Generative Based-Framework for Video Rain Removal

Aymen Hamrouni, Raby Hamadi
Affiliation: ECE M.S. Students
King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia
Email: aymen.hamrouni, raby.hamadi@kaust.edu.sa

Abstract

In rainy weather, the performances of video/image processing tasks may be severally degenerated leading to unreliable decision-making systems. In fact, rain hamper the visibility in images or videos captured in rain and hinder the use of information extracted from these scenes. As nowadays various computer vision algorithms rely on cameras (e.g., autonomous vehicle navigation systems) to gather information about the surrounding environment (e.g., tracking, detection, segmentation), many research interest is focused on restoring rain-free background scenes from degraded videos. In this context, a low-complexity approach is needed to derain the videos in real-time. In this paper, we propose a minimal-delay de-noising pipeline to eliminate noisy and unwanted rain artifacts from a live video feed. We propose a standalone deep learning Convolutional AutoEncoder (CAE) network to reconstruct noise free images and then enhance this architecture by adding a sequential Generative Adversarial Network (GAN) based module that enhances the quality of the image frames. The idea is to disentangle the de-raining problem into two sub-problems where we harness the power of the CAE and the GAN as the same time. Experimental results on the rainstreaks dataset and the raindrop dataset show that the CAE architecture with GAN enhancing outperforms a benchmarking CNN when it comes to SSIM, VIF, and FSIM. Other comparative simulations show that our proposed approach achieves close performances to an attentive standalone GAN de-raining approach with significant time complexity gain. Moreover, other conducted experiments to evaluate the performances of video recognition before and after de-noising show that the proposed approaches have potential gain to computer vision algorithms in detection and recognition.

1 Introduction

Images and videos captured from outdoor vision systems are often affected by rain. Specifically, as a complicated atmospheric process, rain can cause different types of visibility degradation (e.g., drops, streaks, etc). Due to a magnitude of environmental factors including raindrop size, rain density, and wind velocity. When a rainy image is taken, the visual effects of rain on that digital image further hinges on many camera parameters, such as exposure time, depth of field, and resolution. In fact, rain incline to obstruct or distort background scene contents and distant rain accumulation tend to generate atmospheric veiling effects like mist or fog and blur the image contents.

In this context, CCTVs or any outdoor cameras can be hard to rely under such environmental weather resulting in major difficulties of monitoring or tracking. On top of this, and as most of the navigation vehicles, whether they are autonomous (e.g., autonomous cars/Drones), or man-controlled are equipped with high quality cameras to navigate and monitor their surroundings. In terrible

Submitted to 35th Conference on Neural Information Processing Systems (NeurIPS 2021). Do not distribute.

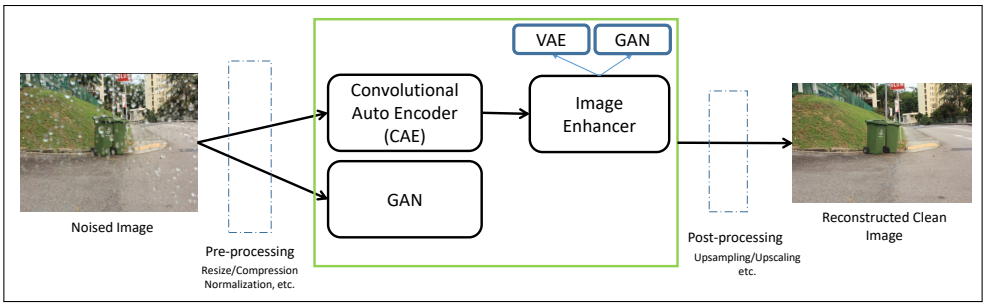

Figure 1: Architecture of the proposed image de-raining pipeline showing the different components.

weather, raindrops, for example, adhered to the camera lens can severely degrade visibility of a background scene in an image. Therefore, these systems' vision gets fuzzy and vague and leads to several navigation difficulties.

Under such poor conditions, not all of these camera-equipped devices can rely on the captured footage for computer vision tasks such as image segmentation, detection, etc. Indeed, state-of-the-art routing and navigation algorithms rely either, on the points of clouds provided by usually LIDAR systems which have very high cost and hard to maintain, or other sensing technologies such as ultra sonic which are very sensitive to variation in the temperature and has more difficulties in reading reflections from soft, curved and thin as well as a small object.

In this paper, we propose to design and implement a video processing pipeline to enhance the camera's vision[2, 6, 7]. Our objective is to put in place a cleaning system where the videos captured by a lower cost camera gets corrected and cleaned instantaneously to provide more accurate information about the surrounding in bad environmental effects[4, 21]. To this end, we take advantage of the Generative models such as Generative Adversarial Networks (GANs) and Convolutional AutoEncoders (CAE). Initially, we propose a two step de-raining process: i) at a first step, we process the videos by applying a rain removal pipeline and ii) in a second step, we enhance the video's quality using an enhancement model.

Several simulations were conducted to evaluate and compare the performances of the different approaches. The enhancer for two-step approach has been enabled via either a GAN or a Variational AutoEncoder (VAE) architecture. We have also implemented a CNN benchmarking approach [25] for fair evaluation. The models were trained and tested on two different datasets types, i.e., rain streaks and rain drops, having over $20,000$ pairs of noisy/clean images with different rain severity. Experimental results showed that the CAE architecture with GAN enhancer outperforms the classical CNN when it comes to Structural Similarity Index Measure (SSIM), Variance inflation factor (VIF), and Feature Similarity Index (FSIM) but yields lower performances on Peak Signal-To-Noise Ratio (PSNR). The standalone GAN rain removal on the other hand yielded better results than the CAE with GAN enhancer but with high running time making it hard to adapt for real-time de-raining systems. Other conducted experiments to evaluate the performances of video recognition before and after de-noising shows that the proposed approaches have potential gain to computer vision algorithms.

The remainder of this paper is organized as follows. Related work is reviewed in Section 2. The proposed de-noising architectures and their different components are presented in Section 3. Section 4 presents the experiments and simulations results. Finally, the paper is concluded in Section 5 with some future directions.

2 Related Work

The existing rain removal works can be generally categorized into two classes, i.e., image based and video based approaches. In these classes, we can also find two sub-groups: priorbased and deep learning-based approaches.

```
Layer (type)                 Output Shape              Param #
=================================================================
input_2 (InputLayer)         [(None, 256, 256, 3)]     0
_________________________________________________________________
conv2d_5 (Conv2D)            (None, 256, 256, 16)      784
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 128, 128, 16)      0
_________________________________________________________________
conv2d_6 (Conv2D)            (None, 128, 128, 32)      12832
_________________________________________________________________
max_pooling2d_3 (MaxPooling2 (None, 64, 64, 32)        0
_________________________________________________________________
conv2d_7 (Conv2D)            (None, 64, 64, 64)        18496
_________________________________________________________________
conv2d_transpose_1 (Conv2DTr (None, 64, 64, 32)        18464
_________________________________________________________________
up_sampling2d_2 (UpSampling2 (None, 128, 128, 32)      0
_________________________________________________________________
conv2d_8 (Conv2D)            (None, 128, 128, 16)      12816
_________________________________________________________________
up_sampling2d_3 (UpSampling2 (None, 256, 256, 16)      0
_________________________________________________________________
conv2d_9 (Conv2D)            (None, 256, 256, 3)       435
=================================================================
Total params: 63,827
Trainable params: 63,827
Non-trainable params: 0
_________________________________________________________________
```

Figure 2: Layers Composition of the CAE.

For images, most approaches [10, 17, 26] formulate rain removal as a signal separation problem. Single image de-raining is a classical problem in low-level computer vision which aims to remove rain streaks from a rainy image. Traditional methods, such as Gaussian mixture model [14], kernel based method [11], sparse coding [17], low rank approximation [3], representation learning [8], and dictionary learning [22] have been employed in de-raining task. Though these methods can improve the quality of rainy images to some extent, they fail to learn the distribution of rain streaks with different scales, shapes, and directions. Thus relatively low performances are obtained. For videos, some early works were proposed to utilize the chromatic properties [16], Gaussian mixture model (GMM) and low rank matrix completion [12] to remove rain. Most recently, benefiting from the great success of deep learning, deep neural networks (DNNs) have been applied for rain removal of videos [1, 15]. For example, Liu et al. [15] proposed a hybrid rain model to remove rain, in which the DNNs extract spatial features with temporal coherence of background. In addition, Chen et al. [1] proposed aligning scene content at super pixel (SP) level and averaging the aligned SPs to obtain the intermediate de-rain output. In [1], a convolutional neural network (CNN) is used to restore rain free details and then obtain the final output of clean content. However, the generalization ability of these DNN-based rain removal approaches suffers from insufficient training data, as the existing video databases for rain removal are of small-scale. Kang et al. [10] clustered rain and non-rain dictionaries based on the histogram of oriented gradients (HOGs) features of rain streaks as priors and then reconstructed clean images. Luo et al. [17] proposed a dictionary learning based approach to separate the background layer from the rain one. In [23], the authors presented a Wasserstein GAN-based framework for image de-raining. In this paper, the authors propose to learn a fusion of both high-level semantic features and rich local-spatial information using the feature aggregating structure of U-Net as the generator and performed extensive simulations using a synthetic dataset. In [**?**], the authors proposed a semi-supervised GAN-based de-raining network that can use both the synthetic and real rain images. The authors also proposed a real image dataset called Real200. Xin et al. [9], proposed an Unsupervised de-raining Generative Adversarial Network (UD-GAN) for single image de-raining which involves self-supervised constraints from the statistics of raw data through two collaboratively optimized modules. AutoEncoders are also widely used in literature for image de-raining. In fact, Du et al. [5] proposed a Conditional Variational AutoEncoder (CVAE) framework that models the latent distributions of image priors, from which the clean images are generated for image de-raining. Li et al. [13] proposed a single image de-raining network considering Rain Embedding Consistency with an ideal rain embedding encoded by the rain-to-rain AutoEncoder to address the difficulty to learn a better rain embedding. In this work, a Mask-GAM with ground-truth rain mask supervision to reweight the encoder features to leave only rain-attentive features which are passed to the decoder.

Table 1: PSNR, SSIM, VIF, and FSIM for CNN, AutoEncoder, and AutoEncoder with GAN in the case of Heavy Rain images

	CNN	CAE	CAE+GAN
PSNR	22.15	19.59	11.1
SSIM	0.674	0.75	0.96
VIF	1.53	1.17	1
FSIM	0.965	0.9	0.968

Table 2: PSNR, SSIM, VIF, and FSIM for CNN, AutoEncoder, and AutoEncoder with GAN in the case of Medium Rain images

	CNN	CAE	CAE+GAN
PSNR	23.13	21.09	11.43
SSIM	0.7	0.78	0.96
VIF	1.21	1.16	1
FSIM	0.965	0.91	0.97

3 Proposed Architectures

In this section, we present the proposed image de-raining architectures, mainly the Convolutional AutoEncoder neural network with enhancer and the Generative Adversarial Network.

As shown in Fig. 1, the idea was to have a pipeline consisting of two main components: a CAE neural network as the image reconstruction component and an image enhancer component. The incoming stream of noised images, in real-time, goes through an pre-processing phase at first where these images gets resized, and normalized. After that, the images go through a CAE consisting of total of 11 layers (9 intermediate layers). As stated in Fig. 2, the latent space have 128 dimensions. This number has been chosen by empirical simulation where we vary the dimension of the latent space and check the training/testing performances along with Principal Component Analysis estimation. The total number of trainable parameters is $63,827$. The job of the AutoEncoder is to reconstruct the original image while removing the rain and keeping the quality as similar as possible to the original image. As the AutoEncoder fails to perfectly reconstruct the input image, the next step consists of a an image enhancer model that tries to surpass this limitation and enhance the quality of the reconstructed image. Here, we propose two kind of image enhancers: i) a GAN-based enhancer and ii) an AutoEncoder-based enhancer. Finally, the resultant image is then post-processed and up-scaled to match the original image scale.

The de-raining problem is intractable, as the regions occluded by raindrops are not given. In addition, the information about the background scene of the occluded regions is completely lost for most part. To resolve the problem, we have implemented an attentive generative network using adversarial training[18]. The main idea is to inject visual attention into both the generative and discriminator networks. During the training, the visual attention learns about rain regions and their surroundings. Hence, by injecting this information, the generative network will pay more attention to the rain regions and the surrounding structures, and the discriminator network will be able to assess the local consistency of the restored regions.

As motivated before, some of the main application of rain removal is real-time video feeds captures by the cameras of several navigation systems such as UAVs, autonomous vehicles, etc. Our aim is to enable such capability. Therefore, we have implemented an interceptor process that capture the frames from either offline videos or live feeds and perform de-noising operations. Such operations include removing the rain streaks from the video's frames and then running the GAN image enhancer. The resultant frame is then instantaneously sent back to feed pipeline.

4 Experiments and Simulation Results

This section discusses the experimental settings and the simulations that test our approaches.

Table 3: PSNR, SSIM, VIF, and FSIM for CNN, AutoEncoder, and AutoEncoder with GAN, in the case of Light Rain images

	CNN	CAE	CAE+GAN
PSNR	36.11	23	11.5
SSIM	0.97	0.826	0.96
VIF	1.2	1.13	1
FSIM	0.971	0.948	0.971

Table 4: VIF, FSIM, PSNR, and SSIM metrics for the rainstreaks dataset for the different image de-raining approaches: CAE, CAE + GAN enhancer, CAE + VAE enhancer, and the standalone GAN.

	CAE	CAE + GAN Enhancer	CAE + VAE Enhancer	Standalone GAN
VIF	1.16	1	1.12	1.04
FSIM	0.91	0.97	0.81	0.88
PSNR	21.09	11.43	24.55	35.2
SSIM	0.78	0.96	0.68	0.99

4.1 Dataset and Experimental Settings

In order to train the models, we have chosen to perform evaluation on two types of rain images: i) rain streak images and ii) rain drops images. For the rain streak, we have utilized eight frequently-used benchmark datasets: Rain1400 synthesized by Fu et al. [7], Rain12 provided by Li et al. [14], Rain100L and Rain100H provided by Yang et al. [24], DID-MDN dataset provided by Zhang et al. [27] Specifically, Rain1400 includes 14000 rainy images synthesized from 1000 clean images with 14 kinds of different rain streak orientations and magnitudes. Among these images, 900 clean images (12600 rainy images) are chosen for training and 100 clean images (1400 rainy images) are selected as testing samples. Rain12 consists of 12 rainy/clean image pairs. Rain100L is selected from BSD200 [20] with only one type of rain streaks, which consists of 200 image pairs for training and 100 image pairs for testing. Compared with Rain100L, Rain100H with five types of streak directions is more challenging, which contains 1800 image pairs for training and 100 image pairs for testing. Since Rain12 has few samples, we directly adopt the trained model on Rain100L to do an evaluation on Rain12. The DID-MDN dataset contains 12000 pairs of images with different rain level (i.e., light, medium, and heavy) from which we pick 90% pairs for training and 10% pairs for testing. The resultant gigantic dataset consists of more than 16000 pairs of images with different resolutions, scenes (i.e., indoor outdoor) and rain degree. For the droplet dataset, we have used [19] which contains more than 8.000 pairs of images. All the simulations were conducted on a Linux server with Tesla P100 GPU and around 1TB of RAM.

4.2 Quantitative Experiments

For the Quantitative experiments, we propose to perform several simulation. Initially, we proposed to compute the average PSNR, the SSIM, the VIF, and FSIM for the three different rain classes: heavy, medium, and light. This has been done on the CNN benchmarking model, the CAE, and the CAE with GAN enhancer. The results of such simulation are shown in Table. 1, Table. 2, and Table. 3.

As we can see, the CNN based approach achieves better PSNR and lower SSIM. This means that the produced image is close the original image pixel wise since PSNR is lower but it somehow still have clear and visual artifacts and not human plausible. On the other hand, as the AutoEncoder with GAN enhancer approaches achieves better SSIM, The produced image is somehow different from the initial image pixel wise but it looks more plausible as it captures visual perception of human and keeps the structural similarity. Also, the AutoEncoder with GAN achieved higher performances than the standalone AutoEncoder and CNN when it comes to VIF and FSIM.

We have also suggested to implement a VAE as an enhancer with the standalone CAE architecture and computed the average PSNR, the SSIM, the VIF, and FSIM for rainstreaks dataset and raindrop

Table 5: VIF, FSIM, PSNR, and SSIM metrics for the raindrops dataset for the different image de-raining approaches: CAE, CAE + GAN enhancer, CAE + VAE enhancer, and the standalone GAN.

	CAE	CAE + GAN Enhancer	CAE + VAE Enhancer	Standalone GAN
VIF	1.55	1.8	1.12	0.9
FSIM	0.56	0.81	0.75	0.88
PSNR	11.3	8.05	11.5	31.57
SSIM	0.88	0.826	0.81	0.9023

Table 6: Monte-Carlo Simulation (average of 1000 Run) for the full pipeline on the two datasets

	CNN	CAE	CAE + VAE Enhancer	CAE + GAN Enhancer	Standalone GAN
Complexity (s)	0.015	0.02	0.03	0.07	0.18

dataset. As Table. 4 and Table. 5 shows, this simulation been done on the CAE, the CAE + GAN enhancer, the CAE + VAE Enhancer, and a standalone GAN. This simulations shows that the CAE + GAN Enhancer outperforms the CAE + VAE Enhancer in all the four metrics expect the PSNR where the latter has higher values than the CAE + GAN Enhancer. This is due to the fact the GAN is purely generative and the missing rained pixels are filled with different alike pixels unlike the VAE where it tries to reconstruct the values of the missing pixels from the same distribution. The Standalone GAN achieves overall higher performances than all the other approaches especially in the SSIM in both datasets. However, this out performance is clearly shown in the raindrop dataset because the model has been originally trained on such type of images.

In another experiment, we compute the running time of the whole de-noising pipeline for the four approaches as it is mentioned in Table. 6. The Monte Carlo simulation shows that the GAN has the highest running time while the CNN has the lowest running time. The AutoEncoder+ GAN Enhancer has a slightly higher running time from the CAE which makes it more adaptable for real-time systems.

At a final quantitative experiment, we perform another Monte Carlo simulation with 500 runs to evaluate the performances of the de-noising approaches when it comes to object detection. The result of YOLO5 for object detection on the images is found in Table. 7. As we can see, the de-raining process has improved the Intersection over Union (IoU), the Precision, the Recall, and the F1-Score. Overall, the standalone GAN is achieving the highest detection performances followed by the CAE + GAN Enhancer.

4.3 Qualitative Experiments

For the Qualitative experiments, we propose to run a simulation on three different images with different rain degrees. The result of this simulation is shown in Fig. 3. As we can see, the CNN failed to remove rain streaks from images experiencing heavy rain unlike the AutoEncoder with GAN where the images is clean with minor artifacts.

In the following simulation, we highlight the effect of the de-noising approaches on the object detection and recognition. The result of YOLO5 for object detection on the images in the rain drop dataset shows that some objects can be more accurately detected and recognized when the image is de-noised. As shown in Fig. 5, the result of YOLO5 for object detection on the images in the rain streak and raindrops dataset shows that important information can be gained from the proposed CAE + GAN enhancer especially in terms of detection and recognition accuracy.

5 Conclusion and Future Work

In this paper, we devised an image denoising pipeline where a real-time video feed corrupted with rain accumulation, either rainstreaks or raindrops, goes through a process to be cleaned and rain-free. The proposed process consists of two steps: a Convolutional AutoEncoder (CAE) network

Table 7: Average values of object detection metrics such as IoU, Precision, Recall, and F1-Score for the Original rainy image, the CAE, the CAE + GAN Enhancer, the CAE + VAE Enhancer, and for the standalone GAN

	Noisy Image	CAE	CAE + GAN Enhancer	CAE + VAE Enhancer	Standalone GAN
IoU	0.65	0.75	0.89	0.77	0.95
Precision (%)	33	81	87	81	91
Recall (%)	23	73	74	73	85
F1-Score	0.27	0.76	0.79	0.76	0.87

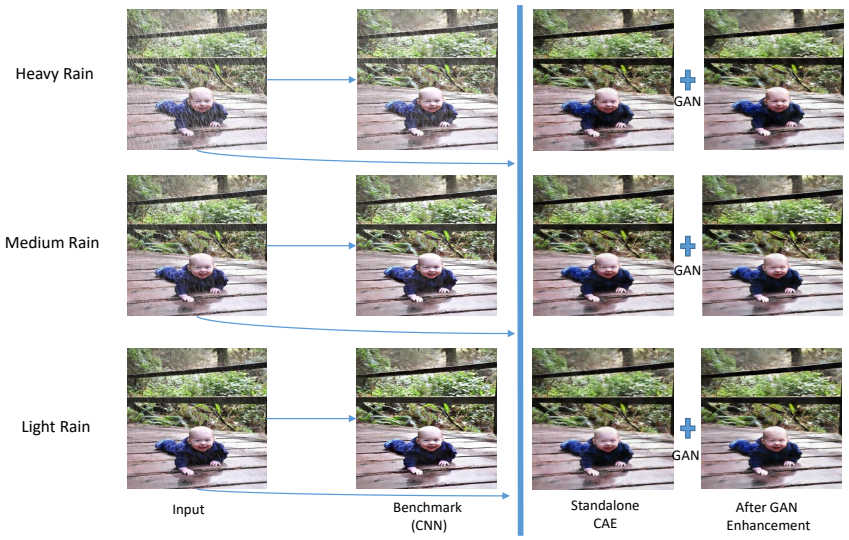

Figure 3: Sample output for the three different approaches: CNN, CAE, and the CAE + GAN Enhancer. The sample corresponds to three different type of rain classes: heavy rain, medium rain, and light rain.

for the rain removal followed by an Generative Adversarial Network (GAN) approach for image enhancement. Simulation results indicates that this approach achieves better performances in terms of FSIM, SSIM, and VIF than a benchmarking CNN and a VAE enhancer. Also, simulations indicated that this approach achieving close result to an attention-based Standalone GAN architecture with significant time complexity gain. Other experimental results also suggest that the proposed CAE + GAN enhancer ameliorate computer vision techniques in terms of detection and recognition accuracy.

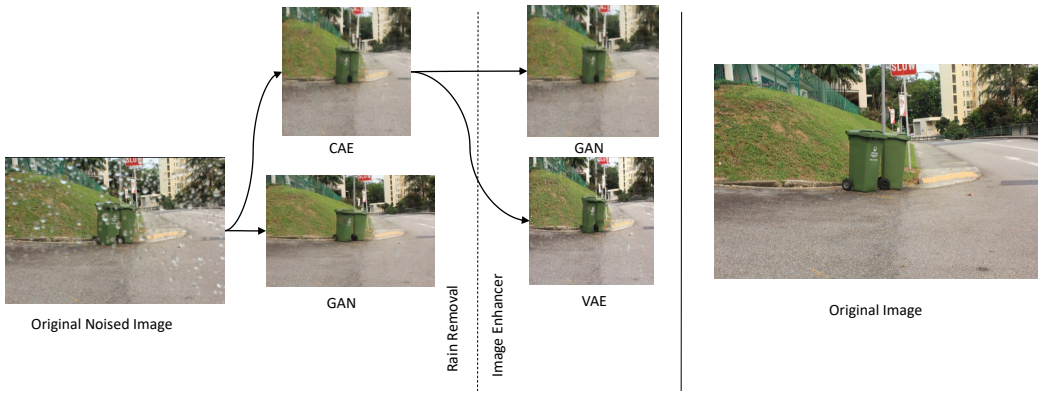

Figure 4: Visualization of de-noising procedure of a given image using CAE + GAN Enhancer, CAE + VAE Enhancer, and standalone GAN.

Figure 5: The result of YOLO5 for object detection on the images in the rain streak and raindrops dataset for the CAE + GAN enhancer approach.

References

[1] Jie Chen, Cheen-Hau Tan, Junhui Hou, Lap-Pui Chau, and He Li. Robust video content alignment and compensation for rain removal in a cnn framework, 2018.

[2] Tianyi Chen and Chengzhou Fu. Single-image-based rain detection and removal via CNN. *Journal of Physics: Conference Series*, 1004:012007, apr 2018.

[3] Yi-Lei Chen and Chiou-Ting Hsu. A generalized low-rank appearance model for spatio-temporally correlated rain streaks. In *2013 IEEE International Conference on Computer Vision*, pages 1968–1975, 2013.

[4] Sen Deng, Mingqiang Wei, Jun Wang, Luming Liang, Haoran Xie, and Meng Wang. Drd-net: Detail-recovery image deraining via context aggregation networks, 2019.

[5] Yingjun Du, Jun Xu, Qiang Qiu, Xiantong Zhen, and Lei Zhang. Variational image deraining. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, March 2020.

[6] Yingjun Du, Jun Xu, Xiantong Zhen, Ming-Ming Cheng, and Ling Shao. Conditional variational image deraining. *IEEE Transactions on Image Processing*, 29:6288–6301, 2020.

[7] Xueyang Fu, Jiabin Huang, Delu Zeng, Yue Huang, Xinghao Ding, and John Paisley. Removing rain from single images via a deep detail network. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1715–1723, 2017.

[8] Shuhang Gu, Deyu Meng, Wangmeng Zuo, and Lei Zhang. Joint convolutional analysis and synthesis sparse representation for single image layer separation. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 1717–1725, 2017.

[9] Xin Jin, Zhibo Chen, Jianxin Lin, Zhikai Chen, and Wei Zhou. Unsupervised single image deraining with self-supervised constraints. In *2019 IEEE International Conference on Image Processing (ICIP)*, pages 2761–2765, 2019.

[10] Li-Wei Kang, Chia-Wen Lin, and Yu-Hsiang Fu. Automatic single-image-based rain streaks removal via image decomposition. *IEEE Transactions on Image Processing*, 21(4):1742–1755, 2012.

[11] Jin-Hwan Kim, Chul Lee, Jae-Young Sim, and Chang-Su Kim. Single-image deraining using an adaptive nonlocal means filter. In *2013 IEEE International Conference on Image Processing*, pages 914–917, 2013.

[12] Jin-Hwan Kim, Jae-Young Sim, and Chang-Su Kim. Video deraining and desnowing using temporal correlation and low-rank matrix completion. *IEEE Transactions on Image Processing*, 24(9):2658–2670, 2015.

[13] Yizhou Li, Yusuke Monno, and Masatoshi Okutomi. Single image deraining network with rain embedding consistency and layered lstm. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 4060–4069, January 2022.

[14] Yu Li, Robby T. Tan, Xiaojie Guo, Jiangbo Lu, and Michael S. Brown. Rain streak removal using layer priors. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2736–2744, 2016.

[15] Jiaying Liu, Wenhan Yang, Shuai Yang, and Zongming Guo. Erase or fill? deep joint recurrent rain removal and reconstruction in videos. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3233–3242, 2018.

[16] Peng Liu, Jing Xu, Jiafeng Liu, and Xianglong Tang. Pixel based temporal analysis using chromatic property for removing rain from videos. *Comput. Inf. Sci.*, 2:53–60, 2009.

[17] Yu Luo, Yong Xu, and Hui Ji. Removing rain from a single image via discriminative sparse coding. In *2015 IEEE International Conference on Computer Vision (ICCV)*, pages 3397–3405, 2015.

[18] Rui Qian, Robby T. Tan, Wenhan Yang, Jiajun Su, and Jiaying Liu. Attentive generative adversarial network for raindrop removal from a single image, 2017.

[19] Rui Qian, Robby T. Tan, Wenhan Yang, Jiajun Su, and Jiaying Liu. Attentive generative adversarial network for raindrop removal from a single image. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[20] Yukai Shi, Keze Wang, Chongyu Chen, Li Xu, and Liang Lin. Structure-preserving image super-resolution via contextualized multitask learning. *IEEE Transactions on Multimedia*, 19(12):2804–2815, dec 2017.

[21] Ye-Tao Wang, Xi-Le Zhao, Tai-Xiang Jiang, Liang-Jian Deng, Yi Chang, and Ting-Zhu Huang. Rain streak removal for single image via kernel guided CNN. *CoRR*, abs/1808.08545, 2018.

[22] Yinglong Wang, Shuaicheng Liu, Chen Chen, and Bing Zeng. A hierarchical approach for rain or snow removing in a single color image. *IEEE Transactions on Image Processing*, 26(8):3936–3950, 2017.

[23] Sahil Yadav, Aryan Mehra, Honnesh Rohmetra, Rahul Ratnakumar, and Pratik Narang. Deraingan: Single image deraining using wasserstein gan. *Multimedia Tools and Applications*, 80(30):36491–36507, 2021.

[24] Wenhan Yang, Robby T. Tan, Jiashi Feng, Zongming Guo, Shuicheng Yan, and Jiaying Liu. Joint rain detection and removal from a single image with contextualized deep networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(6):1377–1393, 2020.

[25] Youzhao Yang and Hong Lu. Single image deraining via recurrent hierarchy enhancement network. In *Proceedings of the 27th ACM International Conference on Multimedia*, MM '19, page 1814–1822, New York, NY, USA, 2019. Association for Computing Machinery.

[26] He Zhang and Vishal M. Patel. Density-aware single image de-raining using a multi-stream dense network, 2018.

[27] He Zhang and Vishal M Patel. Density-aware single image de-raining using a multi-stream dense network. In *CVPR*, 2018.