APCA-Net: Adaptive object detection in rainy weather based on principal component analysis

Xue Wang
School of Information Engineering, Yangzhou University, Yangzhou, China
MZ120210911@stu.yzu.edu.cn

Lu Ding
School of Information Engineering, Yangzhou University, Yangzhou, China
MZ120230950@stu.yzu.edu.cn

Yuan Huang
School of Information Engineering, Yangzhou University, Yangzhou, China
Yuan_Huang@yzu.edu.cn

ABSTRACT

Since rain has a negative impact on the image processing, traditional image-based target detection algorithms perform poorly on rainy days. The traditional sparse coding based rain-removal method requires the selection and adjustment of many parameters, which requires professional manual adjustment and optimization. Additionally, these methods demand significant computational load and resource allocation when processing high-resolution images. To overcome these challenges, we propose an adaptive object detection method on rainy days. This approach incorporates a rain removal model based on Principal Component Analysis (PCA) as a preprocessing step and combines it with a target detector to improve the image recognition effectiveness. It adaptively predicts the hyperparameters for the Preprocessing Refinement Module (PRM) by inputting low-resolution images into Convolutional Neural Networks (CNNs). This step not only eliminates the need for manual intervention but also reduces the computational complexity of the CNN training step. Then, the PRM uses these predicted hyperparameters for rain-removal and image enhancement. Thus, APCA-Net can improve the precision of object detection. Our experimental results on VOC-based synthetic datasets confirm the effectiveness of this method.

CCS CONCEPTS

• : **Computing methodologies**; • **Artificial intelligence**; • **Computer vision**; • **Computer vision problems**; • **Object detection**;

KEYWORDS

Object Detection, Image Adaptive Enhancement, Principal Component Analysis

ACM Reference Format:
Xue Wang, Lu Ding, and Yuan Huang. 2024. APCA-Net: Adaptive object detection in rainy weather based on principal component analysis. In *International Conference on Computing, Machine Learning and Data Science (CMLDS 2024), April 12–14, 2024, Singapore, Singapore*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3661725.3661770

CMLDS 2024, April 12–14, 2024, Singapore, Singapore

ACM ISBN 979-8-4007-1639-3/24/04
https://doi.org/10.1145/3661725.3661770

1 INTRODUCTION

In recent years, the application of object detection has permeated diverse fields, seeking to autonomously discern and precisely pinpoint objects of interest within images or videos. The relentless evolution of computer vision technology underscores the pivotal role of object detection across practical domains, including but not limited to autonomous driving, video surveillance, face recognition, and medical image analysis.

Analyzing abundant empirical data from automatic driving scenarios highlights the proficiency of traditional target detection algorithms in clear weather conditions. However, real-world complexities, particularly intricate weather environments, introduce disruptions in sensor data, resulting in incongruent outputs related to target presence, absence, and associated information. Multi-modal data content experiences degradation and asymmetric distortion.

Currently, adverse weather conditions, especially in the presence of rain, are a prominent research focus. Unique challenges arise from spatial distribution disparities in dynamic scenes of rain and snow within a single image. While fog in static scenes can be addressed through time-domain information processing, the random spatial distribution of rain and snow complicates effective handling. Rainy days, among adverse scenarios, garner significant research attention.

Addressing object detection in rainy scenarios involves two core steps: first, enhancing visibility through a rain removal model as a pre-processing step, and second, actual object detection. Image restoration technology is integral to rainy day target detection, emphasizing that clearer target feature information collected during rainy conditions enhances detection efficacy across diverse scenarios.

Commonly used image rain removal methods include filter-based rain removal [1.]; The rain-removing model is constructed with sparse coding of rain-removing models [2., 3.] based on deep learning [4.] by using various priors and assumptions on the characteristics of the rain-stripe, such as direction, size and density. However, due to the difference between the recovered image domain and the trained image domain, the combination of the rain removal model and the target detection model can not always guarantee the high precision of the target detection. In addition, sparsely coded rain removal methods may reduce the detection speed due to additional rain removal tasks.

To solve the above problems, we propose an adaptive object detection method based on principal component analysis (PCA) for rainy days. In general, the CNN adaptively predicts the hyperparameters of the Preprocess refine Module (PRM) based on the brightness, color, hue, and weather-specific information of the input image, and the PRM then enhances the image data based on the hyperparameters predicted by the CNN. After PRM processing, the interference of specific weather information in the image can be suppressed, and the original data of the image can be restored. We propose an end-to-end joint optimization scheme for learning PRM, CNN and YOLOv3 detection networks. In addition, we use images under normal and harsh weather conditions for training. By using the CNN network, our proposed method can adaptively process images affected by different degrees of weather. Our main contributions can be summarized as follows:

- We propose an adaptive detection framework for rainy day images based on principal component analysis, seamlessly integrating an image enhancement module and detection network. This integration aims to enhance image quality by effectively eliminating rain stripe noise interference and refining target edges. The framework demonstrates robust target detection in both rainy and normal weather conditions.
- Our adaptive image rain removal algorithm introduces a novel approach, where threshold parameters are trained through a CNN prediction network. Leveraging CNN enables automatic learning and prediction of optimal hyperparameter combinations, thereby alleviating the need for manual search. By capturing the intricate relationship between input data and hyperparameters, the CNN identifies the most effective hyperparameter configuration, enhancing overall model performance
- The experimental results obtained on the synthetic rain data set based on VOC prove that the method has better effect.

The rest of the paper is organized as follows. We briefly review related works in Section II. Then, the proposed effective framework for target detection on rainy days is introduced in Section III. The result and comparisons are given and discussed in Section IV. At the end of the paper, we conclude in Section V.

2 RELATEWORK

2.1 *Single Image Derain*

Raindrops will block the target object, resulting in changes in its appearance features or even complete occlusion, making it difficult to identify and detect in the image. The presence of rain will cause the decrease in image contrast, blur the difference between the target and the background, and make the edge information of the target object unclear, which will interfere with the detection of the target. And may produce false detection.

Since there is no time information in a single image, the removal of rain in a single image is more challenging. Kim et al. [5.] proposed that nuclear regression and non-local mean filtering were used to detect and remove the rain streak, but the object outline and details in the image would become blurred. Eigen et al. [6.] first applied convolutional neural networks to remove rain from a single image, constructing a 3-layer convolutional neural network to learn the nonlinear mapping between images with rain and images without rain. Fu et al. [7.] proposed a deep detail network based on residual ResNet, which can enhance deep learning by changing mapping range, divide input into high-frequency detail layer and base layer, and enhance image features by directly mapping input to output. However, the result of this method is easy to lose details. Jiang et al. [8.]used ConvLSTM to process rain streaks under different scales and spatial locations, thus capturing global textures and designing a progressive feature fusion enhancement module. In the form of pyramid representation, multi-scale image information was used to present rain streaks, and rain removal was realized in a single image through a multi-scale progressive fusion network.

2.2 *Image adaptive enhancement*

Adaptive image enhancement technology, capable of automatically adjusting parameters based on varying input image conditions, is crucial for accommodating different environments and application scenarios. This adaptability enhances the universality and flexibility of the enhancement effect, minimizing manual intervention, elevating image quality, and ensuring compatibility with complex scenes. A notable contribution in this realm is presented by Wang et al. [9.], who introduced an adaptive method grounded in local gamma transform and a light-reflection model. This approach features a brightness adjustment function that adaptively tunes enhancement parameters according to the light distribution characteristics of the input image.

In a parallel vein, Polesel et al. [10.] proposed an innovative unsharpen mask method for image contrast enhancement. Employing an adaptive filter to regulate the contribution of the sharpening path allows for contrast enhancement in high-detail areas, while smooth areas experience minimal or no image sharpening. Furthermore, Upadhyay et al. [11.] presented a GAN-based framework that incorporates an adaptive loss function to enhance robustness against out-of-distribution noise data. This framework automatically adjusts norm-size penalty residuals of spatial changes and estimates voxel uncertainty in predictions, contributing to the adaptability and resilience of the model.

3 PROPOSED METHOD

In this section, we will provide a detailed introduction to a proposed effective framework for rainy weather target detection. This framework consists of three modules: the Predict Refine Module (PRM) data enhancement module, the YOLOV3 target detector, and the CNN parameter prediction module, as shown in Fig.1. The framework significantly improves target detection performance in rainy conditions.

3.1 *PRM*

For gradient-based optimization problems, filters need to ensure the differentiability of their parameters to allow the training of CNNs through backpropagation. To simplify the modeling of filters as basic neural network layers, piecewise linear functions are used instead of smooth curves. Most editing operations can be determined at low resolution, and to save time and space costs, operations can be performed on downsampled versions of the images. Specifically, the filter parameters are trained on a low-resolution (256×256) version of the original image. Subsequently, the same trained filters and parameters are applied to the original high-resolution image. Therefore, the filters need to be resolution-independent. Following these principles, our proposed PRM module consists of five differentiable filters with adjustable hyperparameters, including

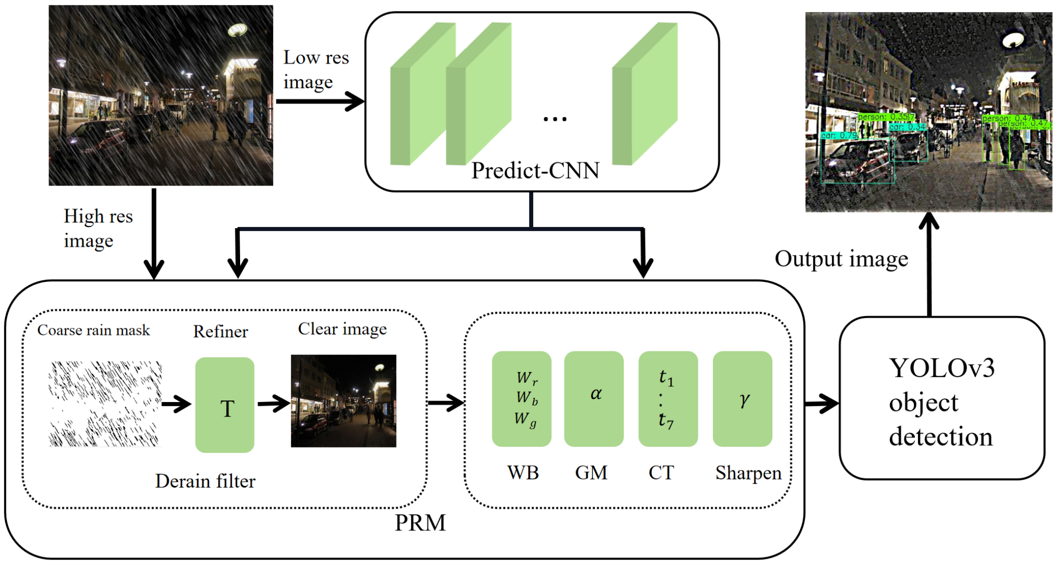

Figure 1: Overview of whole framework. The input high-resolution image is fed into the data enhancement module for image preprocessing; meanwhile, the reduced low-resolution image is input into the Predict-CNN module for training through detection loss learning to predict the parameters of Derain filter, White Balance, Gamma filter, Color and Tone filter and Sharpen filter. Finally, the processed image is input into YOLO for target detection.

the Derain filter, White Balance filter (WB), Gamma filter, Color and Tone filter(CT), and Sharpen filter.

1)Derain Filter Based on PCA

In general, rain images can be decomposed into two parts:

$$I = B + R \tag{1}$$

where I is the rain image, R is the high frequency detail layer. The high frequency detail layer contains the rain line characteristics and high frequency detail information. B is the background layer without rain streaks

Wang et al. [12.] discovered that raindrops typically exhibit strong light reflection, resulting in significantly higher pixel intensities compared to the surrounding background pixels. Exploiting this characteristic, the initial positions of rain can be detected. For a given color rain image I, where each pixel value *I(x, y)* is higher than the average value $\bar{I}^{(k)}$ of its surrounding five windows *w(k)* $(k = 1, 2, 3, 4, 5)$.$I(x, y)$ is located at different positions within the five windows, namely center, top-left corner, bottom-left corner, top-right corner, and bottom-right corner, as shown in Fig.2. If the following inequality is satisfied, it is classified as rain:

$$I\left(x,y\right) > \overline{I^{(k)}} + \varphi \tag{2}$$

$$\overline{I_t^{(k)}} = \frac{\sum_{\{x,y\}\in w^{(k)}} I\left(x,y\right)}{\left|w\left(k\right)\right|}, k = \{1,2,3,4,5\}, t \in \{r,g,b\} \tag{3}$$

$$\overline{I^{(k)}} = \frac{\overline{I_t^{(k)}}}{3} \tag{4}$$

where $|w(k)|$ stands for the window size, and t means r, g, b three color channels. The parameter φ is an empirical value which will be given later in the experimental part. Because there may be some non-rain pixels with intensities that are less than neighboring rain pixels but larger than $\overline{I^{(k)}}$, The set value of φ can avoid mis-classifying this type of non-rain pixels as rain. From (2), we can obtain the coarse binary mask image Mc, In Mc, if $I(x, y)$ is recognized as a rain pixel ,the corresponding term Mc(x, y) is set to be 1; otherwise, Mc(x, y) is assigned as 0. This approach can roughly extract the high-frequency information of the rain image, ignoring some irrelevant details.

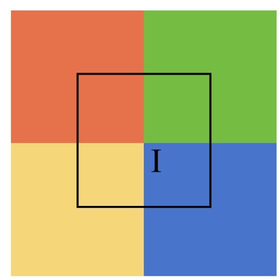

Figure 2: I is located at different positions within the five windows, namely center, top-left corner, bottom-left corner, top-right corner, and bottom-right corner.

However, the Mc contains the high-frequency detailed information and the rain line information. Further separation of the high-frequency rain lines from the high-frequency detail is required for rain removal. How to identify and eliminate those non-rain elements from Mc, other physical features of rain proposed by Chen et al. [13.] can also be used:

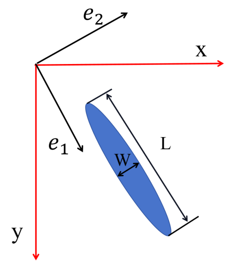

Figure 3: Analysis of rain streak's feature

- Rain streaks or raindrops are usually spindle-shaped, so the width of rain streaks will be very small.
- The color of rain streaks or raindrops is generally light white, and the length of rain streaks and raindrops will be far greater than their width.
- A rain image is stationary, so the inclination direction of all rain streaks in the scene at this moment is close, with most being nearly consistent.

Once the rain streak features are determined, the shape of the rain streaks is described. Therefore, we employ Principal Component Analysis (PCA) to extract refined binary images Mr (contains only the rain line information)from the connected regions in the coarse binary image Mc obtained above. Connected regions generally refer to the image areas composed of foreground pixels with the same pixel value and adjacent positions.

After determining the raindrop features(considering the above mentioned three conditions), we proceed to describe the shape of the rain lines. We utilize the PCA (Principal Component Analysis) method to extract a refined binary image Mr from the connected regions obtained from the coarse binary image Mc. Connected regions typically refer to image areas composed of foreground pixels with the same pixel value and adjacent positions. Thresholding based on the features is employed to separate the high-frequency rain lines from the high-frequency details in the binary image Mc. The average position vector and covariance matrix for the i^{th} connected component are represented as follows:

$$p_x = \frac{\sum_{k=1}^{N} x_k}{N} \tag{5}$$

$$cov_x = \frac{\sum_{k=1}^{N} x_n x_n^T - p_x p_x^T}{N} \tag{6}$$

where$x_n = [i_n, j_n]$, and $i_n, j_n$are respectively the corresponding coordinates of the k^{th} pixel. By performing eigen decomposition on the covariance matrix, we obtain the eigenvalues λ_1, λ_2, as well as the eigenvectors e_1, e_2, as shown in the Fig.3. The length, width, and angle of the connected component can be expressed as follows:

$$L = r\lambda_1 \tag{7}$$

$$W = r\lambda_2 \tag{8}$$

$$\beta = \frac{180^\circ \arctan \frac{e_1}{e_2}}{\pi} \tag{9}$$

where r is a proportional parameter.

In the process of extracting Mr from Mc, threshold segmentation based on above calculated features is used. After extraction, only rain streaks information remained in the binary image Mr. Based on the calculated angles β of each connected component in each image, we obtain the most frequent value β_i in the entire image. Three feature thresholds T_1, T_2, T_3 are needed. If the following conditions are met, we get determined to be a rain streak:

$$\frac{L}{W} > T_1 \tag{10}$$

$$W > T_2 \tag{11}$$

$$|\beta - \beta_i| < T_3 \tag{12}$$

where T_1, T_2 is manually set, and T_3 is derived from CNN network adaptive prediction. According to the above method, we get the refined binary mask Mr which contains rain streaks.

We map each pixel of Mr to the original image to obtain coarse background B(in (1)) as follows:

$$B = \frac{f\left(I \odot (1 - Mr), 21\right)}{f\left(1 - Mr, 21\right)} \odot Mr + I \odot (1 - Mr) \tag{13}$$

where f indicates mean filtering with window size of 21.

2) Enhancement

Hu et al.[14.] proposed five filters We select four filters from them, including: Gamma filter, White Balance filter, Sharpen filter, and Color and Tone filter. The thirteen adjustable hyperparameters in the filter are predicted through CNN.

As rainy images are typically captured in dim lighting conditions, Gamma filter which adjust exposure and White Balance filter are necessary. The deraining algorithm uses a smoothing filter for parameter mapping, which can lead to issues such as image edge blurring. Therefore, Sharpen filter are designed to enhance the edges and details of the image, which makes the image look sharper and more visually striking. Additionally, some distortions and color anomalies can be addressed through Color and Tone filter, as well as color curve adjustments.

3.2 *Predict-CNN*

By learning the complex relationship between image features and target parameters, CNN can effectively realize the task of parameter prediction learning. So we design a CNN prediction module to predict the hyperparameters required by PRM.

In the rain map, Predict-CNN can extract features of the entire image, including the brightness of the image, the color and the density of the rain pattern. It consists of five convolution layers and two fully connected layers. Each convolution layers consists of a 3×3 convolution blocks with step size 2 and a Leaky Relu. The output channels of these five convolution layers are 16, 32, 32, 32, and 32. Finally, the hyperparameters of the PRM module are output from the fully connected layer.

3.3 *YOLOv3 Tar*

The YOLOv3 algorithm, renowned for its outstanding performance, stands out as a prominent single-stage target detection methodology within the industry. Leveraging YOLOv3 [15.] as our chosen detection network, we employ an identical network architecture and loss function as the original YOLOv3, utilizing a single neural network to process the entire image. This network comprises

Table 1: Overview of train and test sets.

Sets	Number of Images	Car	Bus	Bicycle	motorbike	person
train	7779	1061	3262	801	1022	13169
test	2593	768	222	199	139	2964

(a) Synthetic rain image (b) YOLOv3 (c)PRM (d)APCA-Net

Figure 4: (a) Synthetic rain image; (b) detection results of YOLOv3 directly detection on rain image; (c) detection results of network combined PRM and YOLOv3; (d) the adaptive detection results of APCA-Net including the Predict-CNN module, PRM and YOLOV3. It can be found that the accuracy has been significantly improved, although there are still some noise in derain finished pictures, it does not affect the entire detection result.

a Darknet53 backbone network, a multi-scale feature extraction network, and a detection network. Incorporating conventional convolutions of dimensions 1*1 and 3*3, alongside upsampling of small-size feature maps, the algorithm demonstrates rapid and accurate real-time detection of multiple targets across various categories and scales. Its applicability extends to domains such as automatic driving and video surveillance, making it a widely adopted choice in industry applications.

4 EXPERIMENTS

4.1 *Dataset*

Throughout the experiment, variables like lighting conditions and background elements exhibit susceptibility to change. Obtaining a consistent set of rainy and sunny images with identical backgrounds proves challenging. Consequently, prevailing research resorts to extensive synthetic datasets. While existing synthetic datasets encompass diverse rain patterns, they lack corresponding rainfall label information for each synthetic rainy image. Moreover, the limited quantity of such images impedes CNN model training efficacy.

To enhance the detection of five object categories—pedestrians, motorbikes, buses, cars, and bicycles—we have curated a novel dataset. Leveraging the VOC [16.] dataset as raw data, which comprises 7779 training images and 2593 test images, each exhibiting varying target numbers and scales, we applied a two-fold synthesis approach. Specifically, half of the dataset underwent rain mark noise addition using Photoshop, while the remainder employed OpenCv to generate randomized white noise at various densities, simulating diverse margin sizes. This meticulously synthesized dataset aims to address the shortcomings of existing datasets and serve as a robust foundation for CNN model training. The statistics of the datasets are summarized into Table 1.

4.2 *Implementation Details*

APCA-net model is trained end-to-end using the Adam optimizer. Learn domain invariant features of source domain and target domain under Tensorflow framework. The processor is CPU i5-13400F 2.50GHZ and GPU NVIDIA GeForce RTX 3060, and the Windows 11 64GB system is used. The epoch of model is set to 60. The start learning rate is 10^{-4}, and the batch size is 6.

4.3 *Evaluation Results*

We evaluated the influence of different components in our architecture. The mean Average Precision (mAP) represents mean average precision, which is a commonly used metric for evaluating object

Table 2: Performance comparison of four method by mAP are list on the second . The five right columns show the detection precision of each class.

Method	Total(%)	Car(%)	Bicycle(%)	Motorbike(%)	Bus(%)	person(%)
YOLOV3	41.82	54	50	41	39	25
PRM	62.62	69	68	52	57	63
Global sparse model [15.]	60.44	67	67	55	50	64
APRM	65.44	70	70	54	64	68

detection algorithms. It comprehensively evaluates performance by taking into account precision across different IoU (Intersection over Union) thresholds, providing a more comprehensive performance evaluation metric. As the numerical value of mPA for a method increases, its performance capability correspondingly improves. YOLOV3 object detector has a good effect on image detection under clear weather conditions. However, the mean Average Precision(mAP) was only 41.82% under rainy conditions. On this basis, we designed PRM to recover the image, and the detection accuracy of the whole network was significantly improved by 20.8%. In addition, after adding the domain adaptive network Predict-CNN to predict the PRM hyperparameters, the accuracy of the APCA-Net is improved by 2.82% compared to PRM. The results of our network visualization are shown in 4. The overall mAP results can be seen in Table 2

5 CONCLUSION

This paper introduces a new method for image de-raining and image enhancement to improve target detection accuracy on rainy days. The low-resolution images are used in the predict-CNN module. The training is conducted through backpropagation to adaptatively predict the parameters of the PRM. The high-resolution input images are used by the PRM for rain removal based on PCA and data enhancement. Experimental results demonstrate that our proposed method performs well on the VOC synthetic dataset. Quantitative and qualitative comparisons, particularly in terms of mean Average Precision (mAP), affirm the efficiency and superiority of APCA-Net over competing methods. Despite achieving notable success, residual noise persists in our derained images, prompting our future work to focus on refining the accuracy of rain removal.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China under Grant 62301475.

REFERENCES

[1.] He, K., J. Sun, and X. Tang, Guided image filtering. IEEE transactions on pattern analysis and machine intelligence, 2012. 35(6): p. 1397-1409.
[2.] Li, Y., *et al.* Rain streak removal using layer priors. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
[3.] Luo, Y., Y. Xu, and H. Ji. Removing rain from a single image via discriminative sparse coding. in Proceedings of the IEEE international conference on computer vision. 2015.
[4.] Qian, R., *et al.* Attentive generative adversarial network for raindrop removal from a single image. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.
[5.] Kim, J.-H., *et al.* Single-image deraining using an adaptive nonlocal means filter. in 2013 IEEE international conference on image processing. 2013. IEEE.
[6.] Eigen, D., D. Krishnan, and R. Fergus. Restoring an image taken through a window covered with dirt or rain. in Proceedings of the IEEE international conference on computer vision. 2013.
[7.] Fu, X., *et al.* Removing rain from single images via a deep detail network. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
[8.] Jiang, K., *et al.* Multi-scale progressive fusion network for single image deraining. in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020.
[9.] Sindagi, V.A., *et al.* Prior-based domain adaptive object detection for hazy and rainy conditions. in European Conference on Computer Vision. 2020. Springer.
[10.] Polesel, A., G. Ramponi, and V.J. Mathews, Image enhancement via adaptive unsharp masking. IEEE transactions on image processing, 2000. 9(3): p. 505-510.
[11.] Upadhyay, U., V.P. Sudarshan, and S.P. Awate. Uncertainty-aware GAN with adaptive loss for robust MRI image enhancement. in Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021.
[12.] Wang, Y., *et al.* A framework of single-image deraining method based on analysis of rain characteristics. in 2016 IEEE International conference on image processing (ICIP). 2016. IEEE.
[13.] Chen, D.-Y., C.-C. Chen, and L.-W. Kang, Visual depth guided color image rain streaks removal using sparse coding. IEEE transactions on circuits and systems for video technology, 2014. 24(8): p. 1430-1455.
[14.] Hu, Y., *et al.*, Exposure: A white-box photo post-processing framework. ACM Transactions on Graphics (TOG), 2018. 37(2): p. 1-17.
[15.] Redmon, J. and A. Farhadi, Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767, 2018.
[16.] Everingham, M., *et al.*, The pascal visual object classes (voc) challenge. International journal of computer vision, 2010. 88: p. 303-338.