Robust pedestrian detection in thermal imagery using synthesized images

My Kieu, Lorenzo Berlincioni, Leonardo Galteri, Marco Bertini, Andrew D. Bagdanov, Alberto Del Bimbo

{kieu, berlincioni, galteri, bertini, bagadanov, delbimbo}@unifi.it

Introduction

- Pedestrian detection is central problem in a variety of applications, such as video surveillance and autonomous driving.
- Due to the cost of deploying multiple aligned sensors, and the lack of annotated thermal imagery available Thermal-only detectors typically yield lower performance than multispectral detectors
- We propose a novel generative model based on the Least-Squares Generative Adversarial Network (LSGAN) that is able to synthesize thermal imagery from RGB, we then use it in order to adapt a pretrained YOLOv3 pedestrian detector to work in the thermal-only domain
- Experimental results demonstrate the effectiveness of our approach: using less than 50% of available real thermal training data, and relying on synthesized data generated by our model in the domain adaptation phase, our detector achieves state-of-the-art results on the KAIST Multispectral Pedestrian Detection Benchmark

Model

The two main components of our proposed approach are a thermal pedestrian detector based on a fine-tuned to the thermal domain YOLOv3 and a generative model which produces fake thermal images from available RGB images

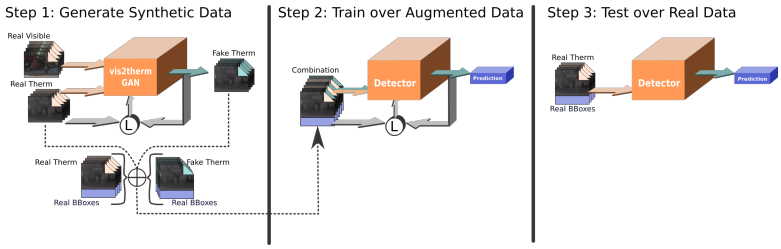

Figure 1: System overview

Mixtures

Real%	Synthetic%	MRall%
0	100	45.88
50	50	33.90
80	20	**25.88**
90	10	25.62
100	0	28.46
100	100	34.29

Table 1: Ablation study on varying quantities of GAN- generated images. Results are on KAIST in terms of log- average miss rate (lower is better). Best results highlighted in underlined bold, second best in bold.

Results

Using the domain adaptation method from [1] and proposed mixed training procedure, we achieve state-of-the- art performance for both all (day and night) and nighttime

Detectors	MR all	MR day	MR night
KAIST baseline	64.76	64.17	63.99
FasterRCNN	47.59	50.13	40.93
R^3-Net Saliency + KAIST	-	**30.40**	21.00
ResNet101-two-stage	42.65	49.59	26.70
Ours Mixed 40 60	34.78	43.45	14.53
Ours Mixed 80 20	25.88	33.01	**11.12**
Ours Mixed 90 10	**25.62**	31.86	12.92

Table 2: Comparison with state-of-the-art single-modality approaches on KAIST Thermal in term of log-average miss rate (lower is better).

Conclusions

Experiments show that even using only 50% of available real thermal images it is possible to obtain state-of-the-art comparable results This suggests that generated images may help to adapt visible spectrum detectors to operate in thermal spectrum.

References

[1] My Kieu, Andrew D Bagdanov, Marco Bertini, and Alberto Del Bimbo. Domain adaptation for privacy-preserving pedestrian detection in thermal imagery. In *International Conference on Image Analysis and Processing*, pages 203–213. Springer, 2019.

MICC
Media Integration and Communication Center

UNIVERSITÀ DEGLI STUDI FIRENZE
DINFO
DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE