Drone Surveillance in Extreme Low Visibility Conditions

Prachi Agrawal*, Anant Verma* and Pratik Narang
Department of CSIS BITS Pilani, Rajasthan, India

Keywords: Drone Surveillance, Object Detection, Image Enhancement.

Abstract: Autonomous surveillance has several applications which include surveilling calamity prone areas, search and rescue operations, military operations and traffic management in smart cities. In low visibility conditions like low-light, haze, fog, snowfall, autonomous surveillance is a challenging task and current object detection models perform poorly in these conditions. Lack of datasets that capture challenging low visibility conditions is one of the reasons that limits the performance of currently available models. We propose a synthetic dataset for Human Action Recognition for search and rescue operations consisting of aerial images with different low visibility conditions including low light, haze, snowfall and also images with combinations of these low visibility conditions. We also propose a framework called ExtremeDetector for object detection in extreme low visibility conditions consisting of a degradation predictor and enhancement pool for enhancing a low visibility image and YOLOv5 for object detection in the enhanced image.

1 INTRODUCTION

UAV and drones have recently emerged as alternatives for surveillance in situations where human involvement is dangerous or unfeasible. One such application is in search and rescue (SAR) operations during disasters where drones can identify affected humans thus aiding in timely rescue. Deep learning models show good performance in object detection and classification tasks (He et al., 2016; Szegedy et al., 2015; Uijlings et al., 2013; Purkait et al., 2017; Girshick, 2015; Ren et al., 2015; Wang et al., 2016) thus motivating their use for autonomous surveillance. However, current deep learning models perform poorly in low visibility conditions such as low light, haze, fog, snowfall making them unfit for deployment in real-world scenarios. Furthermore, recent works on image enhancement and restoration (Zheng and Gupta, 2022; Singh et al., 2020; Cai et al., 2016; Zhang et al., 2021b; Dong et al., 2020; Li et al., 2021; Qin et al., 2020; Fu et al., 2017) are focused towards enhancement of images with a specific kind of degradation and don't take challenging low visibility conditions and neither their combinations into account. Lack of publicly available datasets that capture a variety of degrading factors and real-world challenging scenarios is one of the reasons for poor performance of deep learning models in challenging low visibility scenarios.

In this work, we explore the performance of current object detection models in challenging low visibility conditions and develop an end to end framework for object detection focusing on Human Action Recognition in extreme visibility conditions. Our major contributions are listed below.

- To the best of our knowledge, existing object detection datasets do not contain images with extreme low visibility conditions and combinations of them. Furthermore, lack of aerial datasets capturing such scenarios motivates us to generate a new dataset. Thus, we generate a new aerial images dataset for for Human Action Recognition consisting of five low visibility conditions which include low light, fog, snowfall, combination of low light and fog, combination of snowfall and fog.
- We evaluate performance of current object detection models on the generated low visibility dataset thus laying groundwork for future research.
- We propose a framework, ExtremeDetector shown in Figure 2 for object detection in extreme low visibility conditions including but not limited to the ones listed above.

2 RELATED WORK

Object detection is a task involving localization and classification. Current object detection methods can be classified broadly into two categories - single stage

*These authors contributed equally to this work

Agrawal, P., Verma, A. and Narang, P.
Drone Surveillance in Extreme Low Visibility Conditions.
DOI: 10.5220/0011727600003393
In *Proceedings of the 15th International Conference on Agents and Artificial Intelligence (ICAART 2023) - Volume 3*, pages 658-664
ISBN: 978-989-758-623-1; ISSN: 2184-433X

Figure 1: Sample images from the dataset with original captured images (column 1), foggy images (column 2), low light images (column 3), low light+fog (column 4), snowfall+fog (column 5), snowfall (column 6).

Table 1: Number of images in our proposed dataset per action class in train and test set.

Action	Train	Test
Person Standing	11778	1452
Person Sitting	2988	510
Person Handshaking	468	42
Person Running	2490	234
Person Waving	5562	534
Person Lying	1164	132
Person Walking	6900	1002

detectors and two-stage object detectors. Single stage detectors (Lin et al., 2017; Jocher et al., 2022; Liu et al., 2016) use a single CNN to predict object labels and bounding box coordinates whereas two-stage detectors (Uijlings et al., 2013; Purkait et al., 2017; Girshick, 2015; Ren et al., 2015; Wang et al., 2016), extract regions of interest (RoIs), then classify the RoIs. Object detection in low visibility conditions is an insufficiently researched area. Previous works (Zhang et al., 2021a), (Shan et al., 2019), (Chen et al., 2018b), (Hnewa and Radha, 2021) address this problem by viewing object detection in hazy and rainy scenes as a domain adaptation task. Sindagi et al. (Sindagi et al., 2020) proposed to reduce weather specific features using a prior-adversarial loss that uses additional knowledge about the target domain (hazy and rainy images) for aligning the source and target domain features. Liu et al. (Liu et al., 2022) use a CNN to learn parameters of a differentiable image processing module which takes into account the adverse weather conditions for a YOLOv3 detector. Some previous works (Huang et al., 2021) jointly perform image enhancement and object detection. Most datasets used for image enhancement mainly target at evaluating the difference of enhanced images w.r.t ground truth images using quantitative metrics like PSNR, SSIM or qualitatively. Recent datasets include RawInDark (Chen et al., 2018a), LOL dataset (Wei et al., 2018) for low light enhancement, HazeRD (Zhang et al., 2017b), IHAZE (Ancuti et al., 2018), OHAZE (Ancuti et al., 2018) for dehazing, DIV2K (Timofte et al., 2017), MANGA (Fujimoto et al., 2016) for image super-resolution, Rain 100L/H (Yang et al., 2017), Rain800 (Zhang et al., 2019) for rainfall removal. Challenges in aerial datasets include small objects, objects in different sizes and with different orientations. Datasets collected by aerial vehicles include VIRAT Video Dataset (Oh et al., 2011), UAV123 (Mueller et al., 2016), and a multipurpose dataset (Yao et al., 2007). However, these datasets are not captured in adverse conditions. Commonly used datasets for object detection in adverse conditions include Foggy-Cityscapes (Sakaridis et al., 2018), RTTS (Li et al., 2018) for foggy conditions and ExDARK (Loh and Chan, 2019), UFDD (Nada et al., 2018) for low light conditions.

3 PROPOSED WORK

3.1 Dataset Creation

We use an aerial Human Action Recognition dataset (Mishra et al., 2020) to generate realistic synthetic datasets for 5 low visibility conditions which are - (1) Low light , (2) Fog , (3) Snowfall , (4) Low light + Fog , (5) Fog + Snowfall. The dataset in (Mishra et al., 2020) consists of images of 7 human actions captured from a drone equipped with a high definition camera from the height between 10 m to 40 m. The 7 human actions captured are - Person Standing, Person Sitting, Person Handshaking, Person Running, Person Waving, Person Lying and Person Walking. It consists of a total of 3050 images (split into train and test each having 2560 and 490 images respectively). Our proposed dataset with low visibility conditions has a total of 15360 training images and 2940 test images. The distribution of images for each human action class is presented in Table 1. Some images from the dataset

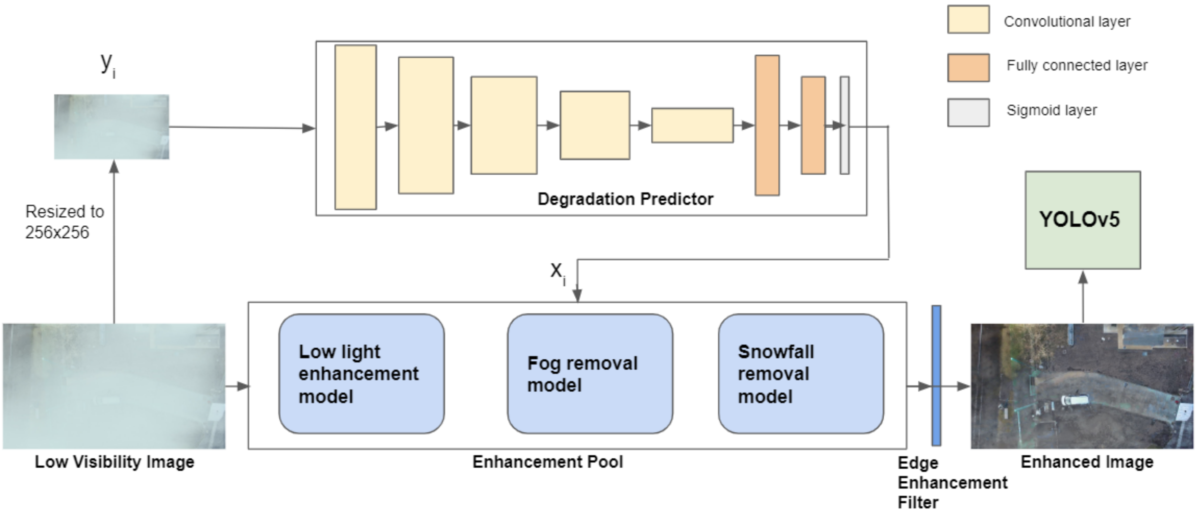

Figure 2: ExtremeDetector. Proposed framework for object detection in low visibility images.

Table 2: mAP@IoU=0.5 of object detection models on proposed dataset. The best results have been highlighted in bold text.

Method	Normal	Fog	Low Light	Low Light+Fog	Fog+Snowfall	Snowfall
RetinaNet(Lin et al., 2017)	0.00027	0.00030	0.00034	0.0004	0.0002	0.00027
Detectron2(Wu et al., 2019)	0.282	0.182	0.293	0.177	0.246	0.252
YOLOv5(Jocher et al., 2022)	0.319	0.1899	0.29	0.25	0.31	0.275
IA-YOLO(Liu et al., 2022)	0.312	0.2984	0.334	0.334	0.339	0.3554
Ours	**0.408**	**0.391**	**0.386**	**0.3577**	**0.3546**	**0.405**

are shown in Figure 1. The process of generating synthetic dataset for each low visibility condition is described below.

- **Low Light Effect:** In order to generate a low-lit image while preserving the underlying information, we follow the pipeline used in (Cui et al., 2021). We generate images with varying amount of darkness.
- **Fog effect:** We use Foggy and Hazy Images Simulator (FoHIS) (Zhang et al., 2017a), a framework based on an atmospheric scattering model which can simulate both fog and haze effects at any elevation in an image. Our dataset consists of images with variation in amount of fog.
- **Snowfall effect:** Image editing tools like Adobe Photoshop were used to add snowfall effect through random layered masks supporting varying amounts of snow and falling angle in each image. Further, blue channels of each image were enhanced in order to simulate a more realistic winter effect.
- **Low light + Fog:** We generated images with fog followed by addition of low-light effect using the approaches discussed above.
- **Snow + Fog:** Fog effect was added to images with snowfall using the above approaches.

3.2 Framework

The proposed framework shown in Figure 2 consists of a degradation predictor module that identifies the degradation in the image and accordingly passes the image into selected models from a pool of pre-trained models specialized to remove a specific degradation. A Laplacian filter for edge enhancement (available in Python Imaging Library) is then applied on the image to improve action recognition. A YOLOv5 (Ultralytics 2020) (Jocher et al., 2022) detector then uses the enhanced image to detect and identify the action. Since our dataset consists of images with extreme

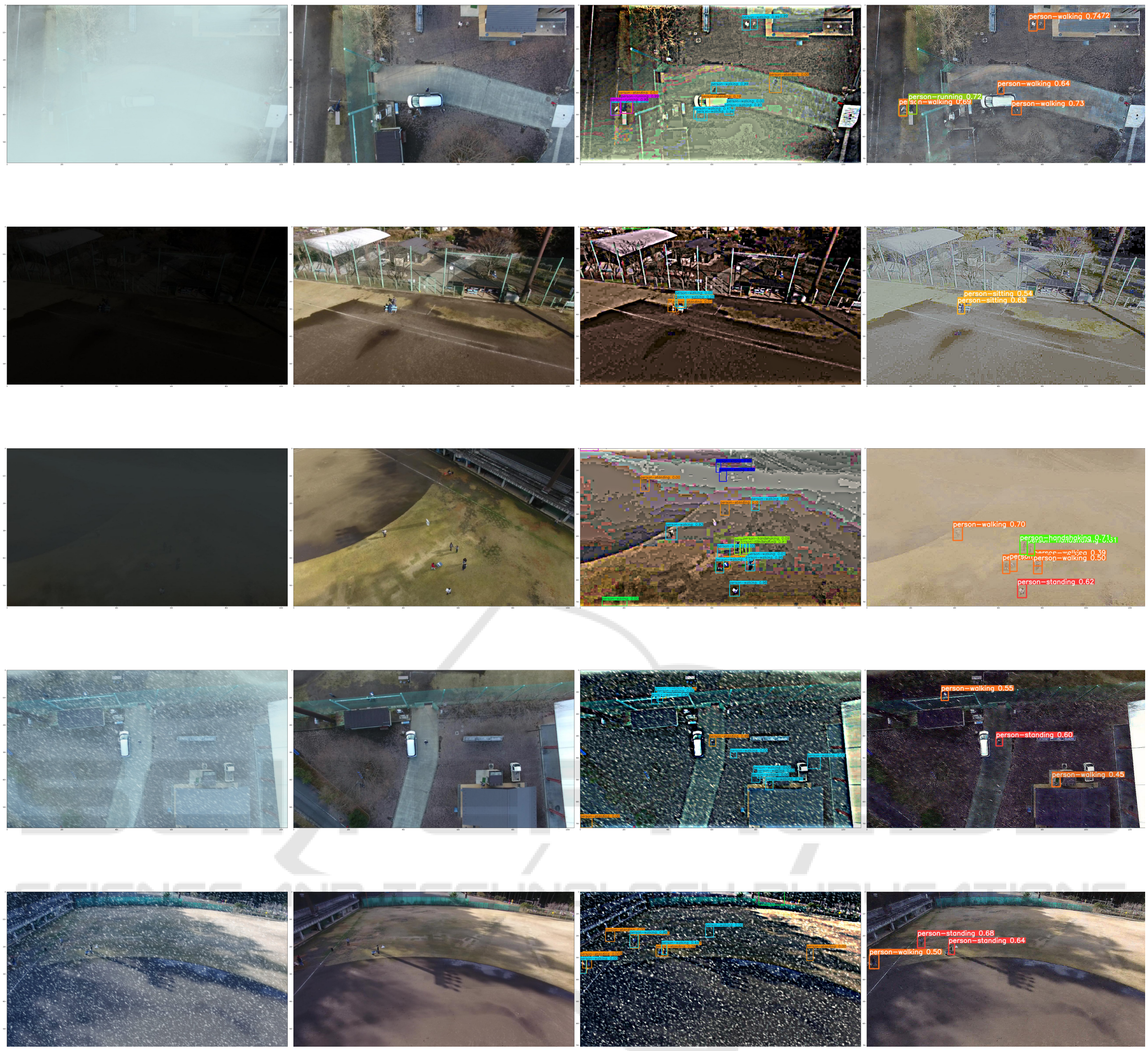

Figure 3: Detection results of IA-YOLO (Liu et al., 2022) (column 3) and our framework (column 4) on augmented low visibility images (column 1) along with their corresponding original images (column 2). Each row correspond to the 5 low visibility conditions - fog (row 1), low light (row 2), low light+fog (row 3), snowfall+fog (row 4), snowfall (row 5).

darkness, fog, snowfall and combination of these conditions, we leverage SOTA models specialized to enhance specific degradation.

3.2.1 Enhancement Pool

Our enhancement model pool consists of Zero-DCE (Li et al., 2021) for low light image enhancement, FFA-Net (Qin et al., 2020) for removing fog and Deep Detailed Network (Fu et al., 2017) for removing snow. Pre-trained Zero-DCE model was used to enhance low light images. FFA-Net and Deep Detailed Network models were fine-tuned using a subset of foggy images and snowfall images respectively and their corresponding clean image pair.

3.2.2 Edge Enhancement Filter

Edge enhancement filter (Laplacian filter) increases the contrast of the pixels around edges thereby making them prominent, aiding the use of object and pattern recognition. The kernel shown below (denoted by *EE*) is applied on image output from the enhancement pool.

$$EE = \begin{matrix} -1 & -1 & -1 \\ -1 & 10 & -1 \\ -1 & -1 & -1 \end{matrix}$$

3.2.3 Degradation Predictor Module

We use a degradation predictor to identify degradation in input image and enhance them accordingly for better feature extraction during object detection. The image is resized to 256x256 before passing to the degradation predictor module. The degradation predictor module is composed of five convolutional blocks, two fully-connected layers followed by a sigmoid layer. Each convolutional block consists of a 3×3 convolutional layer with stride 2 and a leaky ReLU activation. The module outputs 3 values each in the range 0 to 1 which correspond to the probability for the condition that would be present in the input image. We use a threshold of 0.5 to determine the target enhancement model(s), if any, from the enhancement pool. To train the module, we use supervised learning based on manually determined degradation classes of the images. Binary cross entropy loss (denoted by L_{BCE} in Equation 1) is used where the target output value is set to 1 if corresponding degradation is present in image, otherwise set to 0.

$$L_{BCE} = \frac{1}{N}\sum_{i=1}^{N} -(y_i log(x_i) + (1-y_i) log(1-x_i)) \quad (1)$$

In the above equation, N is number of models in enhancement pool. We use 3 models in the enhancement pool. x_i refers to output value of degradation predictor module and y_i refers to target value corresponding to the degradation type label of image.

3.2.4 Object Detection

We choose YOLOv5 (Ultralytics 2020) (Jocher et al., 2022) for object detection because it is suitable for deployment due to fast inference and train it using the enhanced images. The YOLOv5 architecture consists of three parts - (1) CSPDarknet backbone, (2) PANet neck, and (3) YOLO Layer. The enhanced images are first input to CSPDarknet for feature extraction, and then fed to PANet for feature fusion. Finally, YOLO Layer outputs detection results (class, score, location, size).

4 EXPERIMENTAL RESULTS

We trained the object detection models on a hybrid set of images from the five low visibility conditions along with the original ambience. The degradation predictor was trained using Adam optimizer with weight decay 1e-5 and learning rate 1e-4. We use PyTorch for our experiments. We evaluate the performance of these well established object detection models- RetinaNet (Lin et al., 2017), YOLOv5 (Jocher et al., 2022), Detectron2 (Faster RCNN X101-FPN) (Wu et al., 2019) and IA-YOLO (Liu et al., 2022) on our proposed dataset. The evaluation metric used is mean average precision (mAP) at Intersection over Union (IoU) threshold of 0.5. If the ratio of the intersection of a detected region with an annotated object is greater than 0.5, a score of 1 is assigned to the detected region, otherwise 0 is assigned.

4.1 Discussion

Results in Table 2 show that our proposed framework shows a significant increase in mAP over current object detection models which include RetinaNet, Detectron2 and YOLOV5 in all conditions. This indicates that degradation in images hinders extraction of relevant features. Thus, enhancement/restoration of degraded images is essential before detection in extremely low visibility conditions. Our framework also has a better mAP than IA-YOLO (Liu et al., 2022) in all conditions. Further, Figure 3 shows that IA-YOLO has very poor image enhancement especially in fog and low light+fog images. The visual results indicate that the differentiable filters proposed in (Liu et al., 2022) are insufficient for enhancement of images with extreme degradation thus leading to poor extraction of relevant features for object detection. Our framework shows better image enhancement resulting in better object detection. Additionally, our approach has detection results for multiple degradation conditions (combination of fog & snowfall, combination of low light and fog) at par with single type of degradation without additional enhancement models for these conditions thus making our framework robust to multiple degradation. We study the impact of using edge enhancement filter (EE filter) and report the results in Table 3. The results indicate that applying edge enhancement filter on the images output from the enhancement pool has a significant improvement in detection results in all conditions. The mAP values of the methods shown in Table 2 are below 0.45 in all conditions which indicates that our proposed dataset is challenging and there is room for further improvement in detection models to be fit for deployment in challenging scenarios.

Our work is a step in the direction of exploring the challenges of several types of degradation, hostile weather conditions with varying intensities in object detection. Lack of real-world datasets capturing these conditions, varying heights of captured objects and their sizes in aerial images add to the challenges. Going forward we aim to distillate the specialised image enhancement/restoration models into one and

Table 3: Ablation Experiment. Evaluation of mAP@0.5 of our framework with and without edge enhancement filter. Best values are in bold text.

Method	Normal	Fog	Low Light	Low Light+Fog	Fog+ Snowfall	Snowfall
Ours w/o EE filter	0.384	0.238	0.3	0.32	0.341	0.368
Ours	**0.408**	**0.391**	**0.386**	**0.3577**	**0.3546**	**0.405**

get a lighter framework for the task. In addition to this, with the increasing popularity and applications of Transformers in Computer Vision research, we also intend to explore the possibilities of equipping Transformers in such extreme low-visibility conditions for Object Detection.

5 CONCLUSION

In this work, we have proposed a new dataset with challenging low visibility conditions. We also proposed a framework for object detection that is robust to different low visibility conditions (low light, fog, snowfall and their combinations). We perform benchmarking experiments on our generated dataset and surpass the detection results of some of the well-researched object detection architectures. However, the computational overhead of specialized deep learning models for each degradation limits the scalability of our framework. Our work motivates further research in developing a single lightweight model for object detection in images captured in such extreme low visibility conditions with performance at par with favourable visibility conditions.

REFERENCES

Ancuti, C., Ancuti, C. O., Timofte, R., and Vleeschouwer, C. D. (2018). I-haze: a dehazing benchmark with real hazy and haze-free indoor images. In *International Conference on Advanced Concepts for Intelligent Vision Systems*, pages 620–631. Springer.

Cai, B., Xu, X., Jia, K., Qing, C., and Tao, D. (2016). Dehazenet: An end-to-end system for single image haze removal. *IEEE Transactions on Image Processing*, 25(11):5187–5198.

Chen, C., Chen, Q., Xu, J., and Koltun, V. (2018a). Learning to see in the dark. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3291–3300.

Chen, Y., Li, W., Sakaridis, C., Dai, D., and Van Gool, L. (2018b). Domain adaptive faster r-cnn for object detection in the wild. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3339–3348.

Cui, Z., Qi, G.-J., Gu, L., You, S., Zhang, Z., and Harada, T. (2021). Multitask aet with orthogonal tangent regularity for dark object detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 2553–2562.

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*, pages 2157–2167.

Fu, X., Huang, J., Zeng, D., Huang, Y., Ding, X., and Paisley, J. (2017). Removing rain from single images via a deep detail network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3855–3863.

Fujimoto, A., Ogawa, T., Yamamoto, K., Matsui, Y., Yamasaki, T., and Aizawa, K. (2016). Manga109 dataset and creation of metadata. In *Proceedings of the 1st international workshop on comics analysis, processing and understanding*, pages 1–5.

Girshick, R. (2015). Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 1440–1448.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. pages 770–778.

Hnewa, M. and Radha, H. (2021). Multiscale domain adaptive yolo for cross-domain object detection. In *2021 IEEE International Conference on Image Processing (ICIP)*, pages 3323–3327. IEEE.

Huang, S.-C., Le, T.-H., and Jaw, D.-W. (2021). Dsnet: Joint semantic learning for object detection in inclement weather conditions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43:2623–2633.

Jocher, G., Chaurasia, A., Stoken, A., Borovec, J., NanoCode012, Kwon, Y., TaoXie, Fang, J., imyhxy, Michael, K., Lorna, V, A., Montes, D., Nadar, J., Laughing, tkianai, yxNONG, Skalski, P., Wang, Z., Hogan, A., Fati, C., Mammana, L., AlexWang1900, Patel, D., Yiwei, D., You, F., Hajek, J., Diaconu, L., and Minh, M. T. (2022). ultralytics/yolov5: v6.1 - TensorRT, TensorFlow Edge TPU and OpenVINO Export and Inference.

Li, B., Ren, W., Fu, D., Tao, D., Feng, D., Zeng, W., and Wang, Z. (2018). Benchmarking single-image dehazing and beyond. *IEEE Transactions on Image Processing*, 28(1):492–505.

Li, C., Guo, C., and Chen, C. L. (2021). Learning to enhance low-light image via zero-reference deep curve estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P. (2017). Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pages 2980–2988.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., and Berg, A. C. (2016). Ssd: Single shot multibox detector. In *European conference on computer vision*, pages 21–37. Springer.

Liu, W., Ren, G., Yu, R., Guo, S., Zhu, J., and Zhang, L. (2022). Image-adaptive yolo for object detection in adverse weather conditions. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 1792–1800.

Loh, Y. P. and Chan, C. S. (2019). Getting to know low-light images with the exclusively dark dataset. *Computer Vision and Image Understanding*, 178:30–42.

Mishra, B., Garg, D., Narang, P., and Mishra, V. (2020). Drone-surveillance for search and rescue in natural disaster. *Computer Communications*, 156:1–10.

Mueller, M., Smith, N., and Ghanem, B. (2016). A benchmark and simulator for uav tracking. In *European conference on computer vision*, pages 445–461. Springer.

Nada, H., Sindagi, V. A., Zhang, H., and Patel, V. M. (2018). Pushing the limits of unconstrained face detection: a challenge dataset and baseline results. In *2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pages 1–10. IEEE.

Oh, S., Hoogs, A., Perera, A., Cuntoor, N., Chen, C.-C., Lee, J. T., Mukherjee, S., Aggarwal, J., Lee, H., Davis, L., et al. (2011). A large-scale benchmark dataset for event recognition in surveillance video. In *CVPR 2011*, pages 3153–3160. IEEE.

Purkait, P., Zhao, C., and Zach, C. (2017). Spp-net: Deep absolute pose regression with synthetic views. *arXiv preprint arXiv:1712.03452*.

Qin, X., Wang, Z., Bai, Y., Xie, X., and Jia, H. (2020). Ffa-net: Feature fusion attention network for single image dehazing. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 11908–11915.

Ren, S., He, K., Girshick, R., and Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28.

Sakaridis, C., Dai, D., and Van Gool, L. (2018). Semantic foggy scene understanding with synthetic data. *International Journal of Computer Vision*, 126(9):973–992.

Shan, Y., Lu, W. F., and Chew, C. M. (2019). Pixel and feature level based domain adaptation for object detection in autonomous driving. *Neurocomputing*, 367:31–38.

Sindagi, V. A., Oza, P., Yasarla, R., and Patel, V. M. (2020). Prior-based domain adaptive object detection for hazy and rainy conditions. In *European Conference on Computer Vision*, pages 763–780. Springer.

Singh, A., Bhave, A., and Prasad, D. K. (2020). Single image dehazing for a variety of haze scenarios using back projected pyramid network. In *European Conference on Computer Vision*, pages 166–181. Springer.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2015). Rethinking the inception architecture for computer vision. *CoRR*, abs/1512.00567.

Timofte, R., Agustsson, E., Van Gool, L., Yang, M.-H., and Zhang, L. (2017). Ntire 2017 challenge on single image super-resolution: Methods and results. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 114–125.

Uijlings, J. R., Van De Sande, K. E., Gevers, T., and Smeulders, A. W. (2013). Selective search for object recognition. *International journal of computer vision*, 104(2):154–171.

Wang, L., Wang, L., Lu, H., Zhang, P., and Ruan, X. (2016). Saliency detection with recurrent fully convolutional networks. In *European conference on computer vision*, pages 825–841. Springer.

Wei, C., Wang, W., Yang, W., and Liu, J. (2018). Deep retinex decomposition for low-light enhancement. *arXiv preprint arXiv:1808.04560*.

Wu, Y., Kirillov, A., Massa, F., Lo, W.-Y., and Girshick, R. (2019). Detectron2. https://github.com/facebookresearch/detectron2.

Yang, W., Tan, R. T., Feng, J., Liu, J., Guo, Z., and Yan, S. (2017). Deep joint rain detection and removal from a single image. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1357–1366.

Yao, B., Yang, X., and Zhu, S.-C. (2007). Introduction to a large-scale general purpose ground truth database: methodology, annotation tool and benchmarks. In *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, pages 169–183. Springer.

Zhang, H., Sindagi, V., and Patel, V. M. (2019). Image de-raining using a conditional generative adversarial network. *IEEE transactions on circuits and systems for video technology*, 30(11):3943–3956.

Zhang, N., Zhang, L., and Cheng, Z. (2017a). Towards simulating foggy and hazy images and evaluating their authenticity. In *International Conference on Neural Information Processing*, pages 405–415. Springer.

Zhang, S., Tuo, H., Hu, J., and Jing, Z. (2021a). Domain adaptive yolo for one-stage cross-domain detection. In *Asian Conference on Machine Learning*, pages 785–797. PMLR.

Zhang, Y., Ding, L., and Sharma, G. (2017b). Hazerd: an outdoor scene dataset and benchmark for single image dehazing. In *2017 IEEE international conference on image processing (ICIP)*, pages 3205–3209. IEEE.

Zhang, Z., Jiang, Y., Jiang, J., Wang, X., Luo, P., and Gu, J. (2021b). Star: A structure-aware lightweight transformer for real-time image enhancement. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4106–4115.

Zheng, S. and Gupta, G. (2022). Semantic-guided zero-shot learning for low-light image/video enhancement. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 581–590.