Supplementary Material

Our supplementary materials give more details about our DG-Net and more experiment results, which can be summarized as follows:

- We provide more experiment details and parameter sets.
- We provide more visual results on the RTTS dataset to demonstrate the superior performance of the proposed DG-Net.
- We provide more visible results of the ablation study to prove the effectiveness of DGFF.
- We provide more details of the mixed dataset in union training strategy.

1 Experiment details and parameter settings

Table 1 illustrates the experiment details, and Table 2 shows platform configuration information. To improve the utilization of image features further, we add the mosaic and mix-up data augmentation strategies in the training process as YOLOX defaults.

Table 1: Details of experimental parameter settings.

Parameters	Descriptions	settings
mosaic_prob	possibility of mosaic	0.5
mixup_prob	possibility of mosaic	0.5
special_aug_ratio	rate of augmentation epochs	0.7
unfreeze_batch_size	batch size of unfreeze epoch	16
freeze_batch_size	batch size of freeze epoch	16
momentum	/	0.937
num_workers	number of workers	4
ema	exponential moving average	true

2 More detection visual results

Fig. 1 illustrates more visual comparisons with several state-of-the-art methods on the RTTS dataset. As we can see, the proposed DG-Net achieves satisfactory performance and maintains robustness in scenes with heavy and heterogeneous haze.

3 More visible results of the ablation study

We analyze the effectiveness of the DGFF module. Fig. 2 shows that the DGFF module can help the network focus more on potential object regions and detect more instances.

Table 2: Experimental platform configuration information

Properties	Configurations
Operating System	Ubuntu 20.04
CPU	Intel Xeon 6330 * 2
RAM	32GB DDR4 * 16
GPU	Tesla A100 GPU 40GB * 4
PyTorch	1.12
CUDA	11.3
cudnn	8.3.2
Python	3.8
matplotlib	3.6
numpy	1.23
pillow	9.2.0
opencv	4.6.0
scipy	1.9.3

4 More details of the mixed dataset in union training strategy

In fact, the mixed dataset does not artificially introduce additional data. We kept the same number of images in the dataset and randomly selected 10% clear images to replace 10% foggy images one-to-one, so there are no pairs of clear and foggy images, i.e., the dataset contains 90% synthetic foggy images and 10% clear images. Our method can also use those clear images of other scenes as long as they contain these detection categories. Thus, our method is more friendly to the requirements for datasets.


(a) YOLOXs (b) DCP-YOLOXs* (c) TogetherNet (d) DG-Net(Ours)

Fig. 1: More visual comparisons on the RTTS.(Zoom in to see the details)

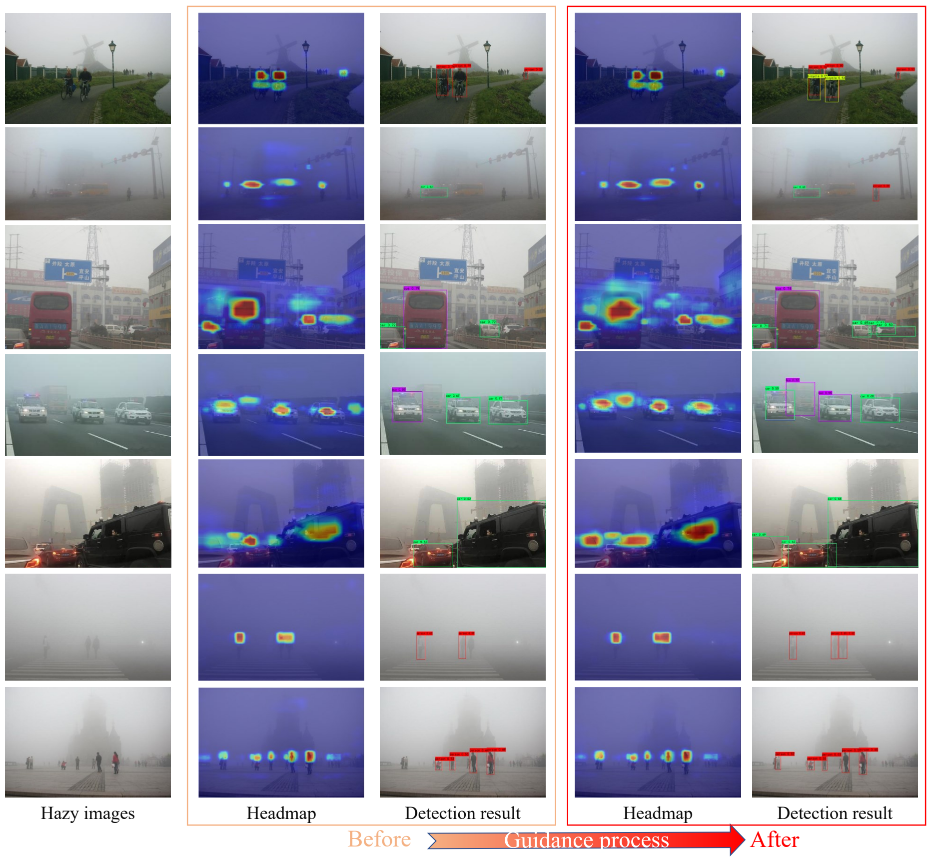

Fig. 2: More ablation study results on the DGFF module.(Zoom in to see the details)