



Real-time Weather Monitoring and Desnowification through Image Purification

Eliott Py, Elies Gherbi, Nelson Fernandez Pinto, Martin Gonzalez, Hatem
Hajri

► To cite this version:

Eliott Py, Elies Gherbi, Nelson Fernandez Pinto, Martin Gonzalez, Hatem Hajri. Real-time Weather Monitoring and Desnowification through Image Purification. Workshop AITA AI Trustworthiness Assessment - AAAI Spring Symposium, Mar 2023, San Francisco, CA, United States. hal-03970670

HAL Id: hal-03970670

<https://hal.science/hal-03970670>

Submitted on 2 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Real-time Weather Monitoring and Desnowification through Image Purification

Elliott Py^{*1}, Elies Gherbi^{*1}, Nelson Fernandez Pinto^{*1,2}, Martin Gonzalez^{*1}, Hatem Hajri^{*1}

¹IRT-SystemX

²Air Liquide

elliott.py@irt-systemx.fr

elies.gherbi@irt-systemx.fr

nelson.fernandezpinto@airliquide.com

martin.gonzalez@irt-systemx.fr

hatem.hajri@irt-systemx.fr

Abstract

Object detection and tracking are essential for reliable decision-making in modern applications such as self-driving cars, drones, and industry. Adverse weather can hinder object detectability and pose a threat to the reliability of these systems. As a result, there is an increasing need for efficient image denoising and restoration techniques. In this study, we investigate the use of image purification as a means of defending against weather corruptions. Specifically, we focus on the effect of snow on an object detector and the benefits of efficient desnowification. We find that the performance of a strong image purifying baseline (PreNet) is not constant across different levels of snow intensity, leading to a reduced overall performance in diverse situations. Through extensive experimentation, we demonstrate that adding a lightweight snow detector significantly improves the overall object detection performance without needing to modify the purification model. Our proposed weather-robust architecture exhibits a 40% performance improvement compared to a strong image purification baseline on the gas cylinder counting task. In addition, it leads to significant reductions of the computational power required to run the purification pipeline with a minimal added cost.

Introduction

Several real-world applications rely on vision-based sensors to navigate and achieve different tasks through their surroundings. Snow is a common weather condition that can adversely affect the performance of deep learning-based models on high-level vision tasks such as image classification, object detection, and video surveillance. These effects can include visual obstructions, changes in lighting conditions, and changes in the background of the video footage, which can all contribute to reduced accuracy and reliability of the surveillance system. As weather conditions impact these sensors, perception pipelines require considerable training on diverse data to increase the robustness of downstream tasks, and it may not be practical to collect such

large-scale datasets. Thus, it becomes essential to remove these undesirable artefacts due to severe weather conditions automatically. This problem has drawn the attention of several researchers driven by the recent impact of deep learning on the computer vision domain. It results in many approaches for denoising and restoration of the corrupted image (Image captured under weather conditions). Generative models based on GANs and Denoising Auto Encoder (Theis, Oord, and Bethge 2015; Zhang, Sindagi, and Patel 2020; Teeti et al. 2022; Gherbi et al. 2019) are mainly proposed as a proxy to purify the image as a preprocessing step for the model responsible for the downstream task like counting and recognising objects in videos. Those methods are often evaluated using Image similarity metrics like PSNR, SSIM (Wang et al. 2004) and not by their effectiveness in achieving results for the targeted application. It is possible that the output image of these methods has high image quality but also learns more about corrupted images. Hence, it achieves good results only on a corrupted image. In this work, we evaluate the effectiveness of restoration techniques for image purification to create a practical object detection pipeline suited for real-time video surveillance. To this end :

- We investigate the use of image purification to defend against weather corruption. We find that the purification performance of PreNets varies across different levels of snow intensity, leading to reduced overall performance in cases of low snow intensity.
- Our experiments have shown that using a lightweight weather monitor detector to trigger an image purification baseline can significantly improve its performance without requiring any modifications. This demonstrates the effectiveness of the approach in purifying images.
- In our proposed weather-robust pipeline, we have demonstrated a 40% improvement on the gas cylinder counting task compared to a strong image purification baseline. Additionally, this approach requires significantly less computational power to run the purification pipeline, while incurring minimal additional cost.

^{*}These authors contributed equally.

Related Work

This section briefly reviews the different components of a weather-robust pipeline for a specific downstream task, such as video object detection.

Weather corruption and restoration

Corrupted images, which are a mixture of pixel intensities from a specific weather distortion (e.g. snowfall) and a clean background image, can be decomposed into two separate images. Traditional model-based restoration methods have attempted to address this problem by incorporating regularizers on the background, and corruption images and solving for the clean image using specific optimization algorithms (Gu et al. 2017; Luo, Xu, and Ji 2015; Cantor 1978; Li et al. 2016; Chen and Hsu 2013). Works such as the atmospheric scattering model (Cantor 1978), Gaussian Mixture Model (GMM) prior (Li et al. 2016), and patch rank prior (Chen and Hsu 2013) have been adopted for modelling the background image or specific corruption weather. However, the actual composition is often more complex, and these approaches have been limited in their ability to accurately restore a clean image due to their insufficient characterization of the background and weather-corruption layers.

With the rise of deep learning and generative models (Theis, Oord, and Bethge 2015) in the computer vision field, one natural approach involves learning a direct mapping from corrupted images y to clean background images x , and to achieve this, several techniques have been developed. Approaches such as CNN-based and GAN-based image-to-image translation models (e.g., pix2pix (Qu et al. 2019), CycleGAN (Teeti et al. 2022), and perceptual adversarial networks (Wang et al. 2018a)) have been successful in modelling the underlying image background structures when applied to the problem of generating clean background images from rainy images. As a result, specific generative models have been developed for the purpose of weather restoration tasks, such as multi-task CNN architecture designed for joint detection and removal of rain streaks (Wang et al. 2018b), also in (Li et al. 2018) the author recurrently utilize dilated CNN and squeeze-and-excitation blocks to remove heavy rain streaks. Image deraining conditional GANs (Zhang, Sindagi, and Patel 2020) and conditional variational image deraining (Du et al. 2020) based on VAEs. With the recent success of diffusion models (Özdenizci and Legenstein 2022) propose Patch-Based Denoising Diffusion Models. There have also been efforts to develop lightweight deep learning models to improve computational efficiencies, such as the cascaded scheme (Fan et al. 2018) and the Laplacian pyramid framework (Fu et al. 2020). However, these approaches often result in a degradation of performance. In contrast, the PRN and PReNet models (Ren et al. 2019) introduce intra-stage recursive computation to reduce network parameters while achieving state-of-the-art deraining performance. In this study, we utilize PReNet as a baseline snow purifier and investigate ways to improve its performance under adverse conditions.

Object Detection in Adverse Weather Conditions

Object detection is a crucial task in computer vision and can be divided into two categories: region proposal-based methods and one-stage regression-based approaches (Zhao et al. 2019). Region proposal-based methods, such as those proposed by (Girshick et al. 2014; Girshick 2015), generate regions of interest (RoIs) from an image and classify them using neural networks. On the other hand, one-stage regression-based approaches, such as YOLO (Redmon et al. 2016; Redmon and Farhadi 2017, 2018; Bochkovskiy, Wang, and Liao 2020; Zhu et al. 2021), predict object labels and bounding box coordinates using a single convolutional neural network (CNN). Although object detection is well-studied, relatively few studies have focused on object detection under adverse weather conditions. One common approach to this problem is to preprocess the image using classical image enhancement methods (Guo et al. 2020; He, Sun, and Tang 2011), which are designed to remove specific weather corruption and improve image quality. However, this may not necessarily lead to improved detection performance. Other prior-based methods (Li et al. 2017; Dong et al. 2020) have attempted to jointly perform image enhancement and detection to mitigate the effects of adverse weather conditions. Under the assumption that there is a domain shift between images captured under normal and adverse weather conditions, some studies have also explored the use of domain adaptation to address this problem (Hnewa and Radha 2021; Zhang et al. 2021).

In this study, we utilize the classic one-stage YOLOv5 detector (Zhu et al. 2021) as a baseline and investigate ways to improve its performance under adverse conditions by adding a purifier model as preprocessing method of the input while focusing on preserving the detector performance and introducing less computational overhead. To this end we favor lightweight solutions to build a practical pipeline.

Methodology

In order to address the issue of poor visibility in images captured during adverse weather conditions such as heavy snowfall, we propose a pipeline of image restoration algorithms. Motivated by the recent progress achieved with state-of-the-art conditional generative models, especially weather purification based on diffusion models, we chose to implement a lightweight alternative in our pipeline. Specifically, we use PReNet (Ren et al. 2019), which can be seen as the Euler discretization of a particular neural differential equation (which we will not discuss in this work), as a substitute for the diffusion model for performance reasons to meet the practical industrial design specification. This allows us to maintain a high level of image quality while reducing the inference time to the order of the tenth of a second. As discussed above, the goal is to provide a lightweight pipeline that enables a well-trained object detection (YOLOv5) model to achieve its task in adversary weather conditions.

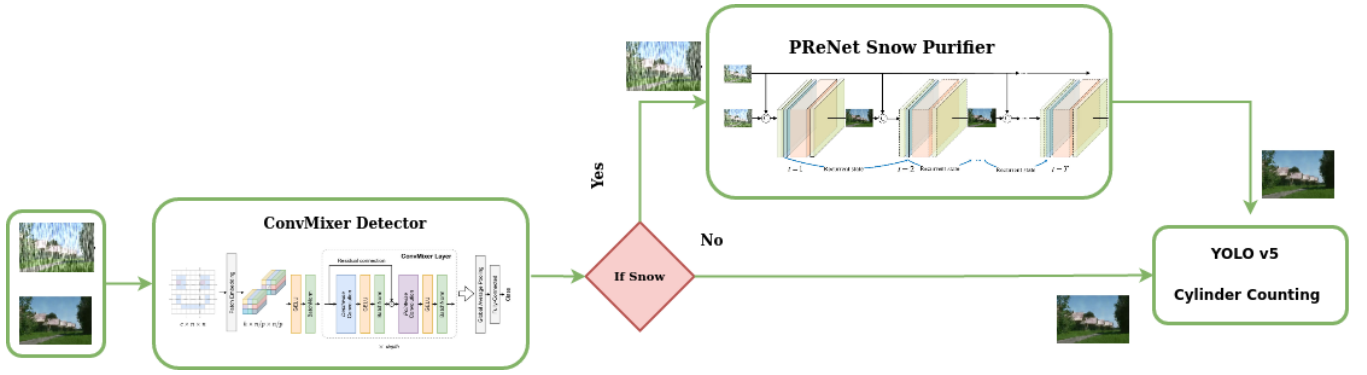


Figure 1: Proposed weather-robust cylinder counting pipeline.

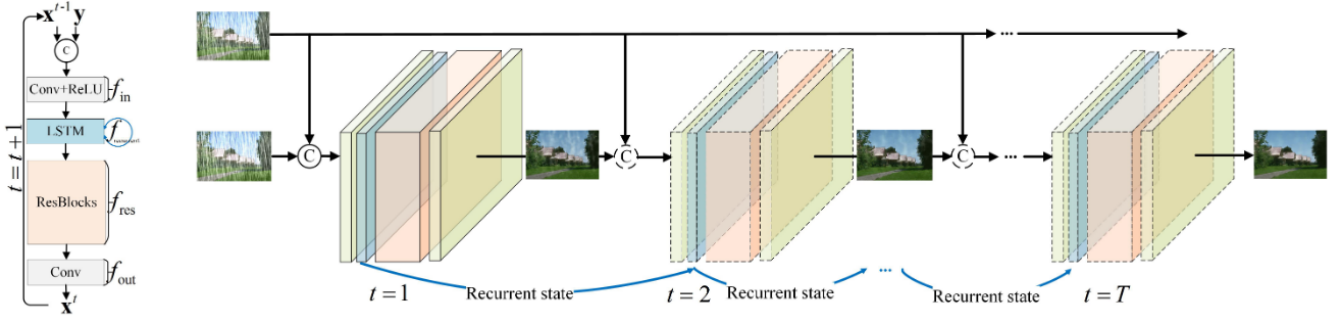


Figure 2: Illustration of PReNet architecture, a progressive network composed of a convolution layer f_{in} followed by an LSTM layer $f_{recurrent}$ and ResBlocks f_{res} into convolution layer output f_{out} . (Ren et al. 2019)

Purification

In this paper, we propose a method for purifying images that have been corrupted by the presence of snow. Our method, called "desnowification," is based on the preprocessing technique described in (Ren et al. 2019), which uses a weighted ResNet to restore a corrupted image to its original, uncorrupted form. The ResNet is implemented as a shallow network with a recurrent layer to capture the dependencies of deep features across stages. The inference of PReNet can be formulated as follow :

$$\mathbf{x}^{t-0.5} = f_{in}(\mathbf{x}^{t-1}, \mathbf{y}), \quad (1)$$

$$\mathbf{s}^t = f_{recurrent}(\mathbf{s}^{t-1}, \mathbf{x}^{t-0.5}), \quad (2)$$

$$\mathbf{x}^t = f_{out}(f_{res}(\mathbf{s}^t)), \quad (3)$$

where the recurrent layer $f_{recurrent}$ takes both $\mathbf{x}^{t-0.5}$ and the recurrent state \mathbf{s}^{t-1} as inputs at stage $t - 1$. The architecture is shown in Fig 2. A recursive procedure is used to unfold a shallow ResNet and add a recurrent layer. The is done in order to capture the dependencies of deep features across stages in the process. This learning goal is to identify the presence of snow as a corruption in the image and gradually restore a clean image by exploiting these dependencies. The objective function of PReNet with T stages is :

$$Loss = - \sum_{t=1}^T \lambda_t SSIM(\mathbf{x}^t - \mathbf{x}^{gt}) \quad (4)$$

where SSIM is the single negative structural similarity loss objective, a measure of the similarity between two images proposed by (Wang et al. 2004), and λ is a tradeoff parameter where $\lambda_{t=T} > \lambda_t$, the tradeoff parameter for the final stage is larger than the others.

One challenge of using this method for desnowification is the lack of paired data, i.e., a set of images that includes both a corrupted and a clean version. Without this data, the purifier may produce poor results on clean images, as it may darken the image in an attempt to remove the snow, and leads to an adverse examples input for YOLOv5 object detection model.

Snow monitoring

To address the issue of snow corruption in our object counting system, we have developed a real-time snow monitoring system that can detect snow events and their intensities. We have chosen to use a supervised learning approach, leveraging the data corruption pipeline we previously developed for the image purifier.

Our snow monitoring system consists of a weather-robust cylinder counting pipeline that includes both a real-time snow detector and a snow purifier. The snow purifier is activated only when the snow detector classifies the image as having an intermediate or high level of snow. For images with little to no snow, the monitoring system sends the image directly to the YOLOv5 counting classifier, saving inference

time and optimizing the overall performance of the pipeline. The Proposed weather-robust cylinder counting pipeline is illustrated in Figure 1.

For the snow detector component of our system, we have trained a ConvMixer image classifier, as proposed in (Trockman and Kolter 2022). This architecture is well-suited to our needs due to its high performance and low inference time. A straightforward model that is similar to the ViT (Visual Transformers) where it operates directly on patches as input. By separating the mixing of spatial and channel dimensions in the input (Patches), ConvMixer can maintain equal size and resolution throughout the network. In contrast, the ConvMixer uses only standard convolutions to achieve the mixing steps. However, ConvMixer outperforms the ViT-based architectures and classical vision models such as the ResNet. A detailed architecture is shown in figure 3. ConvMixer works following these steps see Figure:

- It uses a method called "tensor layout" to embed image patches in a way that keeps the information local.
- Then it applies a series of convolutional blocks, each one consisting of two types of convolution: first, a large-kernel convolution that operates on each channel separately, and second, a pointwise convolution that combines the information from all the channels.
- The number of times this convolutional block is applied is determined by the variable d .
- After all the convolutional blocks are applied, a global pooling is applied to extract features from the entire image, then A simple linear classifier is applied to make a prediction on the image.

Experiments and results

In this study, we conducted ablation studies on the purifier component to evaluate the performance of the restoration process under adverse snow conditions. We then assessed the effectiveness of our pipeline in improving the accuracy of Cylinder-Counting using YOLOv5 in these challenging conditions. The results of these experiments are presented below.

Dataset

The Air Liquide’s cylinder counting dataset contains 40 videos of around 300 images of 780 x 1024 pixels each. Images are drone-like views of trucks containing gas cylinders of different sizes. A YoloV5 object detector pre-trained on clean data to detect gas cylinders is provided as a baseline. In addition, a tracking algorithm (Bewley et al. 2016) is plugged on top of the detector to track gas cylinders across the video frames ensure sure an object is counted once time : the final count is obtained by counting the number of tracks.

To train the PReNet we select four random images from each video batch, and we apply a fixed snow corruption of high intensity see Figure 4. The images are normalized between 0 and 1 and then separated into 100x100 patches. Flipped versions of them are also considered during training to avoid overfitting.

Purifier experiments

In this study, we trained a PReNet (with a LSTM block) to purify images that had been corrupted using the image corruption framework provided by (Hendrycks and Dietterich 2019). The PReNet had a width of $T=6$, a batch size of 18, and a learning rate of $1e-3$ that was decayed three times by a factor of 5 during training. We used the Structural Similarity Index (SSIM) as the loss function to measure the perceptual quality differences between the original and purified images. The Adam optimizer was used, and the model was trained for 100 epochs on two NVIDIA A100 GPUs with 32Gb of memory each. To evaluate the performance of the trained PReNet, we applied three gradually increasing snow intensity levels (low, intermediate, and high) to the test set in a random order, following a uniform random distribution. Figure 4 shows one example of an original image and its corruptions. The results of the model’s performance on this task are presented below.

In our experiment, we applied a model to two images from a test set that were corrupted by high and intermediate intensity snow. The results, shown in Figure 6, indicate that the model was able to effectively remove the snow from the images, revealing small details of the cylinders that were previously hidden. This cleaning effect was consistent across different types of cylinders, regardless of their color and size, and the cylinders were clearly distinguishable from the surrounding context (e.g., floor, truck). These results suggest that the model is effective at purifying images from snow corruption.

In contrast, the purification performance the results in the presence of little to no snow and on clean images shown in Figure 5, indicate that the purification model resulted in overall darker images with less clearly visible small details. This is because, in the absence of snow, the purification model tended to attack small white features in the image, leading to an increase in counting error until it stabilized. As a result, the performance of the purification model was worse than the classifier alone, as shown Table 1. These results suggest that the purification model should only be applied in situations where snow is present in order to improve accuracy.

Snow level	stage						
	0	1	2	3	4	5	6
Clean	0.95	7.48	10.13	8.16	7.37	6.83	6.29
Low	1.24	1.94	2.0	1.51	1.40	1.29	1.32
Intermediate	3.10	2.16	1.54	1.21	1.08	1.05	1.05
High	11.56	3.02	1.64	1.35	1.18	1.24	1.16

Table 1: Mean counting error of the pipeline Counter+Purifier at increasing stage width of the purifier and for increasing levels of snow severity.

Table 1 shows that, as the purification unfolds, the average counting error shrinks from 3.1 cylinders/truck to reach a minima of 1.16 and 1.05 for high and intermediate snow respectively. The counting error tends to stabilize at $T=4$ afterwards further iterations are not beneficial. The lower counting error of intermediate snow indicates that the purification

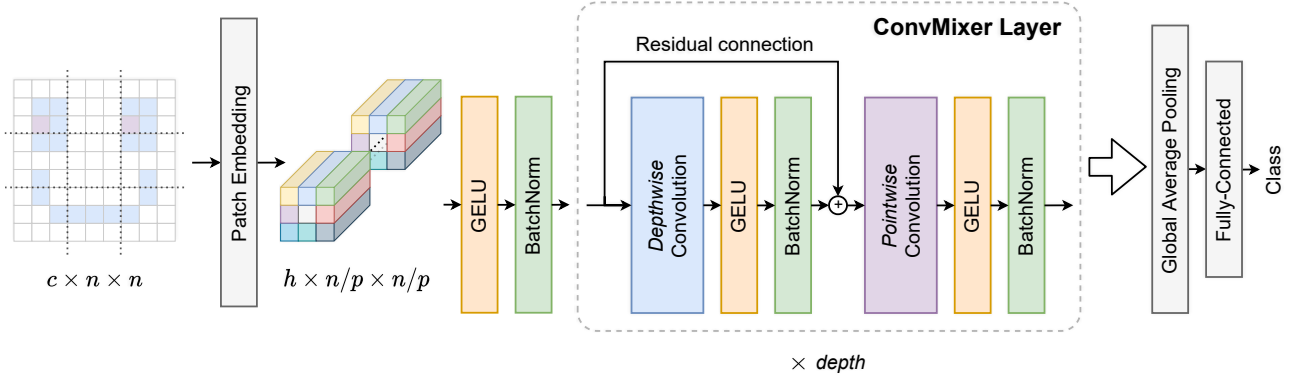


Figure 3: Illustration of ConvMixer architecture (Trockman and Kolter 2022)

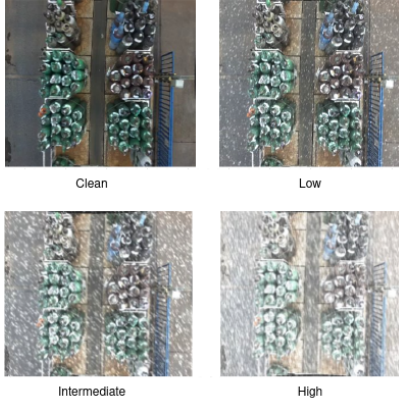


Figure 4: Illustration of examples of the original image and its corruption with different intensities of snowfall

task is easier. Comparing the obtained results with the reference clean data counting error of 0.951 cylinders/truck, we conclude that the purifying model has not only provided a perceptual but an operational advantage compared to the classifier alone on corrupted data.

Real-time weather corruption monitoring:

In light of the above results, we developed a lightweight, real-time snow monitoring that is able to detect snow events and their intensities. We trained an image classifier type ConvMixer, firstly proposed by (Trockman and Kolter 2022). This architecture is compelling due to its high performance and low inference time. The input images were resized and cropped, and the model was trained using Adam optimization with a learning rate of 0.01 and weight decay of 0.01. The model was trained on a dataset of images with four classes: clean, low intensity snow, intermediate intensity snow, and high intensity snow. To prevent overfitting,

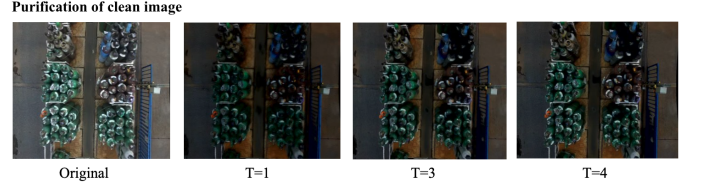


Figure 5: Illustration of the PreNet purification through different stages on a clean image.

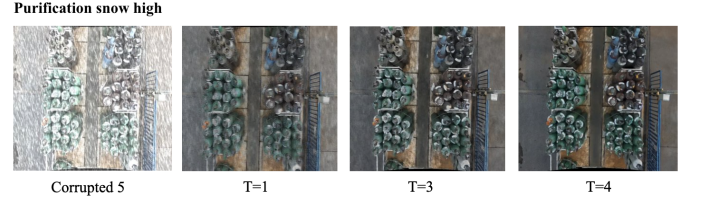


Figure 6: Illustration of the PreNet purification through different stages on a corrupted image.

we varied the corruption level of the images during training and validation. After training, the model achieved an accuracy of 81.7% on a test set of 42 videos.

The trained model demonstrated the ability of ConvMixer model to perform real-time snow intensity monitoring in images of trucks. An example of this is shown in Figure 7, where the model correctly identified intermediate and high intensity snow events with probabilities of 82% and 66%, respectively. However, the model had some difficulty correctly identifying a clean image, incorrectly labeling it as low intensity snow. This is likely because the clean image was not significantly altered by corruption, making the classification task more challenging.

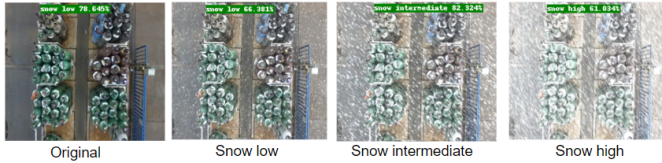


Figure 7: Examples of snow detection of several intensities.

Weather-robust Cylinder Counting pipeline results

In order to evaluate the impact of our weather-robust pipeline, we conducted a series of tests. First, we established the reference counting error and inference time of YOLOv5. Next, we created a corrupted dataset containing 10 clean videos and three collections of 10 videos with snow at intensities 1, 2, and 5, respectively. On this corrupted dataset, we evaluated the performance of YOLOv5 alone, YOLOv5 with PReNet purification preprocessing at 4 stages (YOLOv5+PReNet($t=4$)), and YOLOv5+PReNet($t=4$) passed through a trained ConvMix model that determined whether the amount of snow in the image required purification (i.e., if the snow intensity was greater than or equal to 3). In particular, the ConvMix model would send the image to PReNet($t=4$) if the snow intensity was sufficient to warrant purification. Table 2 summarizes the obtained results.

Model	Dataset	MCE	Inference time(s)
(ref) YOLO	clean	0.9512	0.06
YOLO	corrupt	3.4358	0.06
PReNet+YOLO	corrupt	1.9743	0.13
Ours	corrupt	1.1784	0.131

Table 2: Mean Counting Error (MCE) Benchmark on Air Liquide Cylinder Counting comparing **Ours** pipeline (ConvMix+PReNet+YOLO) with different preprocessing modules.

As shown in the table 2, **ours** method demonstrates superior performance across both high and low snow intensity regimes by lowering the MCE. This is particularly noteworthy as the improved accuracy of Cylinder Counting under adverse weather conditions is achieved with minimal impact on inference time compared to the PReNet+YOLO pipeline. This highlights the effectiveness of our approach in providing accurate results while maintaining computational efficiency.

Conclusion

Based on the results of our experiments, we have developed a weather-robust pipeline that is able to improve the accuracy of cylinder counting in images of trucks with varying levels of snow. Our pipeline consists of a ConvMix model that is able to detect snow events in real-time and a PReNet purification model that is effective at removing snow corruption from images. In situations with intermediate and high-intensity snow, the combination of these two models resulted in a significant reduction in counting error,

up to 10.4 cylinders per truck. However, in the absence of snow or with low-intensity snow, the purification model had a negative impact on accuracy. There are several potential directions for future work. One possibility would be to further optimize the ConvMix, and PReNet model using conditional trick (Mirza and Osindero 2014) to improve its accuracy in classifying low-intensity snow or clean images and reduce the number of preprocessing components. Additionally, it may be worthwhile to explore alternative purification models or techniques that are more effective at preserving image quality in the absence of snow (Özdenizci and Legenstein 2022) despite their lazy inference time. Overall, our weather-robust pipeline represents a promising approach for improving the accuracy of cylinder counting in challenging imaging conditions.

Acknowledgments

This work has been supported by the French government under the France 2030 program, as part of the SystemX Technological Research Institute and AirLiquide for this research. This work was conducted as part of the Con fiance.AI program, which aims to develop innovative solutions for enhancing the reliability and trustworthiness of AI-based systems. The valuable contributions of the teams at IRT-SystemX and AirLiquide have been crucial in driving this research forward and we are grateful for their support and collaboration.

References

- Bewley, A.; Ge, Z.; Ott, L.; Ramos, F.; and Upcroft, B. 2016. Simple Online and Realtime Tracking. *CoRR*, abs/1602.00763.
- Bochkovskiy, A.; Wang, C.-Y.; and Liao, H.-Y. M. 2020. YOLOv4: Optimal Speed and Accuracy of Object Detection.
- Cantor, A. 1978. Optics of the atmosphere—Scattering by molecules and particles. *IEEE Journal of Quantum Electronics*, 14(9): 698–699.
- Chen, Y.-L.; and Hsu, C.-T. 2013. A Generalized Low-Rank Appearance Model for Spatio-temporally Correlated Rain Streaks. In *2013 IEEE International Conference on Computer Vision*, 1968–1975.
- Dong, H.; Pan, J.; Xiang, L.; Hu, Z.; Zhang, X.; Wang, F.; and Yang, M.-H. 2020. Multi-Scale Boosted Dehazing Network With Dense Feature Fusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Du, Y.; Xu, J.; Zhen, X.; Cheng, M.-M.; and Shao, L. 2020. Conditional Variational Image Deraining. *IEEE Transactions on Image Processing*, 29: 6288–6301.
- Fan, Z.; Wu, H.; Fu, X.; Hunag, Y.; and Ding, X. 2018. Residual-Guide Feature Fusion Network for Single Image Deraining.
- Fu, X.; Liang, B.; Huang, Y.; Ding, X.; and Paisley, J. 2020. Lightweight Pyramid Networks for Image Deraining. *IEEE*

- Transactions on Neural Networks and Learning Systems*, 31(6): 1794–1807.
- Gherbi, E.; Hanczar, B.; Janodet, J.-C.; and Klaudel, W. 2019. An Encoding Adversarial Network for Anomaly Detection. In Lee, W. S.; and Suzuki, T., eds., *Proceedings of The Eleventh Asian Conference on Machine Learning*, volume 101 of *Proceedings of Machine Learning Research*, 188–203. PMLR.
- Girshick, R. 2015. Fast R-CNN. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*.
- Girshick, R.; Donahue, J.; Darrell, T.; and Malik, J. 2014. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Gu, S.; Meng, D.; Zuo, W.; and Zhang, L. 2017. Joint Convolutional Analysis and Synthesis Sparse Representation for Single Image Layer Separation. In *2017 IEEE International Conference on Computer Vision (ICCV)*, 1717–1725.
- Guo, C.; Li, C.; Guo, J.; Loy, C. C.; Hou, J.; Kwong, S.; and Cong, R. 2020. Zero-Reference Deep Curve Estimation for Low-Light Image Enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.
- He, K.; Sun, J.; and Tang, X. 2011. Single Image Haze Removal Using Dark Channel Prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(12): 2341–2353.
- Hendrycks, D.; and Dietterich, T. 2019. Benchmarking Neural Network Robustness to Common Corruptions and Perturbations. *Proceedings of the International Conference on Learning Representations*.
- Hnewa, M.; and Radha, H. 2021. Multiscale Domain Adaptive Yolo For Cross-Domain Object Detection. In *2021 IEEE International Conference on Image Processing (ICIP)*, 3323–3327.
- Li, B.; Peng, X.; Wang, Z.; Xu, J.; and Feng, D. 2017. AOD-Net: All-In-One Dehazing Network. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*.
- Li, X.; Wu, J.; Lin, Z.; Liu, H.; and Zha, H. 2018. Recurrent Squeeze-and-Excitation Context Aggregation Net for Single Image Deraining. In *Proceedings of the European Conference on Computer Vision (ECCV)*.
- Li, Y.; Tan, R. T.; Guo, X.; Lu, J.; and Brown, M. S. 2016. Rain Streak Removal Using Layer Priors. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2736–2744.
- Luo, Y.; Xu, Y.; and Ji, H. 2015. Removing Rain from a Single Image via Discriminative Sparse Coding. In *2015 IEEE International Conference on Computer Vision (ICCV)*, 3397–3405.
- Mirza, M.; and Osindero, S. 2014. Conditional Generative Adversarial Nets.
- Qu, Y.; Chen, Y.; Huang, J.; and Xie, Y. 2019. Enhanced Pix2pix Dehazing Network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Redmon, J.; Divvala, S.; Girshick, R.; and Farhadi, A. 2016. You Only Look Once: Unified, Real-Time Object Detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Redmon, J.; and Farhadi, A. 2017. YOLO9000: Better, Faster, Stronger. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Redmon, J.; and Farhadi, A. 2018. YOLOv3: An Incremental Improvement.
- Ren, D.; Zuo, W.; Hu, Q.; Zhu, P.; and Meng, D. 2019. Progressive Image Deraining Networks: A Better and Simpler Baseline. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Teeti, I.; Musat, V.; Khan, S.; Rast, A.; Cuzzolin, F.; and Bradley, A. 2022. Vision in adverse weather: Augmentation using CycleGANs with various object detectors for robust perception in autonomous racing. *CoRR*, abs/2201.03246.
- Theis, L.; Oord, A. v. d.; and Bethge, M. 2015. A note on the evaluation of generative models.
- Trockman, A.; and Kolter, J. Z. 2022. Patches Are All You Need?
- Wang, C.; Xu, C.; Wang, C.; and Tao, D. 2018a. Perceptual Adversarial Networks for Image-to-Image Transformation. *IEEE Transactions on Image Processing*, 27(8): 4066–4079.
- Wang, C.; Xu, C.; Wang, C.; and Tao, D. 2018b. Perceptual Adversarial Networks for Image-to-Image Transformation. *IEEE Transactions on Image Processing*, 27(8): 4066–4079.
- Wang, Z.; Bovik, A.; Sheikh, H.; and Simoncelli, E. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4): 600–612.
- Zhang, H.; Sindagi, V.; and Patel, V. M. 2020. Image De-Raining Using a Conditional Generative Adversarial Network. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(11): 3943–3956.
- Zhang, S.; Tuo, H.; Hu, J.; and Jing, Z. 2021. Domain Adaptive YOLO for One-Stage Cross-Domain Detection. In Balasubramanian, V. N.; and Tsang, I., eds., *Proceedings of The 13th Asian Conference on Machine Learning*, volume 157 of *Proceedings of Machine Learning Research*, 785–797. PMLR.
- Zhao, Z.-Q.; Zheng, P.; Xu, S.-T.; and Wu, X. 2019. Object Detection With Deep Learning: A Review. *IEEE Transactions on Neural Networks and Learning Systems*, 30(11): 3212–3232.
- Zhu, X.; Lyu, S.; Wang, X.; and Zhao, Q. 2021. TPH-YOLOv5: Improved YOLOv5 Based on Transformer Prediction Head for Object Detection on Drone-Captured Scenarios. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, 2778–2788.
- Özdenizci, O.; and Legenstein, R. 2022. Restoring Vision in Adverse Weather Conditions with Patch-Based Denoising Diffusion Models.