor scenarios has become very common where a video or an image need to be captured through a window frame. For example, sitting inside a vehicle or building, a person may wish to capture the outside scene using his or her camera. And similarly in an indoor situation, exhibits showcased behind some protective medium like glass may need to be captured. Moreover, almost all of the outside mounted cameras are enclosed within a transparent window in order to protect them from the natural elements and collisions. The images captured by these cameras can be affected by attenuation, reflection, refraction and scattering etc. caused due to rain or presence of adherent rain droplets, resulting from their exposure to rain or cleaning procedures. The classical solution in these scenarios has been to de-focus these invisible points at the capture phase. But this requires direct vertical camera placements and many delicate adjustments like increasing the aperture, lens etc. thus adjusting the depth of the field [8]. In practice, however, these precautions are very hard to follow for amateurs. Moreover, many of these control features are obsolete in commonly available webcams and latest smartphone cameras.

Rain streaks corrupt the information within an image, and it thus affects the performance of most of the image processing algorithms which are based on the subtle features within the image. Almost all the existing image processing algorithms are developed based on the assumption that there is no disambiguity within the input data [11]. As most of the segmentation, tracking, recognition algorithms are tailor made for videos, rain streak removal within the crucial frames of a video has been very essential, especially in the recent days of emerging technologies like self-driving cars, intelligent UAVs and other automotive or surveillance applications involving computer vision. Due to the unpredictability of the rainy weather and large deployment of image processing systems, rain streak removal has been one of the important problems for a long time. Hence it's role in practical applications, such as object segmentation and detection, video surveillance, automotive arena etc. can never be overstated.

2 Background

A set of sample rainy and rain-free images obtained using our proposed deraining approach are illustrated in Fig. 1. Based on the input data, rain streak removal algorithms may be classified into single image based techniques and video based techniques [11].

Video Based Techniques. The earlier research works were focused on the removal of rain in videos, which often showed a very high correlation among the

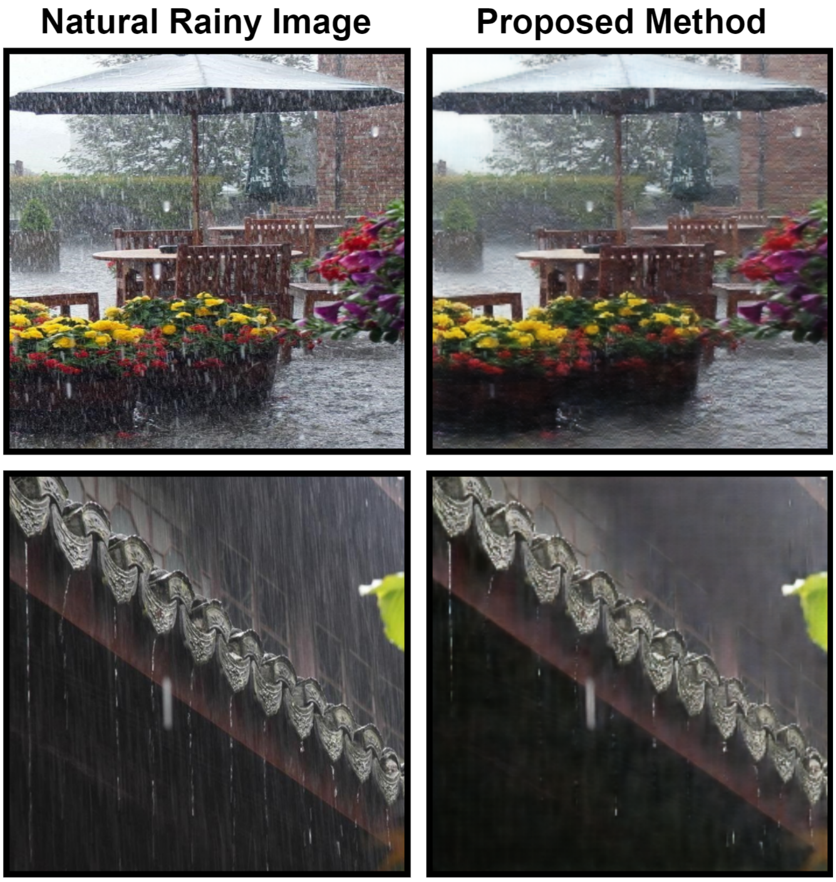

Fig. 1 Sample deraining results of the proposed DerainGAN on real/natural rainy images

nearby video frames. Utilizing both temporal and spatial data from video, rain streaks can be eliminated from a static background with the help of averaging neighboring frame intensities [12]. The use of Fourier transformations [3], and using GMMs (Gaussian mixture model) [5], matrix completion method [20] and low rank approximation techniques [6] are noteworthy in this domain.

Single Image Techniques. As the information is considerably reduced in the individual images, deraining process from a single image is a grueling task. A rainy image can be separated into two distinct parts mathematically: one related to the clear background image and the other to the rain streaks. Hence, the input rainy image can be written as

$$I = I_{clear} + I_{rain} \tag{1}$$

where I_{rain} represents the rain streaks and I_{clear} represents the clear background image. The problem of single image deraining can thus be viewed as decomposing the input image into a clear background image and the rain streaks.

In recent years, significant progress has been made in single image deraining. These methods can be categorized into hand-crafted methods [38], [19], [18], [39],

[6], [23], [24], [34], [26] Convolutional neural network (CNN)-based models [8], [33], [9], [10], [22], [11], [27], [7] and Generative Adversarial Network (GAN) based models [36], [37], [21], [4].

In the hand-crafted methods, both kernel-based and prior-based algorithms have been presented in the literature. Initially Zheng et al. in 2013 [38] used multiple guided filters for deraining. They proposed a method for snow and rain removal from a single image by analyzing the dissimilarity between the high frequency components like back ground edges and rain streaks, with the low frequency parts within the image like non-rain and non-snow regions, which are to be preserved. Finally the image undergoes a guided filtering of low frequency components in order to remove all the residual rain and snow features based on the properties of clear background edges.

In the same year, Kim et al. [19] analyzed the rotation angle and aspect ratio of the elliptical kernel at each pixel to detect the raindrops. They proposed an adaptive single image rain streak removal algorithm using a non-local adaptive Means-filter. Their research into the characteristics features (orientation, rotation angle and aspect ratio) of the elliptical kernel at each location of the image, helped in accurately identifying the rain streaks. After detection, the streaks were eliminated by passing through the adaptive non-local means-filter.

Many methods intelligently exploits the clear image or rain type as priors [18], [39], [6], [23]. Kang et al. [18] reformulated the rainy image into low and high frequency components. The rain streaks were detected with sparse coding in the high frequency components. Based on the MCA (Morphological Component Analysis) authors formulated the problem of removing rain from a single image as an 'Image decomposition problem'. Initially they used a bilateral filter to decompose the input image into low and high frequency components. The high frequency part thus obtained is again divided into 'non-rain component' and 'rain component' – applying sparse coding and automatic dictionary learning. This technique neither distorted the geometrical features of the image nor did it require 'temporal or motion' information from successive related images.

Similarly, Chen and Hsu [6] used the low rank priors to decompose the original image and rain streak. They generalized a low rank matrix model to a tensor structure with the idea of capturing the spatio-temporal correlation between the rain streaks. By this unified way, they were able to eliminate the rain streaks from image/video and other higher dimension input. Yu Luo [24] introduced an approach based on discriminative sparse coding considering the nonlinear generative model of the single rain image, which is also called as screen blend model. Their algorithm is based on dictionary learning, with the idea of sparsely approximating the patches of 2 layers- the rain image and derained image, by large discriminative codes onto the dictionary learning with very good mutual-exclusivity properties. They showed that these sparse codes can help in separating these two layers from its nonlinear composite, more accurately. Zhu et al. [39] modelled the rain streaks with a limited range of directions. They introduced a method of joint bi-layer optimization, in which- after an initial automatic location detection of the rain-dominated regions, gradient statistics were applied to estimate the direction of rain and to extract rain patches. This knowledge is utilized to model the three terms of regularization i.e. sparse representation, gradient deviation and visual similarity between the rain patches, to separate the rain streaks from background details in a step by step manner.

Li et al. [23] divided the image into patch-based priors with Gaussian mixture models (GMMs) to improve the clarity of the image. These GMMs used simple patch-based priors for the rain and background layers, which could accommodate multiple scales and orientation of rain streaks to improve the deraining performance. These approaches depend upon the accurate modeling of priors. Authors applied this technique to the complex real-world rain images like eave, bush, lamp to extract superior results than the existing single image de-raining methods are offering. He-Zhang proposed a CCRR algorithm [34] – Convolutional Coding-based Rain Removal algorithm for the automatic removal of rain-streaks from a single rainy image. CCRR used convolutional filters and low rank learning filters, for working on a single whole image rather than multiple patches and solving the optimization problem by dividing the rainy image into many overlapping patches for learning local dictionaries. Their results were superior to their contemporaries as per the standard image quality metrics like the PSNR and SSIM.

An attentive generative network was designed by Qian et al. [26] in order to deal with the substantial presence of raindrops. They introduced visual attention into the discriminative and generative networks. The discriminative network primarily gives attention to the local consistency of the restored regions, whereas the generative network concentrates on raindrop regions. They tried to inject visual attention map to generative and discriminative recurrent network to retrieve a rain drop free image with the help of a contextual auto-encoder in an adversarial training scheme. The same discriminative neural network checks the global and local validity of the generated output image.

Recently, many convolutional neural network (CNN) based models have been designed to successfully restore the original image from the rainy image. One of the earlier notable works of rain streak removal using CNN was undertaken by Eigen et al. [8] in 2013. Eigen initiated use of CNNs to remove dirt and water droplets adhered to a camera lens or glass window. But when the raindrops are large or the rain streaks are dynamic, the output is a blurry image. They specially designed a CNN which was trained to map corrupted image patched to the clean patches, thus automatically capturing the characteristics features of water droplets and dirt in the input images. They also constructed specialized training sets for the same and got very satisfactory results. Yang et al. [33] proposed a multi-stream network to jointly learn the binary rain streak map, rain streak appearance map, and the clear image. They applied these models on benchmark datasets like Rain100L, Rain20L and Rain100H and have also compared with four standard baseline methods like JORDER and SRCNN, which is a common CNN baseline for image processing.

Xueyang Fu et. al in 2017 developed a convolutional neural network (CNN) termed as DerainNet [10], only on the high frequency (HF) details. This input hungry- DerainNet keeps learning the mapping function of the rainy and clean image layers and gave better results than the state of art techniques available in that period. They also solved the problem of inadequate ground truth of clean images by synthesizing pairs of rainy/clean images for neural learning and showed how these datasets helped in corresponding to real world images. In the same year, Yang et. al [9] introduced a new deep learning based method to eliminate the rain streaks from a single image, by the help of a region dependent rain image model for detection and simulation of heavy rains. Utilizing this model, a convolutional network is designed which detects rain regions and in turn provides additional information for its removal. Along with this accumulation and removal network a recurrent model is also developed to remove the streaks in a progressive manner.

Li et al. [22] tried a recurrent neural network approach for solving the single image deraining problem. A recurrent squeeze-and excitation (SE) based context aggregation network (CAN) was proposed where CAN was used to make a bigger receptive field and SE was utilized to initialize different alpha values. Ren et al. [27] exploited the recursive features by repeatedly unfolding a shallow ResNet module called as progressive ResNet (PRN) to progressively remove the rain drops. Further they tried to exploit the dependencies of deep network features by introducing a Progressive recurrent Network (PReNet). Thus they could reduce the number of network parameters without significant degradation in output performance.

By the end of 2019, Xueyang Fu et al. again proposed a simplified deep convolutional neural networks (CNNs) by using domain specific knowledge or Pyramid networks, where a Laplacian pyramid is used to predict the clean Gaussian pyramid [11]. The advantage of this technique was that by using a recursive and residual network structures, it has very less parameters (only 8K) and still was able to achieve better results than its predecessors in the field of rain removal for low and high-level vision tasks.

Recently, Yingjun Du et al. [7] came up with a Conditional Variational Image Deraining (CVID) network utilizing the inherent and strong generative ability of the CVAE (Conditional Variational Auto Encoder) framework for modeling the latent distributions of clean-image-priors, with the help of which many rainy images are generated. For achieving better adaptive image deraining in multi-color channels the manuscript introduced a separate module for the spatial density estimation on uneven images blurred by the rain streaks.

More recently, GAN based models have also been used to generate rain-free images. In the year 2018, He Zhang [35] proposed a CNN based DID-MDN- (Deraining method with multi-stream densely connected network) which is density-aware method which uses the rain-density estimate for performing the synthesis of the image cleared of rain streaks. The authors also introduce a residual-aware rain-density classier to predict the rain density labels.

In the very next year, Zhang et al.[36] used the conditional GAN with DenseNet based generator for rain-drop removal. Again in 2019 Zheng et al.[37] designed a GAN model to jointly learn the gradient and noise-free image for rain and reflection removal. They came up with an adversarial learning method for de-noising which followed the usual procedure of guided inference of the noise free image after a prior estimation of the gradient map image, for dividing it into positive and negative samples. The final step was running a minmax optimization algorithm for a joint training of the dual optimizers for coming up with the estimated inference and gradient of the noise free image.

Xiaojun Bi et al. [4] proposed a new Multi-scale Weighted fusion spatial Attentive Generative Adversarial Network (MWA-GANet) in order to retain the background features while eliminating the rain streaks. This adversarial network performs multi-scale weighted fusion using attentive generative techniques to convert the rain streak feature map to generate an attention map for guiding the contextual auto-encoder in forming the background image. In the next stage, a discriminative network is used to classify if the image can be real or imperfect. The existing GAN based models perform reasonably well over certain rain drop removal datasets. However, these models demand careful network design choices for both generator and discriminator. Moreover, the selection of model hyper-parameters highly influences the model training convergence.

A more detailed literature review of the existing deraining methods is presented in [21], wherein the authors analyzed the different deraining algorithms available for different problems like rain-streak, rain-drop, and mist detection with the help of benchmark analysis. They applied elimination by applying different criteria for evaluation, such asfull and no-reference objective, task-specific and subjective evaluation techniques.

There are two major challenges faced by almost all current optimization based methods [16] (i) The true composition of the image cannot be perfectly retrieved by the composite models, and therefore regularizers are not satisfactory for widely varying scenarios (ii) The difficulty in defining an optimal derained output image for any real rainy input image, which is due to inherent nature of the process itself – generating (learning) a mapping function and producing deterministic output images [7].

Proposed work and contributions. In this paper, we aim to analyze the effectiveness of a simple well-known generator and discriminator-based Wasserstein GAN (WGAN) framework for rain drop removal with perceptual loss. To learn a robust mapping function between the rainy and clean image, both high-level semantic features and rich local-spatial information are required. Therefore, we use the feature aggregating structure of U-Net [28] as the generator to fuse local-spatial and global-semantic information across same scale layers. The addition of perceptual loss increases the visual capacity of the model while the WGAN is a more stable variant of a regular GAN. Our paper is an effort to robustly remove rain streaks and droplets from a single image using the WGAN framework. The contributions of this work are as follows:

1) We propose to use a simple yet effective U-Net structure as generator which combines same-scale features to represent both local and globally salient information. Moreover, to complement the generator, we use a simple discriminator to produce realistic rain-free image.

2) We analyze the effect of various network design and hyper-parameter changes over the performance of the proposed WGAN-based deraining model through ablation studies.

3) The proposed model produces quantitatively and qualitatively superior results for rain drop removal over synthetic and real-world images.

3 Proposed Method

We focus on exploiting well established principles in deep learning to improve rain drop removal performance by making small adjustments to and fine-tuning the existing CNN models. Thus, instead of designing complex network architectures for the generator and discriminator, we re-purpose a U-Net network to be used as the generator and a simple network as the discriminator. We focus on analyzing the potency of adversarial training for rain-drop removal by keeping other causal factors with minimal complexity. The generator and discriminator used in the proposed DerainGAN framework is shown in Fig. 2 and Fig. 3 respectively.

Generator. Since the raindrops in the image can be of different sizes, shapes and appearances, it is important for the network to identify features at multiple

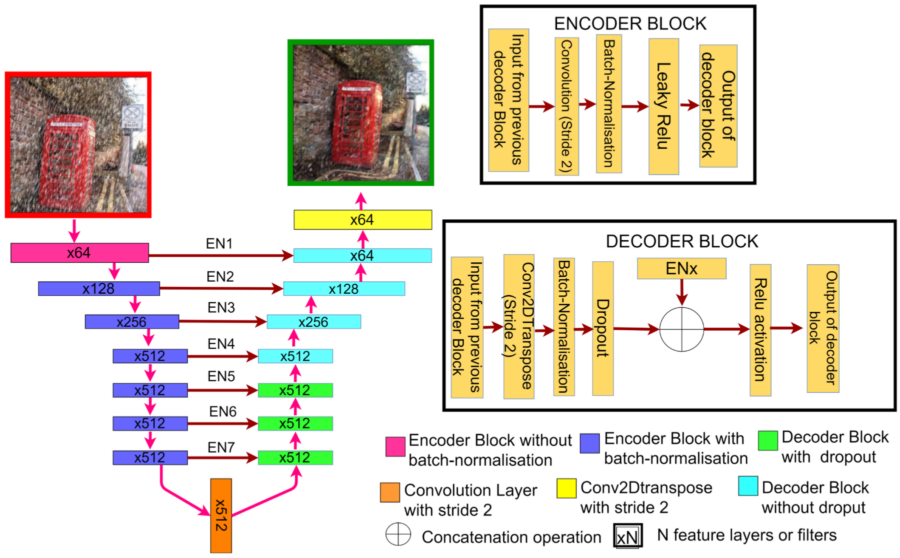

Fig. 2 Architecture of generator in the proposed method. DerainGAN encoder contains one 64 convolution layer without batch-normalization followed by six blocks with batch-normalization. The decoder part consists of three Conv2DTranspose layers with dropout followed by four without dropout. Each such block contains batch-normalization and $ReLU$ activation

scales. Therefore, to robustly remove the raindrops, networks with multi-scale structure are more suitable to capture complex geometric structures. We use a U-Net based network as a generator to extract features from the rainy image. The generator encodes strong semantic features to recognize the rain drops as well as learn the detailed visual patterns from the image. The same-scale connections at multiple stages of the network, as shown in Fig. 2, combine both low-level and high-level features to reconstruct the clean images more accurately. The detailed generator architecture is depicted in Fig. 2.

Discriminator. The GAN model is based on finding a competent adversary to the generator in order to classify the images as real or fake and push the generator to produce visually appealing rain-free images that are almost identical to the clean image. Hence, a 32×32 patch discriminator is used in the model. The patch size refers to the receptive field of the discriminator network. As pointed out by Isola et al. [15], using a smaller patch, like 1×1 (PixelGAN [15]) produces blurred images which makes the already degraded rainy images incomprehensible. Similarly, using a 128×128 patch (ImageGAN [15]) also decreases the quality of the images.

The discriminator takes two images stacked together which is passed through the network as shown in Fig. 3. The final layer of the discriminator uses the sigmoid activation which gives the probability of an image pair being a real (1) or fake (0) one.

Loss function. The loss function used for the model can be defined as a combination of the adversarial loss and perceptual loss:

$$L = L_{adv} + \lambda L_{per} \tag{2}$$

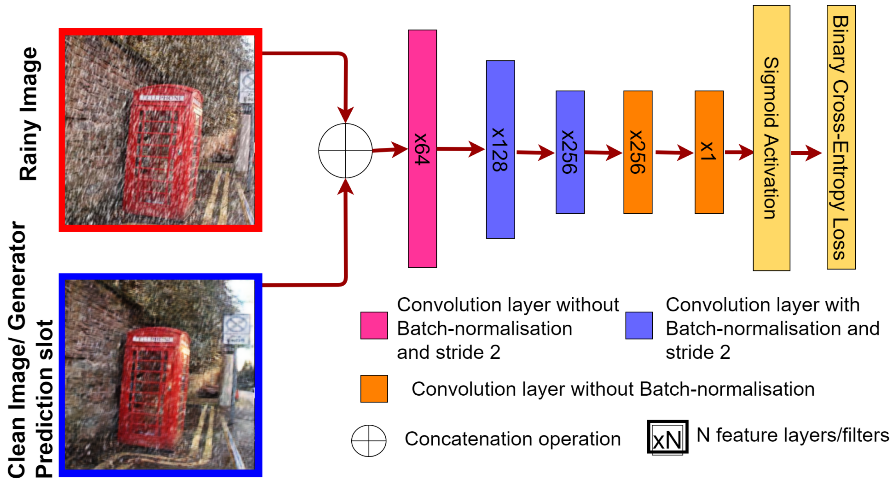

Fig. 3 Architecture of discriminator in the proposed method. The discriminator contains one strided convolution layer without batch-normalization with stride 2 followed by two Convolution layers with batch-normalization and two more layers without batch-normalization.

where λ is taken to be 100 throughout the experiment. GANs are generally prone to mode collapses and are hard to train. In order to mitigate this issue, we use Wasserstein loss [2] for training the generator. This helps the generator to learn even when the discriminator is performing well and prevents it from mode collapsing. For the discriminator, perceptual loss [17] is used with pretrained weights from VGG19 [30]. The perceptual loss is defined as follows:

$$L_{per} = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(I^S)_{x,y} - \phi_{i,j}(G_{\theta_G}(I^B))_{x,y})^2 \tag{3}$$

where $W_{i,j}$ and $H_{i,j}$ are the dimensions of the feature maps, $\phi_{i,j}$ is the feature map from the j^{th} convolution within the VGG19 model before the i^{th} maxpooling layer. For our model, we use activations from $VGG_{3,3}$ convolution layer. This helps in removal of artifacts and enhances the visual performance of the image.

Parameter Setting. Our proposed generator contains one encoder and one decoder with skip connections. The down-sampling in encoder is done by a 4×4 convolution with stride 2, while the upsampling in the decoder is done using a symmetric 4×4 convolution transpose with stride 2. *LeakyReLU* activation is used in all the layers of the encoder while *tanh* activation is used in the last layer of the generator. Our patch discriminator uses a 4×4 convolution with stride 2 and *LeakyReLU* activation, with last layer having a *sigmoid* activation. The loss function used is mean average error for the generator and binary cross entropy for the discriminator.

Training Configuration. All the inputs and outputs are resized to 256×256 pixels. The network is trained on a NVIDIA Titan V GPU. Learning rate of Adam algorithm was 2×10^{-3}. At most 80 epochs were run with a batch size of 1.

Table 1 SSIM and PSNR values of the proposed method along with other models on the Rain700 dataset

Method	SSIM (%inc)	PSNR (%inc)
SPM [18]	0.7632 (7.455)	18.88 (27.91)
PRM [6]	0.7297 (12.388)	20.46 (18.03)
DSC [24]	0.5996 (36.774)	18.56 (30.11)
CNN [9]	0.6013 (36.387)	19.12 (26.30)
GMM [23]	0.7413 (10.629)	22.27 (8.44)
CCR [34]	0.7332 (11.852)	20.52 (17.69)
JORDER [33]	0.7525 (8.983)	21.09 (14.50)
DerainGAN	**0.8201** (−)	**24.15** (−)

4 Experiments and Results

4.1 Datasets

We use two synthetic rainy image datasets and one real rainy image dataset for evaluating the proposed method.

Synthetic Datasets. We use Rain100L [33] dataset which is selected from BSD-200 [25] with 200 pairs of images for training and 100 for testing. We also conduct experiments on Rain700 dataset [36] containing 700 images. Out of 700 images, 500 images are chosen from UCID dataset [29] and remaining 200 images are chosen from BSD-500 [1] training set. For Rain700 dataset, the model is trained and evaluated over 600 and 100 images respectively. We follow the benchmark train-test division for both the datasets in order to have a fair comparison with existing rain drop removal methods.

Real rain images dataset. In order to evaluate the robustness of the proposed method on real rainy images, we also produce results using Rain100L dataset [36]. Since ground truth images are not available for the real rainy images, we qualitatively compare our results with other state-of-the-art models as shown in Figure 5.

4.2 Quantitative Results

We use the Structural Similarity Index (SSIM) [32] and Peak Signal-Noise Ratio (PSNR) for quantitative evaluation of the models, both of which are widely considered the best metrics for evaluation of rain-drop removal algorithms. To evaluate the effectiveness of our method, we quantitatively compare our model with many existing state-of-the-art approaches for rain drop removal [9], [24], [14], [35], [36], [18], [6], [34], [33]. The comparative performance on the commonly used synthetic benchmark datasets Rain700 and Rain100L are shown in Tables 1 and 2. We also tabulate the percentage increase (denoted as %inc) achieved in SSIM and PSNR by the proposed method when compared to the each of the existing methods. From Tables 1 and Table 2, it is evident that the proposed method considerably outperforms other methods in terms of both metrics.

Table 2 SSIM and PSNR values of the proposed method along with other models on the Rain100L dataset

Method	SSIM (%inc)	PSNR (%inc)
SPM [18]	0.6991 (24.46)	23.13 (22.35)
DSC [24]	0.8663 (4.387)	24.16 (17.13)
CNN [9]	0.8142 (6.865)	23.70 (19.40)
DID-MDN [35]	0.8569 (1.54)	28.27 (0.10)
JCAS [14]	0.8520 (2.124)	28.54 (NA)
IDCGAN [36]	0.8186 (6.291)	23.39 (20.99)
DerainGAN	**0.8701** (–)	**28.30** (–)

More specifically, in Table 1, the margin of improvement in SSIM varies between 7.45% over SPM to 36.77% over DSC. Similarly, in PSNR, the margin varies between 8.44% over GMM to 27.91% over SPM. Our work also outperforms deep learning model [33] by an approximate margin of 0.067 and 9.65 in SSIM and PSNR, respectively. In Table 2, we witness a similar improvement in SSIM and PSNR in comparison with all the 6 existing state-of-the-art methods. More specifically, the proposed method outperforms the GAN-based model [36] by an approximate margin of 0.05 and 5 in SSIM and PSNR, respectively. Our model beats JCAS [14] in SSIM and achieves similar results in PSNR.

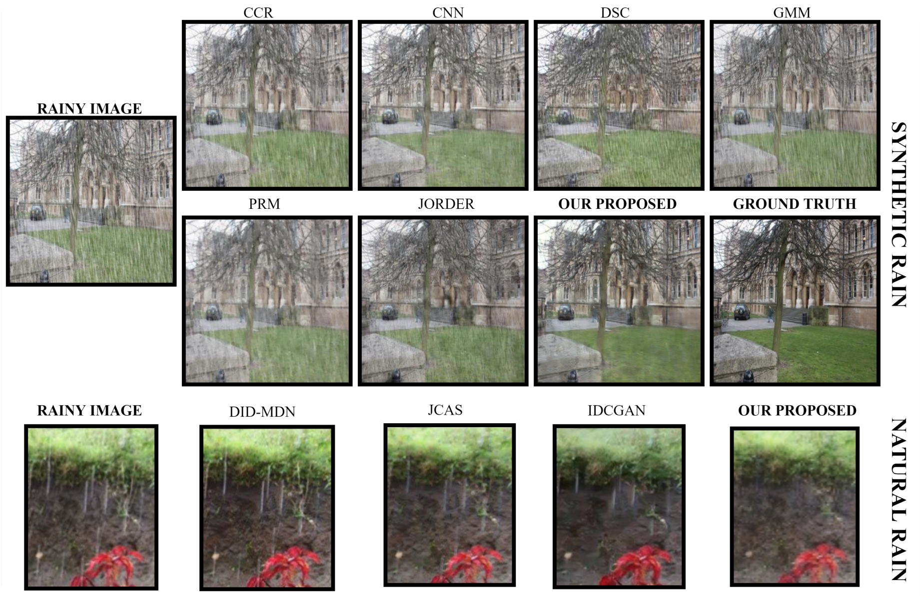

Fig. 4 Qualitative results on different methods and proposed GAN, for 1) synthetic rainy and 2) natural rainy images

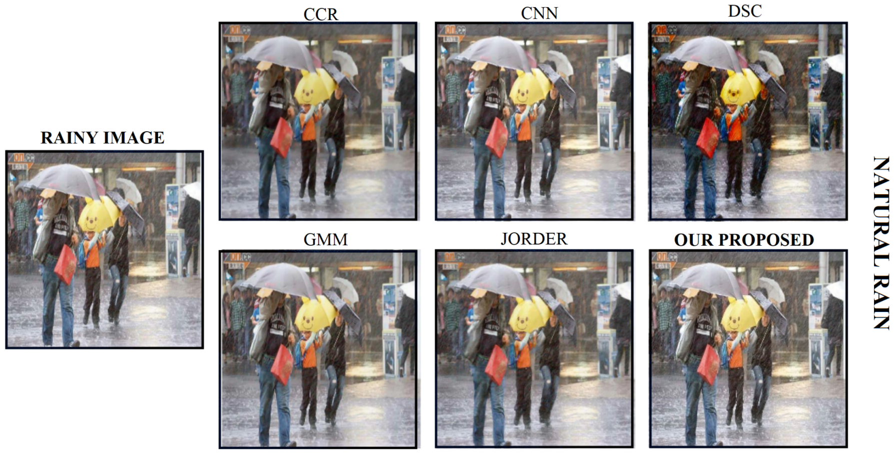

Fig. 5 Qualitative comparison on different methods and DerainGAN, for real rainy images

4.3 Qualitative Analysis

The proposed method generates sharp and realistic rain-free images when compared with the existing state-of-the-art approaches. Figure 5 depicts how the proposed method is more effective for both removing synthetic and natural rain from the images. The qualitative results in natural rain image in Fig. 5 show how the proposed method most effectively reduces the streaks of rain and gracefully merges the pixels with the background, thus reconstructing the image without any loss in context. A similar superiority of the proposed model can be observed for synthetic image as well. It is clear that the use of perceptual loss in the generator greatly enhances the visual quality of the image. Also, the use of Wasserstein loss for the discriminator prevents the occurrence of random black spots in the image and overall smooths the image, giving it a cleaner look. This proposed method also helps to deal with checkerboard artefacts common in GAN generated images. Using perceptual loss from VGG19, the model is able to extract universally applicable features and hence is able to better replace the rainy parts of the image compared to other models using per-pixel loss.

4.4 Time Complexity

Table 3 compares the running time of the proposed method with other state-of-the-art methods. Due to its simpleness and effectiveness, the proposed DerainGAN is computationally more efficient compared to all the other models. On average, the DerainGAN can process an image of size 256x256 in about 0.3s.

Table 3 Time Complexity (in seconds) of different methods

SPM[18]	PRM[6]	DSC[24]	CNN[9]	GMM[23]	CCR[34]	JORDER[33]	**DerainGAN**
400.5s	40.5s	1.3s	54.9s	169.6s	150.2s	0.4s	**0.3s**

Table 4 Ablation analysis of training GANs for deraining on Rain700 dataset

Trial Number/Description	Batch-Norm	Residual Modules	Perceptual Wasserstein Loss	SSIM
Trial 1 – W/o batch-normalization	N	N	N	0.7810
Trial 2 – With residual modules in generator	Y	Y	N	0.7700
Trail 3 – Shallow 2 layered discriminator	Y	N	N	0.7580
Trial 4 – Proposed GAN w/o perceptual and wasserstein loss	Y	N	N	0.8006
Proposed DerainGAN	Y	N	Y	**0.8201**

4.5 Ablation Study

Training adversarial networks, in general, is a difficult task because it involves training two networks that are dependent on each other to converge. With that in mind, an ablation study is conducted with different hyper-parameters and changes in network design for the proposed GAN framework. Table 4 shows how the removal of certain components of the network affects the performance in terms of SSIM. Since batch-normalisation adds to the compute at train and test time, it is essential to note that Trial 1 (without batch-normalisation) performs relatively poorly in comparison to the proposed implementation with batch-normalisation. Similarly, Trial 2 shows that adding residual modules to the generator makes it more complex but only worsens the SSIM. Shallowness in the discriminator had a similar effect. The proposed network gains an edge with the perceptual and wasserstein loss, clearly visible through comparison of Trial 4 and the proposed GAN. All this proves to show that the proposed network has optimum choices for effective and efficient image reconstruction and deraining.

5 Conclusion

This paper presented a Wasserstein GAN based framework for image deraining using a simple well-known generator and discriminator. The proposed GAN performed effective image deraining by learning a fusion of both high-level semantic features and rich local-spatial information, which are obtained using the feature aggregating structure of U-Net as the generator. The effectiveness of this generator coupled with a simple discriminator was also demonstrated with the help of an ablation study wherein the effect of various hyper-parameter or component changes was studied. The proposed GAN was consistently superior than its ablated counterparts. With detailed experiments on two synthetic and one real rain dataset, the proposed GAN delivered superior results – quantitatively (in terms of SSIM and PSNR) as well as qualitatively – for rain drop removal.

Acknowledgements This work is supported by BITS Additional Competitive Research Grant funding under Project Grant File no. PLN/AD/2018-19/5 for the Project titled "Disaster Monitoring from Aerial Imagery using Deep Learning".

References

1. Arbelaez P, Maire M, Fowlkes C, Malik J (2010) Contour detection and hierarchical image segmentation. IEEE transactions on pattern analysis and machine intelligence 33(5):898–916
2. Arjovsky M, Chintala S, Bottou L (2017) Wasserstein gan. arXiv preprint arXiv:170107875
3. Barnum PC, Narasimhan S, Kanade T (2010) Analysis of rain and snow in frequency space. International journal of computer vision 86(2-3):256
4. Bi X, Xing J (2020) Multi-scale weighted fusion attentive generative adversarial network for single image de-raining. IEEE Access 8:69838–69848
5. Bossu J, Hautière N, Tarel JP (2011) Rain or snow detection in image sequences through use of a histogram of orientation of streaks. International journal of computer vision 93(3):348–367
6. Chen YL, Hsu CT (2013) A generalized low-rank appearance model for spatio-temporally correlated rain streaks. In: Proceedings of the IEEE International Conference on Computer Vision, pp 1968–1975
7. Du Y, Xu J, Zhen X, Cheng MM, Shao L (2020) Conditional variational image deraining. IEEE Transactions on Image Processing
8. Eigen D, Krishnan D, Fergus R (2013) Restoring an image taken through a window covered with dirt or rain. In: Proceedings of the IEEE international conference on computer vision, pp 633–640
9. Fu X, Huang J, Ding X, Liao Y, Paisley J (2017) Clearing the skies: A deep network architecture for single-image rain removal. IEEE Transactions on Image Processing 26(6):2944–2956
10. Fu X, Huang J, Zeng D, Huang Y, Ding X, Paisley J (2017) Removing rain from single images via a deep detail network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 3855–3863
11. Fu X, Liang B, Huang Y, Ding X, Paisley J (2019) Lightweight pyramid networks for image deraining. IEEE transactions on neural networks and learning systems
12. Garg K, Nayar SK (2004) Detection and removal of rain from videos. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004., IEEE, vol 1, pp I–I
13. Garg K, Nayar SK (2005) When does a camera see rain? In: Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, IEEE, vol 2, pp 1067–1074
14. Gu S, Meng D, Zuo W, Zhang L (2017) Joint convolutional analysis and synthesis sparse representation for single image layer separation. In: Proceedings of the IEEE International Conference on Computer Vision, pp 1708–1716
15. Isola P, Zhu JY, Zhou T, Efros AA (2017) Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1125–1134

16. Jin X, Chen Z, Li W (2020) Ai-gan: Asynchronous interactive generative adversarial network for single image rain removal. Pattern Recognition 100:107143
17. Johnson J, Alahi A, Fei-Fei L (2016) Perceptual losses for real-time style transfer and super-resolution. In: European conference on computer vision, Springer, pp 694–711
18. Kang LW, Lin CW, Fu YH (2011) Automatic single-image-based rain streaks removal via image decomposition. IEEE Transactions on Image Processing 21(4):1742–1755
19. Kim JH, Lee C, Sim JY, Kim CS (2013) Single-image deraining using an adaptive nonlocal means filter. In: 2013 IEEE International Conference on Image Processing, IEEE, pp 914–917
20. Kim JH, Sim JY, Kim CS (2015) Video deraining and desnowing using temporal correlation and low-rank matrix completion. IEEE Transactions on Image Processing 24(9):2658–2670
21. Li S, Araujo IB, Ren W, Wang Z, Tokuda EK, Junior RH, Cesar-Junior R, Zhang J, Guo X, Cao X (2019) Single image deraining: A comprehensive benchmark analysis. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 3838–3847
22. Li X, Wu J, Lin Z, Liu H, Zha H (2018) Recurrent squeeze-and-excitation context aggregation net for single image deraining. In: Proceedings of the European Conference on Computer Vision (ECCV), pp 254–269
23. Li Y, Tan RT, Guo X, Lu J, Brown MS (2016) Rain streak removal using layer priors. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2736–2744
24. Luo Y, Xu Y, Ji H (2015) Removing rain from a single image via discriminative sparse coding. In: Proceedings of the IEEE International Conference on Computer Vision, pp 3397–3405
25. Martin D, Fowlkes C, Tal D, Malik J, et al. (2001) A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. Iccv Vancouver:
26. Qian R, Tan RT, Yang W, Su J, Liu J (2018) Attentive generative adversarial network for raindrop removal from a single image. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 2482–2491
27. Ren D, Zuo W, Hu Q, Zhu P, Meng D (2019) Progressive image deraining networks: a better and simpler baseline. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 3937–3946
28. Ronneberger O, Fischer P, Brox T (2015) U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention, Springer, pp 234–241
29. Schaefer G, Stich M (2003) Ucid: An uncompressed color image database. In: Storage and Retrieval Methods and Applications for Multimedia 2004, International Society for Optics and Photonics, vol 5307, pp 472–480
30. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:14091556
31. Wang Y, Liu S, Xie D, Zeng B (2020) Removing rain streaks by a linear model. IEEE Access 8:54802–54815
32. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP, et al. (2004) Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing 13(4):600–612

33. Yang W, Tan RT, Feng J, Liu J, Guo Z, Yan S (2017) Deep joint rain detection and removal from a single image. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 1357–1366
34. Zhang H, Patel VM (2017) Convolutional sparse and low-rank coding-based rain streak removal. In: 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), IEEE, pp 1259–1267
35. Zhang H, Patel VM (2018) Density-aware single image de-raining using a multi-stream dense network. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 695–704
36. Zhang H, Sindagi V, Patel VM (2019) Image de-raining using a conditional generative adversarial network. IEEE Transactions on Circuits and Systems for Video Technology
37. Zheng Q, Shi B, Jiang X, Duan LY, Kot AC (2019) Denoising adversarial networks for rain removal and reflection removal. In: 2019 IEEE International Conference on Image Processing (ICIP), IEEE, pp 2766–2770
38. Zheng X, Liao Y, Guo W, Fu X, Ding X (2013) Single-image-based rain and snow removal using multi-guided filter. In: International Conference on Neural Information Processing, Springer, pp 258–265
39. Zhu L, Fu CW, Lischinski D, Heng PA (2017) Joint bi-layer optimization for single-image rain streak removal. In: Proceedings of the IEEE international conference on computer vision, pp 2526–2534