The Effect of Camera Data Degradation Factors on Panoptic Segmentation for Automated Driving

Yiting Wang, Haonan Zhao, Kurt Debattista, and Valentina, Donzella

Abstract— **Precise scene understanding based on perception sensors' data is important for assisted and automated driving (AAD) functions, to enable accurate decision-making processes and safe navigation. Among various perception tasks using camera images (e.g. object detection, semantic segmentation), panoptic segmentation shows promising scene understanding capability in terms of recognizing and classifying different types and objects, imminent obstacles, and drivable space at a pixel level. While current panoptic segmentation methods exhibit good potential for AAD perception under 'ideal' conditions, there are no systematic studies investigating the effects that various degradation factors can have on the quality of the data generated by automotive cameras. Therefore, in this paper, we consider 5 categories of camera data degradation models, namely light level, adverse weather, internal sensor noises, motion blur and compression artefacts. These 5 categories include 11 potential degradation models, with different degradation levels. Based on these 11 models and multiple degradation levels, we synthesize an augmented version of Cityscape, named the Degraded-Cityscapes (D-Cityscapes). Moreover, for the environmental light level, we propose a new synthetic method with generative adversarial learning and zero-reference deep curve estimation to simulate 3 degraded light levels including low light, night light with glare, and extreme light. To compare the effect of the implemented camera degradation factors, we run extensive tests using a panoptic segmentation network (*i.e.* EfficientPS), quantifying how the performance metrics vary when the data are degraded. Based on the evaluation results, we demonstrate that extreme snow, blur, and light are the most threatening conditions for panoptic segmentation in AAD, while EfficientPS can cope well with light fog, compression, and blur, which can provide insights for future research directions.**

I. INTRODUCTION

Perception through sensors (*i.e.* camera, LiDAR, RaDAR) allows automated vehicles (AVs) to understand and interpret their surroundings in order to enable the necessary safety-critical decision-making processes needed to navigate the environment. Panoptic segmentation [1] aims at simultaneously segmenting pixels at the instance and semantic level. It not only predicts categories of pixels for the background (*e.g.* sky, grass) but also countable objects (*e.g.* cars, pedestrians) [2]. Panoptic segmentation is critical for AVs to accurately determine how to navigate roads safely and efficiently, as it not only gives them a more comprehensive understanding of the surrounding environment, it also allows AV navigation to accurately track individual objects and predict their movements, which is necessary for taking informed decisions on how to react to potential hazards. For example, different instances of vehicles and pedestrians can be classified with their distinct shapes and boundary information. If several pedestrians are crossing the street, panoptic segmentation can support the AAD systems to make more accurate movement predictions for each pedestrian and adjust its speed and trajectory accordingly. Such results cannot be achieved with merely object detection, semantic segmentation, or instance segmentation [3]. Despite the benefits of this new scene-understanding method with the potential to boost the automation level for AAD [4], real-world noise factors can hinder it from full deployment on existing AAD systems.

Fig. 1. Visual examples of the newly proposed Degraded-Cityscapes (D-Cityscapes) with 11 types of degradation, from the top to the bottom columns are categorized as light level, adverse weather, internal sensor noises, motion blur and compression artefacts.

Most of the existing panoptic segmentation methods are trained with datasets such as Cityscapes [5], KITTI [6] and COCO[7]. As these datasets are captured under good weather conditions with high visual quality, the data-driven panoptic segmentation models trained on them will show performance degradation when faced with various unseen or less-seen noise factors as are common in real-world scenarios. Examples of these degradation factors have been studied in details in [8], and include adverse weather [9], [10], [11], sensor

Funded by the European Union (grant no. 101069576). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Climate, Infrastructure and Environment Executive Agency (CINEA). Neither the European Union nor the granting authority can be held responsible for them. UK and Swiss participants in this project are supported by Innovate UK (contract no. 10045139) and the Swiss State Secretariat for Education, Research and Innovation (contract no. 22.00123) respectively.

The work was partially supported by High Value Manufacturing CATAPULT. This research is partially sponsored by the Centre for Doctoral Training to Advance the Deployment of Future Mobility Technologies (CDT FMT) at the University of Warwick.

Yiting Wang, Haonan Zhao, Kurt Debattista, and Valentina, Donzella are with WMG, University of Warwick,Coventry CV4 7AL, UK. Corresponding author: `Yiting.Wang.1@warwick.ac.uk`

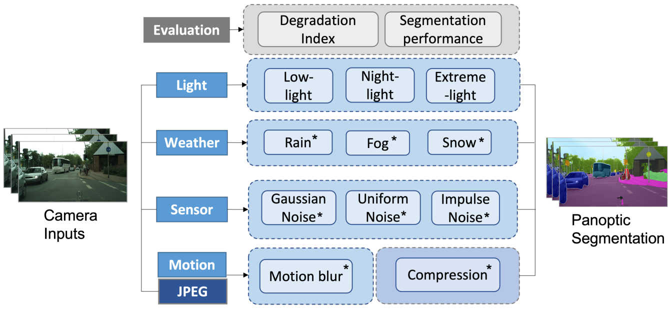

Fig. 2. Experimental pipeline and classification of the 11 degradation models. "*" denotes three different degradation levels modelled for that factor.

internal noise [12], compression[13], [4] and unfavourable lighting scenarios [14], [15]. The lack of robustness in the existing panoptic segmentation methods can lead to safety problems in AAD functions. Therefore, researchers have conducted experiments to compare the effect of various camera data degradation factors by using the evaluation with image classification [16], [4], object detection [12], [17], and semantic segmentation [18]. Nevertheless, none of the above tasks can provide the same level of detail as panoptic segmentation. Due to the fact that noise caused by the various degradation factors is randomly distributed in the whole image, a better way of evaluation from the machine vision perspective with the ability to indicate the effect of corruption at a pixel level is needed.

Existing datasets featuring diverse degradations and panoptic labels exhibit uneven distributions of degradation factors, resulting in reduced generality and reliability when comparing factors, *e.g.* in BDD100k only the 0.2% of the images is collected in foggy conditions [19]. Furthermore, degraded frames have often lower numbers of dynamic objects, such as fewer pedestrians and vehicles at nighttime. The scarcity of extreme degradation levels within these datasets, which primarily capture mild severity [14], [11], limits their comprehensiveness. Lastly, while soem previous studies have addressed adverse weather, noise, and blur, the impact of scene illumination has received less attention, Tab.III-B.

To solve these challenges, in this paper, we apply 11 types of degradation models to the datasets (29 variations in total when considering the severity levels for each degradation), see Fig. 1). The degradation models designed for AAD are categorized into *light level, adverse weather, internal sensor noises, motion blur, compression artefacts*, Fig. 2). Following [10], each noise, except the light level, has three severity levels. We use these degradation models to augment the clean Cityscapes to compile a new Degraded-Cityscapes (D-Cityscapes), which can serve as a useful dataset for future research for scene understanding with AAD camera data degradation, Fig. 1.

We conduct experiments on D-Cityscapes to evaluate the effect of each one of the considered types of degradation on panoptic segmentation. The contributions of this work are: (**I**) We are the first to compare the effect of 11 camera data degradation factors on panoptic segmentation for AAD; (**II**) we augment the D-Cityscapes dataset with 11 types of degradation factors (29 types considering the severity levels of degradation) to boost the future robustness research for automated and autonomous driving.

The presented results demonstrate that panoptic segmentation is particularly affected by some of the degradation factors considered (*e.g.* internal noise, snow) and the design of panoptic segmentation network in the future needs to be optimised to be robust to the AAD noise factors.

II. Related Work

Degradation Simulation There are mainly two different methods for image simulation, which are the conventional methods and the DL-based methods. Conventional methods generate images under various weather conditions or simulate noises based on their physical models [20], [21], [9], [22], [23]. The DL-based methods are mostly data-driven methods while some also use the physical priors as guidance [9]. Besides that, there are also some works that render images from virtual environments for various simulations. For example, the SHIFT synthetic dataset [24] is generated through the CARLA simulator. However, they lack a sense of realism and there is a huge domain gap between these synthetic images and the real-world images [24]. There is no perfect simulation method, the conventional methods require complex modelling and formulations which are highly dependent on certain scenarios and assumptions but also with better explainable ability. The DL-based methods, on the other hand, can flexibly generate more visually pleasant images, however, they are highly dependent on the dataset without easy solutions to control the degradation severity levels. Furthermore, the performance of the DL-based methods, especially the GAN-based methods do not perform consistently (*e.g.* may fail to capture the structural consistency or darken the image) [25]. *Therefore, we choose a combination of both conventional or DL-based methods for the D-Cityscapes simulation.*

Panoptic Segmentation The existing DL-based panoptic segmentation methods can be categorized into top-down methods, [1], [26], [27], bottom-up methods [28], [29], single path methods [30] and other methods[31], [32]. For example, panoptic FPN [1] is first proposed to connect the instance segmentation and semantic segmentation through a shared feature pyramid network. UPSNet [26] proposed a unified network with a parameter-free panoptic head that classifies unknown class pixels to solve the confusion between the 'stuff' and the 'things'. DeeperLab [29] predicts the instance segmentation depending on the corner and centre of the bounding box for class-independent prediction. Similarly, DeepLab [28] uses the foreground mask and the center prediction and regression for instance segmentation. They abandon the processes of proposal generation and post-process to have better real-time ability. Different from the previous methods, FCN Panoptic [30] predicts the stuff and things with kernel generation to directly produce the results. Recently, a new and powerful transformer structure is also implemented in the panoptic segmentation methods [32]. SegFormer [32] considers the robustness of the natural corruptions on the COCO-C dataset, however, it does not discuss the robustness of the complex driving scenes and the low-light corruptions are not being considered.

III. Methodology

In this section, we introduce the overall experimental pipeline, which consists of the selected or created degradation models, the dataset augmentation, and the evaluation via panoptic segmentation using specific metrics for comparison. We highlight our proposed light model and put it in the first category of the degradation simulation models.

A. Degradation Simulation

In the degradation simulation, we categorized the degradation factors into *light level, adverse weather, internal sensor noises, motion blur, and compression artefacts* (See Fig. 2).

Light Level For light levels we propose a new model to generate the data. There exist complex light conditions for real-world AAD, especially at nighttime with multiple light sources. It will not only produce unfavorable light conditions where the camera can capture images with less information but also influence the perception accuracy [33]. Based on the level of illumination and glare, we transfer the input images I into low light l, night light n, and extreme light e images. The low light refers to the uniform darkening of the images. The night light simulates AAD in urban areas with traffic lights, headlights, and colorful street lights with a glare effect. Extreme light refers to darker illumination compared with night light. Gamma correction is widely used for adjusting the light levels[34], which is defined via the equation: $l = \alpha I^{\gamma}(\gamma > 1)$, where constant α is usually set to 1 and the light can be adjusted non-linearly when changing the values of γ. While this works well for static images taken under evenly distributed illumination conditions, they are not suitable for simulating real-world low-light images, with uneven light distribution and saturated pixels from various light sources. Therefore, to generate more realistic and natural low-light driving images, we use the EC-Zero-DCE [35] to retain the saturated pixels while darkening the other regions of the images via implementing a reversed curve adjustment. Therefore, we can obtain the synthetic low-light Cityscapes via $l = EC_{dce}(I, \theta)$, where EC_{dce} is the darkening process with the pre-trained darkening model θ. In addition, nighttime images are always combined with flare and glare effects which cannot be simply generated by the darkening process. Therefore, we use cycleGAN [25] to simulate the night light images with the following cycle consistency loss function.

$$\begin{aligned} L_{cyc}(G,F) = E_{d\sim p_{data}(I)}\|F(G(d)) - d\|_1 \\ + E_{n\sim p_{data}(n)}\|G(F(n)) - n\|_1 \end{aligned} \tag{1}$$

Where the night images n and daytime d can be generated from d with the generators $G(d)$ and $F(G(d))$, respectively. Another cycle reverses the process by generating d from n with the two generators. This process is capable of learning the day-to-night pixel-wise image-to-image translation with glares and flares added to the images.

For darkening images, we find that the performance of cycleGAN is not sufficient as the illumination of a small portion of images remains broadly the same as in the daytime. Therefore, the extreme light Cityscapes can be obtained by $e = EC_{dce}(n, \theta)$, where we perform the EC-Zero-DCE [35] on the cycleGAN generated night images to make sure each picture is properly or further darkened.

Weather Conditions' Models Automated driving frequently experiences weather changes, which may seriously affect camera sensors. For example, the particles in raindrops, snow, and fog can heavily affect visibility [14]. The sight of visual information at far distances under extreme weather conditions is easily compromised. Therefore, we consider the three most common weather conditions for AAD: *rain, fog, and snow*. For convenience, we choose several existing multi-weather datasets from Foggy Cityscapes [20], Rain Cityscapes [9] and Snow Cityscapes [23], respectively, to constitute our multi-weather data. Foggy Cityscapes [22] proposed an automatic pipeline by modelling the fog effect with a mapping function of the clean outdoor scene radiance into the camera sensor radiance observation, using the depth information. Rain Cityscapes [20] synthesized rain streaks images based on the theory from Garg and Nayar [21] that the visual intensity of a rain streak depends on the scene depth. Both datasets can be found on the official website of the Cityscapes. Photoshop was used in SnowCityScapes [23] to create three levels of on-street synthetic snow (*i.e.* light, medium and heavy).

Sensor Noise Model The camera sensors of AVs can suffer from many types of noises when they are impacted by internal or external variables [12]. Following the recent paper [12], we consider 3 types of noise: Gaussian noise, Uniform noise, and Impulse noise.

1) Gaussian Noise. We add different levels of Gaussian noise to images in the validation set by generating random samples from a normal Gaussian distribution with several specified values of mean {0} and standard deviation {5, 15, 25}. Therefore the noisy image $I^n(x, y)$ can be formulated as $I^n(x, y) = J(x, y) + N(x, y)$ where J is the noise-free ideal image, and N is the noise which satisfies the probability density function of the Gaussian distribution: $P(N) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{N^2}{2\sigma^2}\right)$, with zero-mean and variance standard deviation σ.

2) Uniform Noise. We add different levels of Uniform noises to images in the validation set by using a random number generator that follows the probability distributions:

$$\begin{aligned} P(N) = \begin{cases} \frac{1}{b-a} & \text{if } a \le N \le b \\ 0 & \text{otherwise} \end{cases} \\ \mu = \frac{a+b}{2} \quad \sigma^2 = \frac{(b-a)^2}{12} \end{aligned} \tag{2}$$

where P(N) represents the probability density function of noise N, which follows a uniform distribution. The values

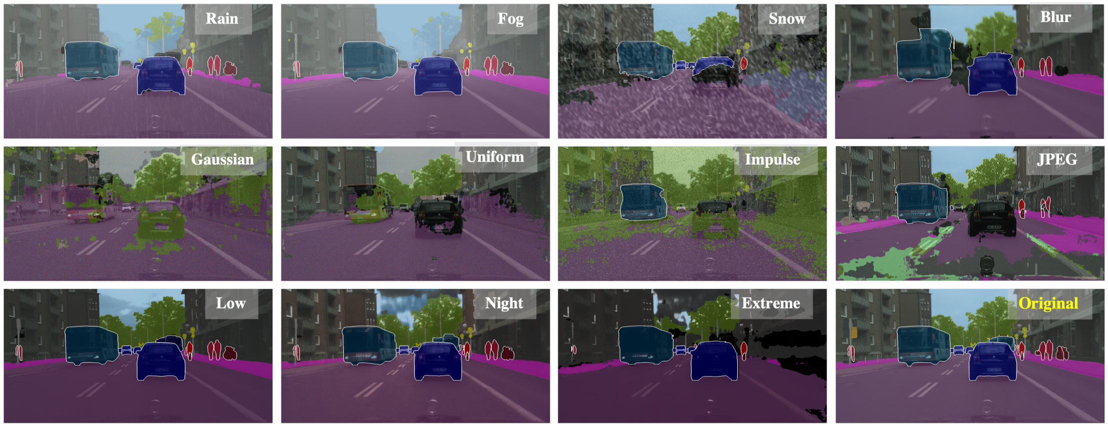

Fig. 3. Visual results of the panoptic segmentation under different degradation. Note that the conditions from left to right, in the first row are: rain, fog, snow, and motion blur. The second row is Gaussian noise, uniform noise, impulse noise, and compression. The third rows are low-light, night light, and extreme light and the original Cityscapes.

of the mean μ $\{0\}$ and standard deviation σ $\{25, 50, 75\}$ of the distribution are applied in order to simulate different levels of noise.

3) Impulse Noise. We add different levels of Impulse noise to images in the validation set. We define different values of probabilities $\{0.01, 0.02, 0.03\}$ for salt and pepper noise, which are the two types of impulse noise. Then the probabilities are used to generate various random binary masks that specify the locations where to add noise.

Motion Blur and Compression Model The unavoidable relative motion between a camera sensor and a scene will result in motion blur in camera pictures as camera sensors collect data by accumulating incoming light over time. In this paper, we use the imgaug library [36] to implement motion blur and use a predefined severity $\{1, 3, 5\}$ to simulate different levels of motion blur. Previous work already demonstrated the need of reducing the data size for transferring the camera images and the perception performance will be degraded with different compression rates [13], [4]. Here, we use the imgaug library [36] to apply different levels of JPEG compression to images in the validation set. The compression rates are set to 74.94%, 58.80%, 42.20%, with the three different compression indexes [20, 50, 80].

B. *Models and Evaluation Metrics*

Panoptic Segmentation Model According to the survey [2], compared with other SOTA panoptic segmentation methods [28], [1], [29], [26] which are also trained on the Cityscapes dataset, EfficeintPS [27] achieves the best performance in terms of time and quality. EfficeintPS speeds up the long-reference time caused by the proposal generation with the proposed bidirectional FPN and efficient EfficientNet architecture.

Evaluation Metrics We use the commonly used panoptic quality (PQ) metric for evaluating the performance of the panoptic model on the different degradation factors [1]. PQ is the combination of both the segmentation quality (SQ) and the recognition quality(RQ), which can be formulated as $PQ = SQ \times RQ$. SQ, is similar to the segmentation evaluation metrics, where the results are considered a match only when IOU is greater than 0.5 for the prediction to be overlapping with the ground truth. RQ is the F1 score, where true positive (TP), false positive (FP), and false negative (FN) represent the correct matches, incorrect matches, and missed matches, respectively. Therefore, the PQ can be formulated as:

$$PQ = \frac{\sum_s^g IOU(s,g)}{|TP|} \times \frac{|TP|}{|TP| + \frac{1}{2}|FP| + \frac{1}{2}|FN|} \tag{3}$$

where s represents the segmentation results, while g represents the ground truth, and $(s,g) \in TP$. In our evaluation experiments, the larger the PQ values the better the quality. Since the index PQ indicates the performance of both the segmentation and detection, it acts as the perfect metric for us to validate the degradation of various factors.

TABLE I

COMPARISON OF RELATED METHODS IN TERMS OF: WEATHER, NOISES, BLUR, COMPRESSION, LIGHT, AND DEGRADATION LEVELS.

	Task	Wea.	Noi.	Blu.	Com.	Lig.	Lev.
[37]'21	sem. seg.	✓	✓	✓	✓	x	x
[14]'22	pan. seg.	✓	x	✓	x	✓	✓
[12]'23	3D det.	✓	✓	✓	x	x	✓
Ours	pan. seg.	✓	✓	✓	✓	✓	✓

C. *Implementation*

Settings for Training Light Model The experiments were performed using a Ubuntu 20, a Quadro P5000 GPU serving as the tensor core, and a Conda environment with Python 3.8. The neural networks were trained and tested using PyTorch. BDD100K dataset [19] which contains both daytime and nighttime driving images is used for training the cycleGAN network for the day-to-night simulation, with 2k epochs.

Degraded-Cityscapes We apply the aforementioned 11 degradation factors with different levels on the

TABLE II

THE PQ VALUES FOR CLEAN AND D-CITYSCAPES VALIDATION SET WITH EFFICIENTPS. LEVELS 3, 2, AND 1 REPRESENT HEAVY, MEDIUM, AND LIGHT FOR WEATHER, NOISE, MOTION, AND COMPRESSION WHEREAS EXTREME LIGHT, NIGHT LIGHT, AND LOW LIGHT FOR LIGHT.

		Light	**Weather**				**Internal Sensor Noise**			**Motion**	**Compression**
Severity	Metrics	Light	Clean	Rain	Fog	Snow	Gaussian	Uniform	Impulse	Blur	JPEG
Level 3	PQ	24.0	62.8	33.5	34.0	22.4	4.4	9.7	7.2	11.9	17.5
	PQ^{Th}	23.9	57.8	30.4	29.1	15.9	0.1	3.1	5.0	3.1	19.5
	PQ^{St}	24.0	66.4	35.4	37.5	27.1	7.5	14.5	8.7	18.3	16.0
Level 2	PQ	44.0	62.8	42.0	47.3	32.1	14.9	15.8	15.7	26.3	40.2
	PQ^{Th}	42.8	57.8	37.1	42.4	25.6	9.9	10.0	18.0	14.5	39.0
	PQ^{St}	44.8	66.4	45.2	50.8	36.3	18.6	20.1	13.9	34.9	41.0
Level 1	PQ	45.7	62.8	49.6	55.2	43.9	52.3	31.6	40.2	49.1	53.8
	PQ^{Th}	43.2	57.8	44.8	49.7	37.7	47.9	28.4	41.5	42.0	48.7
	PQ^{St}	47.5	66.4	52.6	59.1	48.4	55.5	33.9	39.3	54.2	57.6

clean Cityscapes [5] to get the Degraded-Cityscapes (D-Cityscapes). The Cityscapes dataset is the most frequently used dataset in panoptic segmentation for AAD. The Cityscapes dataset contains high-quality daytime images from 50 European cities with 19 classes of pixel-level and 30 classes of instance-level annotations [5]. These make it an ideal clean dataset for us to inject various degradation factors, with the pixel-wise and instance-wise panoptic segmentation ground truth for validation.

All 500 images are used for validation. We resize the image resolution into 1024×512 using bicubic interpolation, for a faster and more uniform comparison.

IV. RESULTS AND DISCUSSION

We show our results with EfficeintPS under different degradation factors in Tab. II in terms of the overall average PQ and the PQ towards objects (*i.e.* person, car, rider, bus) PQ^{Th} and background (*i.e.* road, building, traffic light, sky) PQ^{ST}. Higher values indicate the image is less affected by corruption. Level 3, 2, and 1 represents heavy, medium, and light for the severity of the weather, internal sensor noises, motion blur and compression, whereas for light, they represent the low-light, low-light with glares (night light), and darker low-light with glares (extreme light). Visual examples under 11 degradation factors are illustrated in Fig. 3). Based on the overall tables and figures, we can draw several insightful conclusions regarding the impact of different degradation factors on the perception models trained on clean datasets. These findings are valuable to the intelligent vehicle community to develop better-performing automated vehicles in challenging scenarios. We discuss the results for each category below.

Light Level. Our results reveal that extreme light conditions have the most substantial degradation effect on panoptic segmentation, followed by night level and low-light scenarios. This highlights the need for developing robust perception models that can handle complex light distribution and poor illumination conditions, as these situations pose considerable challenges to automated driving systems.

Adverse Weather. Among adverse weather conditions, snow has the most considerable impact on panoptic segmentation performance, followed by rain and fog. This observation aligns with the fact that snow particles are larger than rain or fog particles, leading to a more significant obstruction to the camera sensors. It is crucial for the AAD community to design systems capable of handling such challenging weather conditions to ensure safety and reliability.

Internal Sensor Noise. As illustrated from the visual and the quantitative results, severe sensor noise, particularly Gaussian noise, causes the most significant degradation in performance. Given the vast variations in performance between different severity levels, it is essential for perception models to be resilient to various types and degrees of sensor noise to ensure reliable operation in real-world environments.

Motion Blur. The motion blur degradation primarily affects moving objects rather than the background (*e.g.* 18.3), resulting in significantly lower PQ values for objects (*e.g.* 3.1). This observation suggests that future research should focus on improving the perception models ability to handle scenes with high-speed objects and severe motion blur.

Compression Artefacts. The impact of compression artefacts on performance is relatively mild compared to other degradation factors. However, it is worth noting that the change in PQ values remains stable within a certain range of compression ratios, indicating that perception models should still be able to handle mild compression artifacts without significant performance loss.

V. CONCLUSION

This paper establishes a unifying framework on the newly proposed D-Cityscapes dataset, that is generated by injecting Cityscapes with 5 categories of degradation factors: light level, adverse weather conditions, internal sensor noises, motion blur, compression artefacts. We compare how different degradation factors affect the accuracy of a panoptic segmentation with EfficientPS. According to the evaluation findings, we illustrate that panoptic segmentation is most affected by severe snow, blur, and extreme light in AAD. On teh contrary, EfficientPS performance are less degraded in handling light fog, compression, and blur. These findings indicate potential areas for future research exploration in the field of AAD function robustness. Furthermore, we anticipate that our generated D-Cityscapes will be useful for future research as well as for fairly and comprehensively comparing the robustness of panoptic segmentation techniques.

Limitation and future work. In this work, camera data degradation has been quantified using one SOTA panoptic segmentation method, more networks need to be explored to generalise the results. Moreover, in the future, better and validated simulation models can be used to improve the realism of the dataset.

REFERENCES

[1] A. Kirillov, K. He, R. Girshick, C. Rother, and P. Dollár, "Panoptic segmentation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 9404–9413.

[2] X. Li and D. Chen, "A survey on deep learning-based panoptic segmentation," *Digital Signal Processing*, vol. 120, p. 103283, 2022.

[3] C. Wei, "Vehicle detecting and tracking application based on yolov5 and deepsort for bayer data," in *2022 17th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, 2022, pp. 843–849.

[4] Y. Wang, P. H. Chan, and V. Donzella, "Semantic-aware video compression for automotive cameras," *IEEE Transactions on Intelligent Vehicles*, 2023.

[5] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The cityscapes dataset for semantic urban scene understanding," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 3213–3223.

[6] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The kitti dataset," *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1231–1237, 2013.

[7] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*. Springer, 2014, pp. 740–755.

[8] B. Li, P. H. Chan, G. Baris, M. D. Higgins, and V. Donzella, "Analysis of automotive camera sensor noise factors and impact on object detection," *IEEE Sensors Journal*, vol. 22, no. 22, pp. 22 210–22 219, 2022.

[9] M. Tremblay, S. S. Halder, R. De Charette, and J.-F. Lalonde, "Rain rendering for evaluating and improving robustness to bad weather," *International Journal of Computer Vision*, vol. 129, pp. 341–360, 2021.

[10] O. Zendel, K. Honauer, M. Murschitz, D. Steininger, and G. F. Dominguez, "Wilddash-creating hazard-aware benchmarks," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 402–416.

[11] V. Mușat, I. Fursa, P. Newman, F. Cuzzolin, and A. Bradley, "Multi-weather city: Adverse weather stacking for autonomous driving," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 2906–2915.

[12] Y. Dong, C. Kang, J. Zhang, Z. Zhu, Y. Wang, X. Yang, H. Su, X. Wei, and J. Zhu, "Benchmarking robustness of 3d object detection to common corruptions in autonomous driving," *arXiv preprint arXiv:2303.11040*, 2023.

[13] Y. Wang, P. H. Chan, and V. Donzella, "A two-stage h. 264 based video compression method for automotive cameras," in *2022 IEEE 5th International Conference on Industrial Cyber-Physical Systems (ICPS)*. IEEE, 2022, pp. 01–06.

[14] O. Zendel, M. Schörghuber, B. Rainer, M. Murschitz, and C. Beleznai, "Unifying panoptic segmentation for autonomous driving," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 21 351–21 360.

[15] C. Guo, C. Li, J. Guo, C. C. Loy, J. Hou, S. Kwong, and R. Cong, "Zero-reference deep curve estimation for low-light image enhancement," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 1780–1789.

[16] Y. Dong, Q.-A. Fu, X. Yang, T. Pang, H. Su, Z. Xiao, and J. Zhu, "Benchmarking adversarial robustness on image classification," in *proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 321–331.

[17] P. H. Chan, A. Huggett, G. Souvalioti, P. Jennings, and V. Donzella, "Influence of avc and hevc compression on detection of vehicles through faster r-cnn," 2022.

[18] C. Kamann and C. Rother, "Benchmarking the robustness of semantic segmentation models," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 8828–8838.

[19] F. Yu, H. Chen, X. Wang, W. Xian, Y. Chen, F. Liu, V. Madhavan, and T. Darrell, "Bdd100k: A diverse driving dataset for heterogeneous multitask learning," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 2636–2645.

[20] X. Hu, C.-W. Fu, L. Zhu, and P.-A. Heng, "Depth-attentional features for single-image rain removal," in *Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition*, 2019, pp. 8022–8031.

[21] K. Garg and S. K. Nayar, "Vision and rain," *International Journal of Computer Vision*, vol. 75, pp. 3–27, 2007.

[22] C. Sakaridis, D. Dai, and L. Van Gool, "Semantic foggy scene understanding with synthetic data," *International Journal of Computer Vision*, vol. 126, pp. 973–992, 2018.

[23] K. Zhang, R. Li, Y. Yu, W. Luo, and C. Li, "Deep dense multi-scale network for snow removal using semantic and depth priors," *IEEE Transactions on Image Processing*, vol. 30, pp. 7419–7431, 2021.

[24] T. Sun, M. Segu, J. Postels, Y. Wang, L. Van Gool, B. Schiele, F. Tombari, and F. Yu, "Shift: a synthetic driving dataset for continuous multi-task domain adaptation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 21 371–21 382.

[25] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223–2232.

[26] Y. Xiong, R. Liao, H. Zhao, R. Hu, M. Bai, E. Yumer, and R. Urtasun, "Upsnet: A unified panoptic segmentation network," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 8818–8826.

[27] R. Mohan and A. Valada, "Efficientps: Efficient panoptic segmentation," *International Journal of Computer Vision*, vol. 129, no. 5, pp. 1551–1579, 2021.

[28] B. Cheng, M. D. Collins, Y. Zhu, T. Liu, T. S. Huang, H. Adam, and L.-C. Chen, "Panoptic-deeplab: A simple, strong, and fast baseline for bottom-up panoptic segmentation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 12 475–12 485.

[29] T.-J. Yang, M. D. Collins, Y. Zhu, J.-J. Hwang, T. Liu, X. Zhang, V. Sze, G. Papandreou, and L.-C. Chen, "Deeperlab: Single-shot image parser," *arXiv preprint arXiv:1902.05093*, 2019.

[30] Y. Li, H. Zhao, X. Qi, L. Wang, Z. Li, J. Sun, and J. Jia, "Fully convolutional networks for panoptic segmentation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 214–223.

[31] Q. Yu, H. Wang, D. Kim, S. Qiao, M. Collins, Y. Zhu, H. Adam, A. Yuille, and L.-C. Chen, "Cmt-deeplab: Clustering mask transformers for panoptic segmentation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 2560–2570.

[32] Z. Li, W. Wang, E. Xie, Z. Yu, A. Anandkumar, J. M. Alvarez, P. Luo, and T. Lu, "Panoptic segformer: Delving deeper into panoptic segmentation with transformers," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 1280–1289.

[33] C.-T. Lin, S.-W. Huang, Y.-Y. Wu, and S.-H. Lai, "Gan-based day-to-night image style transfer for nighttime vehicle detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 2, pp. 951–963, 2020.

[34] F. Lv, F. Lu, J. Wu, and C. Lim, "Mbllen: Low-light image/video enhancement using cnns." in *BMVC*, vol. 220, no. 1, 2018, p. 4.

[35] S. Zhou, C. Li, and C. C. Loy, "Lednet: Joint low-light enhancement and deblurring in the dark," in *ECCV*, 2022.

[36] A. B. Jung, K. Wada, J. Crall, S. Tanaka, J. Graving, C. Reinders, S. Yadav, J. Banerjee, G. Vecsei, A. Kraft, Z. Rui, J. Borovec, C. Vallentin, S. Zhydenko, K. Pfeiffer, B. Cook, I. Fernández, F.-M. De Rainville, C.-H. Weng, A. Ayala-Acevedo, R. Meudec, M. Laporte *et al.*, "imgaug," https://github.com/aleju/imgaug, 2020, online; accessed 01-Feb-2020.

[37] C. Kamann and C. Rother, "Benchmarking the robustness of semantic segmentation models with respect to common corruptions," *International journal of computer vision*, vol. 129, pp. 462–483, 2021.