This is the author's final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

Deposited on: 30 October 2018

Removal of Visual Disruption Caused by Rain using Cycle-Consistent Generative Adversarial Networks

Lai Meng Tang, Li Hong Lim, and Paul Siebert

University of Glasgow, Glasgow G12 8QQ, UK

Abstract. This paper addresses the problem of removing rain disruption from images for outdoor vision systems. The Cycle-Consistent Generative Adversarial Network (CycleGAN) is proposed as a more promising rain removal algorithm, as compared to the state-of-the-art Image Deraining Conditional Generative Adversarial Network (ID-CGAN). The CycleGAN has an advantage in its ability to learn the underlying relationship between the rain and rain-free domain without the need of paired domain examples. Based on rain physical properties and its various phenomena, five broad categories of real rain distortions are proposed in this paper. For a fair comparison, both networks were trained on the same set of synthesized rain-and-ground-truth image-pairs provided by the ID-CGAN work, and subsequently tested on real rain images which fall broadly under these five categories. The comparison results demonstrated that the CycleGAN is superior in removing real rain distortions.

Keywords: CycleGAN, ID-CGAN, generative adversarial network.

1 Introduction

It has been widely acknowledged that severe weather effects caused by rain can badly affect the performance of many computer vision algorithms. This is primarily due to the fact that these algorithms are typically trained using images which have been captured under well-controlled conditions. Rain can manifest itself in the form of low contrast, blurred and distorted scene content, and highly saturated image specularities produced by falling raindrops, which are always brighter than the original background [12]. Figure 1 shows two examples of rain distorted images and their rain removal results using the Cycle-Consistent Generative Adversarial Networks (CycleGAN)[42] algorithm. Our rain images and their removal results clearly show the potential for automatic rain removal from the types of image captured when undertaking many real outdoor tasks in computer vision. One such example is drivable path detection for an autonomous driving system [27], which must be able to detect both drivable and non-drivable paths for successful navigation [1]. Some researchers using camera sensors have already been successful to some extent in drivable path detection. Regardless of this achievement, environmental noise such as rain and/or snow can cause

misdetection of drivable path which can lead to autonomous driving system accident. This is because environmental noises have the capability to affect the color properties of the image with significant effects of misclassification of road as non-road and vice versa [29].

Fig. 1. Rain distortion on outdoor images as shown in Figure 1(a) and (b), and their rain removal results by the CycleGAN were shown in Figure 1(c) and (d) respectively.

This paper makes a number of fundamental contributions to removing real rain effects from images. The CycleGAN [42] is proposed for the first time as a practical and effective way to reconstruct images under visual disruption caused by real rain. The proposed approach does not require synthetically generated paired rain and rain-free training data for learning, as required by other GANs methods ([41], [34]), to address the disruption problem posed by real rain. In other words, the proposed CycleGAN [42] has the distinct advantage compared to the other GAN methods in its ability to learn a mapping from an input rain image to a rain-free image, in the absence of paired rain and rain-free (ground truth) training examples for real-world image reconstruction tasks. Hence, issues of the practicality in collecting similar aligned rain and rain-free image pairs to train the GAN algorithms and the unproven assumption that synthetically generated rain streaks represent real rain, can be addressed by the CycleGAN.

Based on a rain physics model [12], removal of five broad categories of real rain distortion is proposed and this methodology can be applied to the majority of outdoor rain conditions. Using a synthetic training data set provided by a recent ID-CGAN study [41], we have evaluated the proposed CycleGAN[42] network against the previously reported state-of-the-art ID-CGAN [41], in terms of its rain removal performance on real rain test data sets. Our results demonstrate that the above CycleGAN [42] was able to remove all types of rain distortion better than the ID-CGAN [41] algorithm and therefore the CycleGAN represents the new state-of-the-art for the removal of real rain distortion in images.

This paper is organized as follows. A brief overview of existing single image rain removal techniques is given in Section II. In Section III, five categories of rain distortion we propose to address by rain removal algorithms are discussed, based on the physical properties and various types of rain phenomena mentioned in [12]. The network, training and testing of the proposed CycleGAN [42] method are presented in Section IV. The results of the rain removal experiments for both the state-of-the-art ID-CGAN [41] and the proposed CycleGAN [42] algorithms were then analyzed both qualitatively and quantitatively in Section V. Section VI concludes the paper with a brief summary and discussion.

2   Background and related work

In past decades, many researchers have focused on image recovery from video sequences by taking advantage of the additional temporal information contained within rain video, as the actual scene content is not occluded by rain in every video frame of any given sequence ([12], [11],[22],[15],[2],[35]). However, single-image rain detection and removal is a more challenging task, compared to multi-frame based techniques, due to the lack of predictable spatial and temporal information that can be obtained by comparing successive image frames in order to compute rain physics and statistical models [12]. In order to tackle this ill-posed problem, many early single-image based methods considered signal or layer separation ([36], [38], [8], [19], [20], [7], [5], [6], [26], [37], [30], [40], [9], [10], [32], [33]), or relied on rain properties as image priors to detect rain patches, to allow filtering methods to be applied to remove rain ([4], [16], [31], [36], [38]).

Early single-image based methods include sparse coding or morphological component analysis based dictionary learning methods ([8], [19], [20], [7], [5], [6], [26], [37], [30], [40]), and rain prior approaches based on rain prior information or properties ([18], [17], [4], [16], [31], [36], [38]), for single-image rain removal. Due to their common assumption of rain streaks having similar patterns and orientations in the high frequency components of the image, some success in applying both dictionary learning and rain prior approaches has been observed. In more recently reported research using Deep Learning approaches, such as Convolutional Neural Networks (CNN) ([9], [10], [32], [33]) and Generative Adversarial Networks (GAN) ([41], [34]), it was highlighted that early

approaches suffered from incomplete removal of rain streaks and unintended removal of certain global repetitive patterns, such as brick and texture, as well as rain-free components, from the background image.

Due to their ability to learn end-to-end mappings, Convolutional Neural networks (CNN) approaches have been successfully applied by researchers to tackle the rain removal problem ([9], [10], [32], [33]). Their results have shown that CNN-based approaches can out-perform other non-Deep Learning based methods, particularly when applied to heavy rain images. However, these approaches failed to remove rain streaks completely, in contrast to the Generative Adversarial Networks (GANs) approach, recently introduced by Zhang et al. [41].

Inspired by the recent success of GAN- [14],[21] for pixel-level vision tasks such as image generation [24],[25], image inpainting [23] and image super-resolution[3], the GAN approach seemed natural and promising in removing rain streaks from a single image without affecting the background scene details. Using the discriminative model in GANs to ensure that the rain-removed reconstructed images are indistinguishable from their original ground truth counterparts, Zhang et al. [41] recently introduced their special conditional GAN called the Image De-raining Conditional General Adversarial Network (ID-CGAN) to remove rain streaks from a single image, which was inspired by the success of the general purpose Conditional General Adversarial Network (CGAN) solution proposed by Isola et al. [24] for image-to-image translation such as mapping an object edges to its photo, semantic labels to a scenes image, etc. Apart from learning a mapping function, Isola et al. [24] argued that the network also learnt a loss function, eliminating the need for specifying or hand-designing a task-specific loss function. Instead of using a decomposition framework to address the single image rain removal problem, the ID-CGANs [41] framework is based on the CGANs network to directly learn a mapping from an input rain image to a rain-removed (background) image.

The ID-CGAN consists of two models: a generator model (G) and a discriminator model (D). The generator model acts as a mapping function to translate an input image corrupted with rain to a reconstructed rain-removed image such that it fools the discriminator model which is trained to distinguish rain images from images without rain. In other words, by directly incorporating this criterion into the optimization framework, the CGAN approach ensures that rain-removed images are indistinguishable by a given discriminator from their corresponding ground truth images. In addition, a perceptual loss function is also defined in their optimization function to ensure the visual appeal of the end result using the ID-CGAN [41]. Due to the above contributions, the ID-CGAN now represents the state-of-the-art method for rain removal in a single image.

In the next section, five categories of rain distortion phenomena that we propose to address using the rain removal algorithms we discuss. This is followed by

the introduction of the Cycle-Consistent Generative Adversarial Network (Cy-cleGAN) [42], as a practical method to address the single-image rain degradation problem, due to it's unique ability to learn a mapping from an input rain image to a rain-free (ground truth) image, in the absence of paired rain and rain-free training examples.

3   Rain distortions on images

There are various detrimental effects of rain on images ,which can be analyzed using rain physics models [12]. Rain brings complex intensity changes due to its unique physical properties: its small size, high velocity and wide spatial distribution. In addition, rain consists of large numbers of drops falling at high speeds (terminal velocity) in the same direction. Typically, raindrops are water droplets of size between 0.1-10mm, with a wide distribution of size. The drops that make up a significant fraction of rain are less than 1mm in size, which are not severely distorted and their shapes can be well approximated and modelled as transparent spheres of water.Individual raindrops are distributed randomly and uniformly in 3D volume. The probability $P(k)$ that k number of drops exist in a volume V is given by a Poisson distribution [12]. These rain physics models have been used to detect and remove rain effect from images.

The terminal velocity $\overrightarrow{v}$ of a raindrop is directly proportional to the square root of its radius a [12] as shown by the equation:

$$\overrightarrow{v} = 200\sqrt{a} \tag{1}$$

To achieve this terminal velocity(constant maximum velocity) which is typically between 5-9m/s in air, raindrops need distances of at least 12m to accelerate to terminal velocity [12]. For free-fall rain drops travelling at their terminal velocities, motion-blur significantly affects the appearance of rain. Garg and Nayar ([12], [11]) examined the irradiance of the pixel over the exposure (integration) duration T and showed that the time τ that a drop projects onto a pixel is far less than exposure (integration) duration T, i.e., the maximum value of τ is approximately 1.18ms, which is much less than the typical exposure time $T \approx 30ms$ of a video camera. The short exposures produced stationary and bright raindrops and they do not appear transparent. However, at long exposures, due to fast motion, raindrops produce severely motion-blurred rain streaks and makes it look transparent.

Garg and Nayar [12] derived the dependence of rain visibility (intensity standard deviation of acvolume of rain, $\sigma_r(I)$) on rain properties ($k_0 \frac{a^2 \sqrt{\rho}}{\sqrt{v}}$), scene properties (($L_r - L_b$))and camera parameters ($\frac{\sqrt{G(f,N,z_0)}}{\sqrt{T}}$) as:

$$\sigma_r(I) = \int \sigma_r(I,z)\mathrm{d}z = k_0 \frac{a^2\sqrt{\rho}}{\sqrt{v}}(L_r - L_b)\frac{\sqrt{G(f,N,z_0)}}{\sqrt{T}}, \tag{2}$$

where σ is the standard deviation, I is the rain pixel's intensity, z is the distance of the rain drops in front of the camera, k_0 is a constant camera gain, ρ is the rain water density, a and v are the rain drops' radius and velocity respectively, L_r and L_b are the rain drops' and background's radiance respectively, and G is a camera function defined by the focal length f, the F-number N, the Focus Distance z_0 and the exposure time T.

Based on the above physical properties and various types of rain phenomena, this paper proposes five broad categories of real rain distortions, which can be applied to the majority of outdoor rain conditions. The effectiveness of rain removal algorithms can then be evaluated, based on these five categories, which are listed as follows.:

i.) different severity of rain streaks [33];
ii.) different camera settings ([12], [11]);
iii.) rain images taken in an indoor setting behind a glass window;
iv.) rain velocity reduction and splashing at obstructing structures; and
v.) splashing and accumulation of rain water on ground surface.

4   Network, training and testing

4.1   Network parameter and training data set

The same architecture of the CycleGAN networks from Zhu et. al. [42] is proposed in this paper to remove real rain disruptions as it has shown impressive results for general purpose unpaired image-to-image style transfer, object transfiguration, season transfer and photo enhancement. For this purpose, the general purpose CycleGAN [42] can be used to translate an image from a source domain (rain) X to a target domain (rain-free) Y. Its goal is to learn a mapping $G : X \rightarrow Y$ such that the distribution of images from $G(X)$ is indistinguishable from the distribution Y using an adversarial loss. This highly under-constrained mapping is coupled with an inverse mapping $F : Y \rightarrow X$, thereby using the cycle consistency loss introduced to push $F(G(X)) \approx X$ (and vice versa). Combining this cycle consistency loss ($\mathcal{L}_{\mathrm{cyc}}(G, F)$) with adversarial losses on domains X ($\mathcal{L}_{\mathrm{GAN}}(G, D_Y, X, Y)$) and Y ($\mathcal{L}_{\mathrm{GAN}}(F, D_X, Y, X)$) yields the overall CycleGAN objective expressed as:

$$\mathcal{L}(G, F, D_X, D_Y) = \mathcal{L}_{\mathrm{GAN}}(G, D_Y, X, Y)$$
$$+ \mathcal{L}_{\mathrm{GAN}}(F, D_X, Y, X)$$
$$+ \lambda \mathcal{L}_{\mathrm{cyc}}(G, F), \qquad (3)$$

where D_Y is the associated adversarial discriminator that encourage G to translate X into outputs indistinguishable from domain Y, and vice versa for D_X, and λ is a constant that control the relative importance of the two objective

functions G and F in the cycle consistency loss.

The objective of the forward mapping function $G : X \rightarrow Y$, $\mathcal{L}_{\text{GAN}}(G, D_Y, X, Y)$, is expressed as $\mathcal{L}_{\text{LSGAN}}(G, D_Y, X, Y)$ for training stability reason:

$$
\begin{aligned}
\mathcal{L}_{\text{LSGAN}}(G, D_Y, X, Y) =& \mathbb{E}_{y \sim p_{\text{data}}(y)}[(D_Y(y) - 1)^2] \\
&+ \mathbb{E}_{x \sim p_{\text{data}}(x)}[D_Y(G(x))^2],
\end{aligned}
\tag{4}
$$

where x and y are the images in X and Y domain respectively. A similar objective was used for the reverse mapping function $F : Y \rightarrow X$, $\mathcal{L}_{\text{GAN}}(F, D_X, Y, X)$.

The $\mathcal{L}_{\text{cyc}}(G, F)$ is expressed as:

$$
\begin{aligned}
\mathcal{L}_{\text{cyc}}(G, F) =& \mathbb{E}_{x \sim p_{\text{data}}(x)}[\|F(G(x)) - x\|_1] \\
&+ \mathbb{E}_{y \sim p_{\text{data}}(y)}[\|G(F(y)) - y\|_1].
\end{aligned}
\tag{5}
$$

The entire proposed CycleGAN network is trained on a Nvidia GTX 1080 using the Pytorch implementation [13]. Same as [42], this CycleGAN network was trained with a learning rate of 0.0002 for the first 100 epochs and linearly decaying rate to zero for the next 100 epochs. Weights were initialized from a Gaussian distribution with mean 0 and standard deviation 0.02. The discriminators D_X and D_Y were updated using a history of previously generated 50 images rather than the ones produced by the latest generative networks to stabilize the training procedure. For all experiments, λ, which controls the relative importance of the two objective functions G and F in the cycle consistency loss, was set to 10 in Equation (3). The Adam solver was used with a batch size of 1.

The same generator architecture from [42] was adopted to accomodate different image sizes for the rain-removal CycleGAN network. It contains two stride-2 convolutions, several residual blocks, and two fractionally-strided convolutions with stride $\frac{1}{2}$. For images with sizes 128×128, 6 residual blocks were used; and for images with sizes 256×256 and larger, 9 blocks were used. In addition, instance normalization was used.

Let $c7s1$-k denote a 7×7 Convolution-BatchNorm-ReLU layer with k filters and stride 1; and dk denotes a 3×3 Convolution-BatchNorm-ReLU layer with k filters, and stride 2, with reflection padding used to reduce artifacts. Rk denotes a residual block that contains two 3×3 convolutional layers with the same number of filters on both layer. uk denotes a 3×3 fractional-strided-Convolution-BatchNorm-ReLU layer with k filters, and stride $\frac{1}{2}$. Using similar naming convention used in [42] to describe the generator network architecture, the network with 6 blocks consists of:

$c7s1$-32,$d64, d128, R128, R128, R128, R128, R128, R128, u64, u32, c7s1$-3

and the network with 9 blocks consists of:

$c7s1$-32,$d64, d128, R128, R128, R128, R128, R128, R128, R128, R128, R128,$
$u64, u32, c7s1$-3

The same discriminator architecture from [42] was adopted to accomodate different image sizes for the rain-removal CycleGAN network as well. It used 70 × 70 PatchGANs [42] to try to classify whether the 70 × 70 overlapping image patches were real or fake. Such a patch-level discriminator architecture has fewer parameters than a full-image discriminator, and can be applied to arbitrarily-sized images in a fully convolutional fashion. In a similar naming convention as the Generator network, let Ck denote a 4 × 4 Convolution-BatchNorm-LeakyReLU layer with k filters and stride 2. BatchNorm was not used for the first $C64$ layer. Slope of 0.2 was used for the leaky ReLUs. After the last layer, a convolution was applied to produce a 1-dimensional output. The discriminator architecture is:

$C64$-$C128$-$C256$-$C512$

The CycleGAN approach does not require synthetically created rain and rain-free image pairs for training the networks. But for a fair comparison of the proposed CycleGAN rain removal approach with the state-of-the-art ID-CGAN, the same synthesized rain image pairs provided by [41] were used for the training of both networks. A total of 700 synthetically created rain-and-ground-truth image-pairs were used as training samples. All the training samples were resized to 256 × 256 for evaluation purpose. Note that although rain pixels of different intensities and orientations were added to generate this diverse training set, these "fake rain" added may not represent the "real rain" statistics in the natural images. Thus, both algorithms need to be tested using only real rain images. Real rain images with different types of distortions were collected to investigate the rain removal capability of the CycleGAN compared to the ID-CGAN visually, as explained in the next section.

4.2   Testing and evaluation data set

Existing objective image quality measures require some measurement of the closeness of a test image to its corresponding reference (ground truth). These measures are either based on mathematically defined measures such as the widely used mean squared error (MSE), peak signal to noise ratio (PSNR), structure similarity information measures (SSIM) [39], or the human visual system (HVS) based perceptual quality measures such as the visual information fidelity (VIF) [28]. Most existing literature used such generated image pairs for quantitative comparison of their results.

Based on the five types of rain distortions discussed in Section III, the qualitative performance of real rain removal performance was evaluated. As their real rain ground truth reference images were not available in the test data set,

the performance of the proposed CycleGAN and the ID-CGAN is evaluated visually. The results reveal the superiority of the proposed CycleGAN method. In addition, although the corresponding rain-free (ground truth) images for a real rain quantitative comparison of results were not available, the quantitative comparison results of a subset of images from the synthetic data set are used to compare their performance for a complete analysis, as shown in the next section.

5   Experiment results

5.1   Type I: Different severity of rain streaks

As discussed in Section III, rain drops show a wide distribution of size, volume and rate [12]. Hence, it is expected that rain properties affect the appearance of rain streaks in a wide variety of manner [12], [33]. Light rain streaks below 1 mm in rain drop's diameter are common; they are less visible and blur the background scene in a rain image. Heavy nearby rain streaks above 1 mm in rain drops' diameter are more visible and reduce the visibility by occluding the background scene. Severe distant rain with large rain drops' diameter show their individual rain streaks are overlapping and cannot be seen, occluding the background scene in a misty manner [33]. Two Type I sample images are shown in Figure 2 (a) and (c) with different severity of rain streaks, and their corresponding magnified rain streaks are as shown in Figure 2 (b) and (d) respectively. The rain removal results using the ID-CGAN and CycleGAN are as shown in Figure 2(e) to (h) and Figure 2(i) to (l) respectively. The subsequent figures for other types of rain distortion are presented in the similar manner.

As shown by the results, CycleGAN removed the rain streaks of different severity equally well, while the ID-CGAN was unable to remove the rain streaks and many original rain streaks remained, especially for heavy rain. In addition, it was observed that the contrast of background scenes was enhanced with the ID-CGAN.

5.2   Type II: Different camera settings

As discussed in Section III, camera parameters such as exposure time affect the visibility of the rain. Garg and Nayar [12] compared rain images taken with a short exposure time of $1ms$ and normal exposure time of $30ms$, and discovered that the short exposures produced stationary and bright raindrops and they do not appear transparent. However, at long exposures, due to fast motion, raindrops produce severely motion-blurred rain streaks. Type II distortion is typically due to a short exposure time that increases rain visibility and produces stationary, bright and non-transparent raindrops. Due to the high speed of rain, rain drops appear as bright spheres occluding the background scene. Figure 3(a) to (d) shows examples of such rain degradation. The rain removal results using the ID-CGAN and CycleGAN are as shown in Figure 3(e) to (h) and Figure 3(i)

Fig. 2. Type I distortion for different severity of rain streaks as shown in Figure 2(a) to (d), and their rain removal results by the ID-CGAN and CycleGAN were shown in Figure 2(e) to (h) and 2(i) to (l) respectively.

Fig. 3. Type II distortion for different camera setting as shown in Figure 3(a) to (d), and their rain removal results by the ID-CGAN and CycleGAN were shown in Figure 3(e) to (h) and 3(i) to (l) respectively.

to (l) respectively.

It was observed that the CycleGAN was able to remove the bright rain spheres well, although it was not trained to remove such a type of defect. In comparison, the ID-CGAN was unable to remove such defects and left behind many bright rain spheres in the zoomed regions-of-interest. This may be because such real rain defect are not covered in the synthetic training data set. In addition, ID-CGAN is known to suffer from white-round rain streaks due to the high-level features from CNN network inherently enhancing white round particles [41]. Hence, the CycleGAN performed better than the ID-CGAN for the rain distortions in Type II.

5.3 Type III: Indoor rain images behind a glass window

Since a glass window affect the radiance or scene properties of an image, it affects the visibility of rain streaks as shown by Equation (2), as shown in Section III. Hence rain streaks and its background scene viewed behind a transparent or translucent glass window should be considered separately as a different defect. The adherent rain water behind the glass window also occludes the rain streaks and its background scene. The reflection of light by, and the refraction of light through, the adherent water stain behind the glass window produces very low brightness scene captured by a camera or observed by a human, as shown in Figure 4 (a) to (d).

Fig. 4. Type III distortion for scene behind a glass window as shown in Figure 4(a) to (d), and their rain removal results by the ID-CGAN and CycleGAN were shown in Figure 4(e) to (h) and 4(i) to (l) respectively.

Figure 4(e) to (h) and Figure 4(i) to (l) show the results of removing rain using the ID-CGAN and CycleGAN respectively. As shown by the results, the ID-CGAN would brighten the adherent water drops as shown by Figure 4(e) and (f), regardless of the sizes of the drops. This may be due to the same reasons, as discussed in Type II. In comparison, the CycleGAN does not show such defects, as shown in Figure 4 (i) and (j). Also, as shown in Figure 4 (g) and (h), although the ID-CGAN managed to enhance the contrast of the low brightness background scenes, its contrast was still not as good as the CycleGAN, as shown in Figure 4 (k) and (l). This may be due to the nature of the learning of the cycle consistency objective that prevent the learned mappings G and F from contradicting each other, in such low brightness situations. Although none of the algorithms was trained to remove such defects, the CycleGAN has shown that it is more superior to remove such defect and manage to enhance the contrast of the low brightness scene well.
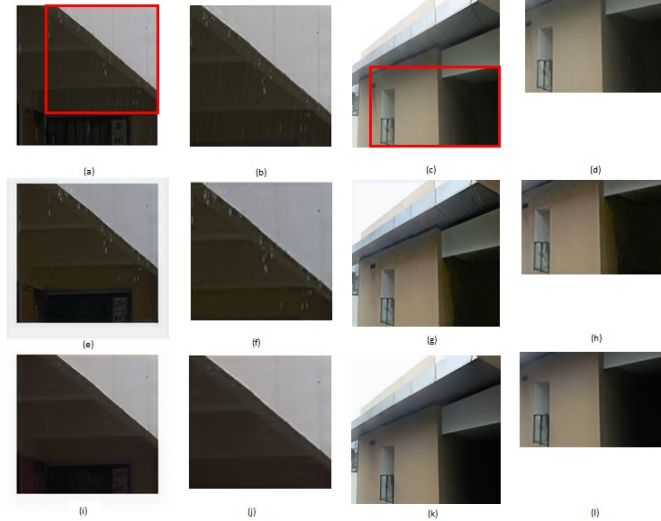

Fig. 5. Type IV distortion for rain velocity reduction and splashing at obstructing structures as shown in Figure 5(a) to (d), and their rain removal results by the ID-CGAN and CycleGAN were shown in Figure 5(e) to (h) and 5(i) to (l) respectively.

5.4   Type IV: Rain velocity reduction and splashing at obstructing structures

For free falling rain drops where the raindrops' velocities were suddenly reduced by a structure (e.g. the roof) of a building as shown in Figure 5(a) to (d), the rain streaks appeared almost stationary and bright as they are not falling at terminal velocities (see Equation (1)). This kind of distortion consists of both

the usual motion blurred long rain streaks as well as the brighter and shorter streaks, as shown in Figure 5(a) to (d).

It is illustrated in Figure 5(i) to (l) that the CycleGAN was able to remove both fast and slow rain streaks, while the ID-CGAN was only able to remove the faster rain streaks. This may be due to the same reasons, as discussed in Type II. As shown in Figure 5(f) and (h), most of the slow rain streaks remained, in the case of the ID-CGAN. Based on these observations, the CycleGAN is more robust for a wide range of real rain defects, as compared to the ID-CGAN.

5.5 Type V: Splashing and accumulation of rain water on ground surface

Rain water tends to accumulate on surfaces such as the road surface or the roof of a building. Hence, distortions due to water splashing defects are common in rain images. Figure 6(a) to (d) show samples of such rain distortion.
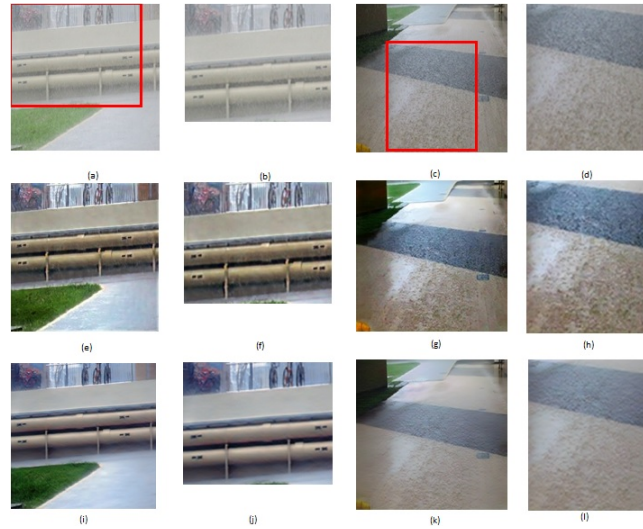

Fig. 6. Type V distortion for splashing and accumulation of rain water on ground surface as shown in Figure 6(a) to (d), and their rain removal results by the ID-CGAN and CycleGAN were shown in Figure 6(e) to (h) and Figure 6(i) to (l) respectively.

Figure 6(e) to (h) and Figure 6(i) to (l) show the results of removing rain using the ID-CGAN and CycleGAN respectively. As shown by the results, the CycleGAN was able to remove water splashes and ripples of water accumulated on the surface completely. As shown in Figure 6 (e) and (f), the ID-CGAN has introduced many white artefacts. This was expected as the ID-CGAN was not

trained to remove such a type of defect, But the CycleGAN was able to remove the defect very well with good contrast, as shown in Figure 6 (i) and (j). The ID-CGAN also created a large patch of bright defect on the accumulated surface water, as shown in Figure 6 (g) and (h). Such artefacts are not observed in the CycleGAN, as shown in Figure 6 (k) and (l). Based on the above observations, the CycleGAN has shown to be superior to the ID-CGAN in removing Type V rain distortion.

5.6   Quantitative Comparison Results

As both the ID-CGAN and the CYCLEGAN are learned using the same synthetic data set provided by the ID-CGAN, we can also perform a quantitative comparison to compare both performance based on the same synthetic data set provided by the ID-CGAN [41], as shown in Table 1. The results show that our proposed CYCLEGAN method achieve superior quantitative performance over the more recent ID-CGAN method using the mathematically defined SSIM and perceptually defined VIF measures.

Table 1. Quantitative Comparison between the ID-CGAN and the CycleGAN using the synthetic data set

	CycleGAN	ID-CGAN
SSIM	0.9992	0.8133
VIF	0.6376	0.4148

6   Conclusions

This paper is addressing the impracticality in collecting aligned rain and rain-free image pairs to train rain removal GANs for real outdoor task in computer vision as synthetic rain's statistics may not faithfully representing real rain. Hence the CycleGAN is proposed to re-construct images under visual disruption caused by rain as it can be trained using real rain images. A qualitative study based on rain physics for comparing the effectiveness of rain removal in five types of real rain distortion is presented, along with quantitative evaluation using the limited synthetic data set provided by the ID-CGAN work [41]. Our results demonstrate that the proposed CycleGAN approach is more robust and effective in removing all five types of real rain defects from images than the ID-CGAN, while preserving image scene details. In comparison, the ID-CGAN was unable to remove bright and short rain streaks, leaving behind many white artefacts and was not able to enhance images with low brightness or contrast. Hence, we could potentially extend the CycleGAN framework as a new state-of-the-art single-image rain disruption removal method from single-image to video for rain disruption removal in the future.

References

1. Birdal, T., Ercil, A.: Real-time automated road lane and car detection for autonomous driving. In: Sabanci University, Faculty of Engineering and Natural Science (2007)
2. Brewer, N., Liu, N.: Using the shape characteristics of rain to identify and remove rain from video. In: Joint IAPR International Workshops on SPR and SSPR. pp. 451–458 (2008)
3. C. Ledig, L. Theis, F.H.J.C.A.C.A.A.A.A.A.T.J.T.Z.W.e.a.: Photo-realistic single image super-resolution using a generative adversarial network. In: arXiv preprint arXiv:1609.04802 (2016)
4. Chen, Y.L., Hsu, C.T.: A generalized low-rank appearance model for spatio-temporally correlated rain streaks. In: IEEE International Conference on Computer Vision (ICCV), IEEE Conference Publications. pp. 1968–1975 (2013)
5. De-An Huang, Li-Wei Kang, M.C.Y.C.W.L., Wang, Y.C.: Context-aware single image rain removal. In: IEEE International Conference on Multimedia and Expo (ICME), IEEE Conference Publications. pp. 164–169 (2012)
6. De-An Huang, Li-Wei Kang, Y.C.W., Lin, C.W.: Self-learning based image decomposition with applications to single image denoising. IEEE Transactions on Multimedia (TMM) **16**(1), 83–93 (2014)
7. Duan-Yu Chen, C.C.C., Kang, L.W.: Visual depth guided color image rain streaks removal using sparse coding. IEEE Transactions on Circuits and Systems for Video Technology (TCSVT) **24**(8), 1430–1455 (2014)
8. Fu, Y. H., K.L.W.L.C.W., T., H.C.: Single-frame-based rain removal via image decomposition. In: Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing. Prague, Czech Republic (May 2011)
9. Fu, X., Huang, J., Ding, X., Liao, Y., Paisley, J.W.: Clearing the skies: A deep network architecture for single-image rain removal. IEEE Transactions on Image Processing **26**, 2944–2956 (2017)
10. Fu X, Huang J, Z.D.H.Y.D.X., J, P.: Removing rain from single images via a deep detail network. In: IEEE Conference on Computer Vision and Pattern Recognition (2017)
11. Garg, K., Nayar, S.K.: Detection and removal of rain from videos. In: Proc. CVPR, 1. pp. 528–535 (2004)
12. Garg, K., Nayar, S.K.: Vision and rain. Int'l Journal of Computer Vision **3**(27) (2007)
13. https://github.com/pytorch/pytorch:
14. Ian Goodfellow, Jean Pouget-Abadie, M.M.e.a.: Generative adversarial nets. In: NIPS (2014)
15. J. Bossu, N.H., Tarel, J.P.: 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 (2011)
16. Jin-Hwan Kim, Chul Lee, J.Y.S., Kim, C.S.: Single-image deraining using an adaptive nonlocal means filter. In: IEEE International Conference on Image Processing (ICIP), IEEE Conference Publications. pp. 914–917 (2013)
17. Jing Xu, Wei Zhao, P.L., Tang, X.: An improved guidance image based method to remove rain and snow in a single image. Computer and Information Science **5**, 49–55 (2012)
18. Jing Xu, Wei Zhao, P.L., Tang, X.: Removing rain and snow in a single image using guided filter. In: IEEE International Conference on Computer Science and Automation Engineering (CSAE), IEEE Conference Publications. pp. 304–307 (2012)

19. Li-Wei Kang, Chia-Wen Lin, C.T.L., Lin, Y.C.: Self-learning-based rain streak removal for image/video. In: IEEE International Symposium on Circuits and Systems (ISCAS), IEEE Conference Publications. pp. 1871–1874 (2012)
20. Li-Wei Kang, C.W.L., Fu, Y.H.: Automatic single-image-based rain streaks removal via image decomposition. IEEE Transactions on Image Processing (TIP) **21**(4), 1742–1755 (2012)
21. Mirza, M., Osindero, S.: Conditional generative adversarial nets. CoRR **abs/1411.1784** (2014)
22. P. C. Barnum, S.N., Kanade, T.: Analysis of rain and snow in frequency space. Int'l Journal of Computer Vision **86**(2-3), 256–274 (2010)
23. Pathak, D., Krähenbühl, P., Donahue, J., Darrell, T., Efros, A.: Context encoders: Feature learning by inpainting (2016)
24. Phillip Isola, Jun-Yan Zhu, T.Z., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: CVPR (2017)
25. Radford, A., Metz, L., Chintala, S.: Unsupervised representation learning with deep convolutional generative adversarial networks. CoRR **abs/1511.06434** (2015)
26. S.-H. Sun, S.P.F., Wang, Y.C.F.: Exploiting image structural similarity for single image rain removal. In: IEEE Int'l Conf. Image Processing. pp. 4482–4486 (2014)
27. S. Zhou, J. Gong, G.X.H.C., Lagnemma, K.: Road detection using support vector machine based on online learning and evaluation. In: Intelligent Vechile Research Center, Beijing (2010)
28. Sheikh, H.R., Bovik, A.C.: Image information and visual quality. IEEE Transactions on Image Processing **15**(2), 430–444 (2006)
29. Shengyan, Z., Karl, L.: Self- supervised learning method for unstructured road detection using fuzzy support vector machines. In: Proceeding of Robotic Mobility Group, Massachusetts Institute of Technology. Boston, USA (2010)
30. Son, C.H., Zhang, X.P.: Rain removal via shrinkage of sparse codes and learned rain dictionary. In: IEEE ICME (2016)
31. Soo-Chang Pei, Y.T.T., Lee, C.Y.: emoving rain and snow in a single image using saturation and visibility features. In: IEEE International Conference on Multimedia and Expo Workshops (ICMEW), IEEE Conference Publications. pp. 1–6 (2014)
32. W. Yang, R. T. Tan, J.F.J.L.Z.G., Yan, S.: Joint rain detection and removal via iterative region dependent multi-task learning. CoRR **abs/1609.07769** (2016)
33. W. Yang, Robby T. Tan, J.F.J.L.Z.G., Yan, S.: Deep joint rain detection and removal from a single image. In: Computer Vision and Pattern Recognition, CVPR (2017)
34. Wang, C., Xu, C., Wang, C., Tao, D.: Perceptual adversarial networks for image-to-image transformation. IEEE Transactions on Image Processing **27**, 4066–4079 (2018)
35. X. Zhang, H. Li, Y.Q.W.K.L., Ng, T.K.: Rain removal in video by combining temporal and chromatic properties. In: Proc. IEEE Int'l Conf. Multimedia and Expo. pp. 461–464 (2006)
36. Y. Li, R. T. Tan, X.G.J.L., Brown, M.S.: Rain streak removal using layer priors. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 2736–2744 (2016)
37. Y. Luo, Y.X., Ji, H.: Removing rain from a single image via discriminative sparse coding. In: ICCV (2015)
38. Yu Li, Robby T. Tan, X.G.J.L., Brown, M.S.: Single image rain streak separation using layer priors. IEEE Transactions on Image Processing (2017)

39. Z. Wang, A. C. Bovik, H.R.S., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing **13**(4), 600–612 (2004)
40. Zhang, H., Patel, V.M.: Convolutional sparse coding-based image decomposition. In: British Machine Vision Conference (2016)
41. Zhang, H., Sindagi, V., Patel, V.M.: Image de-raining using a conditional generative adversarial network. CoRR **abs/1701.05957** (2017)
42. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: IEEE International Conference on Computer Vision (ICCV) (2017)