Sub-data Reloader for Weather Removal

Jung Hwan Kim[1], Yun-su Chung[2], and Dong Seog Han[1,*]
[1]School of Electronic and Electrical Engineering, Kyungpook National University,
Republic of Korea
[2]Electronics and Telecommunications Research Institute,
Republic of Korea
jkim267@knu.ac.kr, yoonsu@etri.re.kr, dshan@knu.ac.kr*

Abstract

The weather effects degrade the image screens and hinder the view of object detection. Ultimately, it could cause a safety issue for autonomous vehicles and result in malfunctioning object detection. In fact, developing weather removal and image reconstruction is an essential component for object detection. Most state-of-the-art generative architectures are still challenging in terms of reliable performance for image cleaning under different weather conditions. Thus, in this paper, we propose to use the SFHformer architecture with KITTI with weather effects. Then, we also developed the sub-data reloader for speeding up training. After applying the sub-data reloader, we managed to obtain relatively close performance of SSIM and PSNR as 97.43% and 36.11, respectively, as the training time per epoch is 1039 seconds, while the training with the whole dataset per epoch has the PSNR as 38.81 and the SSIM as 98.23% which took 2782 seconds.

I. Introduction

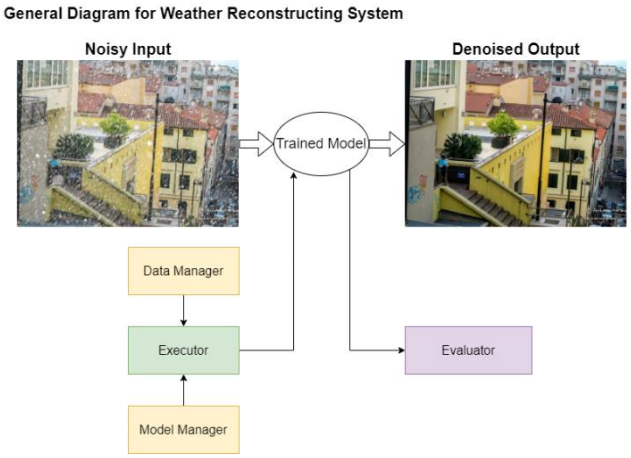

Fig. 1 The built foundation of our weather removal system.

Any vehicle during bad weather could face dangers while hindering the driver's and camera imaging views. Autonomous driving transportation has the same problem with object detection during bad weather, as they occlude potential danger scenes through the windshield or camera lens. Thus, predicting a clean image without weather is vital for protecting drivers and passengers and enhancing the reliability of object detection. From Fig. 1, we created the foundation for removing the weather effects, such as rain, snow, fog, and raindrops on the windshield, and reconstructing images without any trace of weather effects. The problems we face in even highly sophisticated generative architectures, such as pix2pix and diffusion models, are the problem of reconstructing images without any trace of weather effects. In this paper, we propose the SFHformer [1] architecture to clear out images' weather effects and apply the sub-data reloader to shorten the training time.

II. Propose

Jiang et al. introduced the SFHformer architecture by applying the Fourier transformation to identify different weather effects with the frequency domain feature maps. [1] The innovative architecture not only extracts the spatial feature map of the image locally and globally. The frequency domain feature maps represented well on multiple weather degradations with edges of snow and rain, the gradient strength of blur, and foggy weather. It has robust weather removal and is lightweight, as it replaced the transformer architecture with its own feed-forward network due to the high computational cost. The equation for the loss function is used for the L1 loss function for comparing the pixel-by-pixel between the generated image and ground truth, and frequency-domain feature maps between the generated and ground truth images. The lambda as the hyperparameter for the frequency domain feature map is 0.1, to less emphasize the frequency domain map results. The equation of total loss for training follows as,

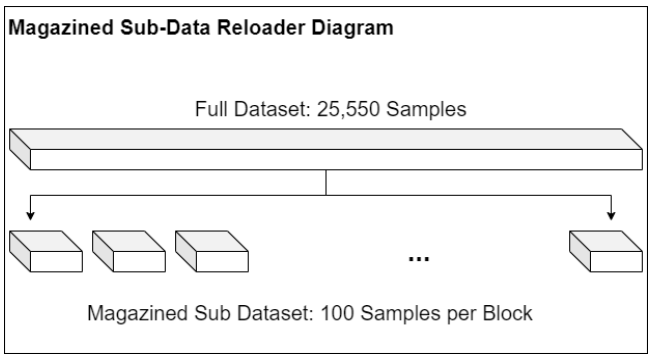

Fig. 2 The diagram of the proposed sub-data reloader.

$$L = \left\|\hat{I} - G\right\|_1 + \lambda\left\|F(\hat{I}) - F(G)\right\|_1 \quad (1)$$

Where $\hat{I}$ is the reconstructed pixels from the SFHformer's prediction, and G is the ground truth image from the KITTI All Weather dataset.

After obtaining datasets of the KITTI [2] and all-weather[3], we merged both datasets into one and produced our synthetic weather effects on those samples, which we call KITTI All Weather, and every sample has the ground truth and synthetic images in pairs. The synthetic images may have fog, rain, raindrops, and snow. The dataset has 25,000 samples in pairs for training and 7518 samples in pairs for testing.

From Fig. 2, our sub-data reloader can partially load from a whole trainable dataset. We randomly load only 8,000 samples in every epoch out of the 17,500 trainable samples, and every epoch has a different subgroup of the trainable dataset. In contrast, the whole validating samples were loaded, as the validating dataset only has 7,500. Thus, the duration of training time reduced as the trainable samples within a single epoch reduced.

III. Experimental Results

The hardware specifications were intel® Core™ i5-10600KCPU @ 4.10Ghz(12 CPUs), 4.1 Ghz, 32GB of RAM, and NVIDIA GeForce RTX 2070 Super.

From Table I, the SFHformer architecture with the whole dataset took 2782 seconds per epoch with 38.81 of the PSNR [5] and 98.23% of the SSIM [6]. After the SFHformer architecture was applied with the sub-data reloader, the PSNR and the SSIM were 36.11 and 97.43%, respectively. The time duration for a whole dataset per epoch from SFHformer was 2782 seconds, which is 46 minutes to complete one epoch. By loading only 8000 samples per epoch, the reconstruction performance could be marginally degraded, yet the time duration for training is drastically reduced. As a result, without much emphasis on the performance, sub-data reloader could be a practical method for mass producing the pre-trained model.

Table I: EXPERIMENTAL RESULTS WITH KITTI+ALL WEATHERS WITH SFHFORMERS ARCHITECTURE.

Methods	Training Time per Epoch (s)	PSNR	SSIM (%)
SFHformer [1] with whole dataset	2782	**98.23**	**38.81**
SFHformer [1] with sub-data loader within 8000 samples	**1039**	97.43	36.11

Fig. 3 The sample results from SFHformer with whole dataset.

Fig. 4 The diagram of the proposed sub-data reloader.

IV. Conclusion

In this paper, we presented how we reconstruct the image degraded by bad weather conditions and reduce the training time with such a highly performing architecture. Without training a whole dataset for each epoch, the sub-data reloader can be utilized for mass production of pre-trained models within a reduced time if performance is not the major issue.

ACKNOWLEDGMENT

This work was supported by the Commercialization Promotion Agency for R&D Outcomes(COMPA) grant funded by the Korean Government(Ministry of Science and ICT). (RS-2023-00304695.

References

[1] X. Jiang, X. Zhang, N. Gao, and Y. Deng, "When fast fourier transform meets transformer for image restoration," in European Conference on Computer Vision, 2024, pp. 381-402.

[2] M. Menze and A. Geiger, "Object scene flow for autonomous vehicles," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 3061-3070.

[3] J. M. J. Valanarasu, R. Yasarla, and V. M. Patel, "Transweather: Transformer-based restoration of images degraded by adverse weather conditions," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2022, pp. 2353-2363.

[4] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.

[5] D. Sethi, S. Bharti, and C. Prakash, "A comprehensive survey on gait analysis: History, parameters, approaches, pose estimation, and future work," Artificial Intelligence in Medicine, vol. 129, p. 102314, 2022.

[6] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE transactions on image processing, vol. 13, pp. 600-612, 2004.