Supplementary Material : Lightweight Thermal Super-Resolution and Object Detection for Robust Perception in Adverse Weather Conditions

Pranjay Shyam
Faurecia IRYStec Inc.
Montreal, Canada
pranjay.shyam.psm@forvia.com

HyunJin Yoo
Faurecia IRYStec Inc.
Montreal, Canada
hyunjin.yoo@forvia.com

1. Appendix A

In this section we provide qualitative evaluation of histogram peak driven image degradation detection. From Fig. 2 we can observe that using histogram peaks we can easily identify different degradation conditions using hard coded thresholded values, represented by β_1, β_2 in the main paper.

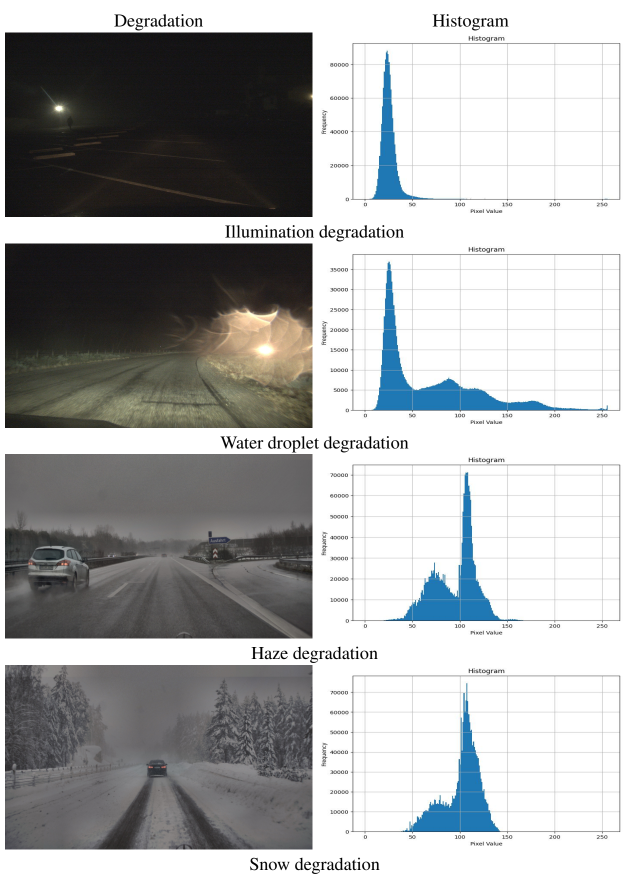


Figure 1. Qualitative evaluation of mechanism to detect different weather degradations using images from FLIR [1] and DENSE [2] datasets.

2. Appendix B

In this section we briefly describe the mechanism for generating bounding box annotations using a combination of segment anything model and domain adaptation version of panoptic segmentation to ensure accurate bounding box information for objects of interest. We provide a brief overview of the mechanism in Fig. 2 with dataset summaries in Tab. 1. Specifically to ensure utilization of segment-anything model with thermal images we observe the results generated by direct application to contain inaccurate bounding details as demonstrated in Fig. 3. Hence to overcome this limitation we combine the segment anything model with domain adapted panoptic segmentation wherein we perform mask matching to accurately enclose objects of interest.

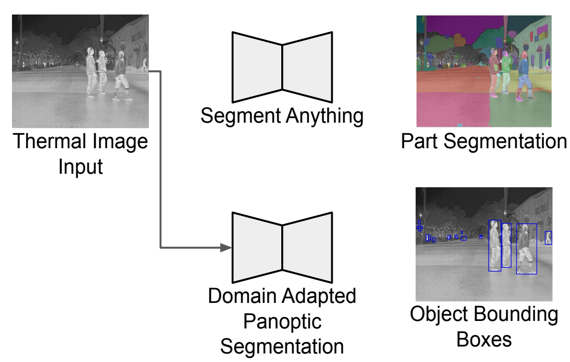


Figure 2. Overview of the mechanism to generate annotations for thermal images using segment-anything [5] model and EDAPS [6].

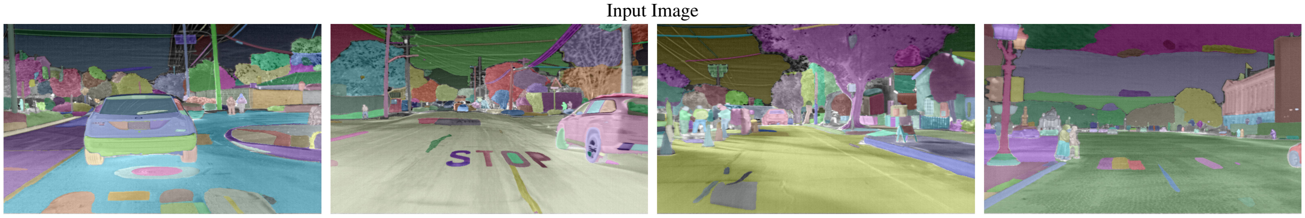
3. Appendix C

In this section, we evaluate the implications of changing the input modality between RGB and thermal images on pretrained object detectors trained only on RGB images or

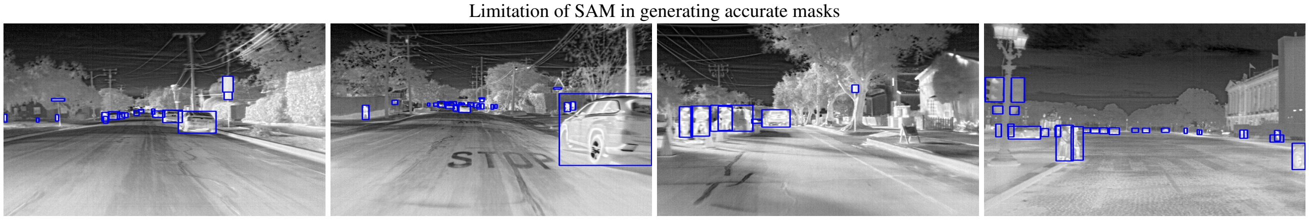
Table 1. Summary of attribute distribution for different datasets.

Dataset	Person	Bike	Car	MotorCycle	Bus	Train	Truck	Traffic Light	Fire Hydrant	Street Sign	Dog	Stakeboard	Stroller	Scotter
FLIR [1]	245451	258912	231549	4125	4120	3214	5021	15212	4126	4128	2510	6120	102	412
KAIST [3]	124582	341269	241582	12361	1528	5662	4857	12362	1025	3698	248	6324	4125	6528
DENSE [2]	154782	25698	41489	341589	41256	41856	9544	41256	516	4122	8456	5125	588	4258
Freiburg Thermal [7]	45812	4125	6523	54111	26684	5249	41690	41699	41	2582	4169	8954	4112	125

Input Image



Limitation of SAM in generating accurate masks



Object bounding boxes generated by proposed approach.

Figure 3. Qualitative evaluation of mechanism to generate bounding boxes for objects of interest using combination of SAM and Domain adapted Panoptic Segmentation.

thermal images. From the quantitative results in Tab. 2, we observe that a change in modality results in a significant performance drop caused by a change in the distribution of data capture. RGB images represent the amount of light the imaging sensor receives that is converted to a perceivable image. In contrast, thermal images are representative of the heat dissipated by different objects of interest. Hence, this change in capture of different characteristics of the scene is represented in the image distribution. While there exist prior works that focus on the fusion of images before being used as input to the object detector [4], we argue against this approach since RGB images contain more information than thermal images in well-lit conditions. Hence, we propose to use thermal images only in scenarios wherein RGB image is affected by weather degradation.

For our evaluation, to ensure the modality change is consistent with archtiecture of object detector we convert the RGB image to grayscale whereas for thermal image, we convert it to RGB by duplicating the values along the channel dimension.

Table 2. Quantitative performance of SoTA real time object detection algorithms on FLIR dataset.

Method	FLIR-RGB	FLIR-Thermal
YOLOX-tiny-Ours [RGB trained]	61.29	25.47
RTMDET-tiny-Ours [RGB trained]	70.91	26.81
YOLOX-tiny-Ours [Thermal trained]	29.98	60.59
RTMDET-tiny-Ours [Thermal trained]	30.41	69.84
YOLOX Proposed System	63.38	
RTMDET Proposed System	72.40	

Table 3. Quantitative performance of SoTA real time object detection algorithms on FLIR dataset.

Method	DENSE	KAIST	Freiberg Thermal
YOLOX-tiny-Ours	51.28	42.57	53.67
RTMDET-tiny-Ours	56.27	46.02	57.93

Figure 4. Qualitative performance of proposed algorithm on DENSE dataset [2] with object detection results super imposed on super resolved thermal image.

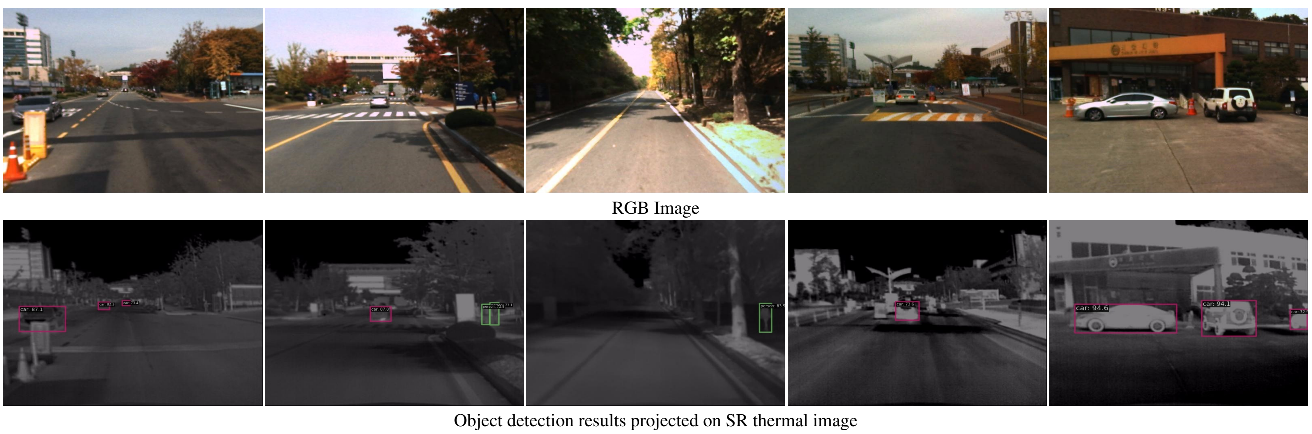


Figure 5. Qualitative performance of proposed algorithm on KAIST dataset [3] with object detection results super imposed on super resolved thermal image.

4. Appendix D

We now summarize the quantitative and qualitative results of integration of thermal super resolution on baseline object detection performance in Tab. 2 and Fig. 4, Fig. 5. We observe the integration of thermal super resolution alongside the RGB input results in improved performance of the underlying object detection algorithms. We also include quantitative results of proposed approach on DENSE, KAIST and Freiberg Thermal datasets in Tab. 3.

References

[1] Flir thermal dataset, 2021. https://www.flir.com/oem/adas/adas-dataset-form/ [Accessed: (Dec. 22, 2022)]. 1, 2

[2] Mario Bijelic, Tobias Gruber, Fahim Mannan, Florian Kraus, Werner Ritter, Klaus Dietmayer, and Felix Heide. Seeing through fog without seeing fog: Deep multimodal sensor fusion in unseen adverse weather. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11682–11692, 2020. 1, 2, 3

[3] Yukyung Choi, Namil Kim, Soonmin Hwang, Kibaek Park, Jae Shin Yoon, Kyounghwan An, and In So Kweon. Kaist multi-spectral day/night data set for autonomous and assisted driving. *IEEE Transactions on Intelligent Transportation Systems*, 19(3):934–948, 2018. 2, 3

[4] Shruthi Gowda, Bahram Zonooz, and Elahe Arani. Does thermal data make the detection systems more reliable? *arXiv preprint arXiv:2111.05191*, 2021. 2

[5] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C. Berg, Wan-Yen Lo, Piotr Dollár, and Ross Girshick. Segment anything. *arXiv:2304.02643*, 2023. 1

[6] Suman Saha, Lukas Hoyer, Anton Obukhov, Dengxin Dai, and Luc Van Gool. Edaps: Enhanced domain-adaptive panoptic segmentation. *arXiv preprint arXiv:2304.14291*, 2023. 1

[7] Johan Vertens, Jannik Zürn, and Wolfram Burgard. Heatnet: Bridging the day-night domain gap in semantic segmentation with thermal images. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 8461–8468. IEEE, 2020. 2