Heng **ZHANG**

Deep learning R&D Engineer

22 Square Louis Armand, 35000, Rennes, France

(+33) 6-4647-0816 | heng.zhang@irisa.fr | zhangheng19931123.github.io | zhangheng19931123 | heng-zhang-628177112

"Maximum effort!"

About Me

I am a second year PhD student in LACODAM team at INRIA Rennes laboratory. I am working under the supervision of Prof. Elisa FROMONT and Prof. Sébastien LEFEVRE. In the same time, I work as a Deep learning R&D Engineer at ATERMES in Paris. My current research interest is deep learning for multispectral object detection, small object detection and video object detection.

Work Experience

Deep learning R&D Engineer
Paris area, France

ATERMES — *Dec. 2018 - now*

- Built Deep Learning models for accurate object detection (car, pedestrian, bicycle, etc) at long distance (>3km).
- Adaptively fusing information from multiple sensors (e.g., thermal camera & visible camera) to improve the detection precision.
- Accelerate CNN model inference for efficient deep learning applications on embedded systems.

Apprenticeship
Saint-étienne, France

Hubert Curien laboratory — *Sep. 2017 - Oct. 2018*

- Implement different deep learning models (Fater RCNN, SSD, YOLO, RetinaNet) for face/person detection in public transport (bus, tramway, subway, etc).
- Proposed efficient video object detection methods for video surveillance applications (6.9% of absolute precision improvement compared to perframe detection).

Education

PhD student in Deep Learning and Computer Vision
Rennes, France

INRIA Rennes — *Dec. 2018 - now*

- Subject: Multispectral Object detection for all-day video surveillance.
- Industry-oriented PhD program, cooperation with ATERMES company.

Engineer's degree in Computer Vision and Image Processing
Saint-étienne, France

Télécom Saint-étienne — *Sep. 2015 - Oct. 2018*

- One-year Apprenticeship in Hubert Curien laboratory.

Bachelor's degree in Computer Science and Electronic Engineering
Xi'an, China

Xidian University — *Sep. 2012 - Jun. 2016*

Publications

Guided Attentive Feature Fusion for Multispectral Pedestrian Detection
Hawaii, United States

2021 Winter Conference on Applications of Computer Vision (WACV 2021)) — *Jan. 2021*

- Multispectral image pairs can provide complementary visual information, making pedestrian detection systems more robust and reliable. To benefit from both RGB and thermal IR modalities, we introduce a novel attentive multispectral feature fusion approach. Under the guidance of the inter- and intra-modality attention modules, our deep learning architecture learns to dynamically weigh and fuse the multispectral features. Experiments on two public multispectral object detection datasets demonstrate that the proposed approach significantly improves the detection accuracy at a low computation cost.

Localize to Classify and Classify to Localize: Mutual Guidance in Object Detection [pdf]
Kyoto, Japan

15th Asian Conference on Computer Vision (ACCV2020) — *Dec. 2020*

- We propose a new anchor matching criterion guided, during the training phase, by the optimization of both the localization and the classification tasks: the predictions related to one task are used to dynamically assign sample anchors and improve the model on the other task, and vice versa. This strategy significantly improves the object detection accuracy of different deep learning architectures on PASCAL VOC and MS COCO dataset without any additional parameter or computation.

Multispectral Fusion for object detection with Cyclic Fuse-and-Refine blocks [pdf] *Abu Dhabi, United Arab Emirates*

27th IEEE International Conference on Image Processing (ICIP2020) *Oct. 2020*

- We propose a new feature fusion method for neural networks that leverages the complementary/consistency balance existing in multispectral features by adding to the network architecture, a particular module that cyclically fuses and refines each spectral feature. We obtain state-of-the-art results on KAIST Multispectral pedestrian detection benchmark.

Improving video object detection by Seq-Bbox Matching [pdf] *Prague, Czech Republic*

14th International Conference on Computer Vision Theory and Applications (VISAPP2019) *Feb. 2019*

- We propose a novel and highly effective box-level post-processing method to improve the accuracy of video object detection. The proposed method can be applied to online/offline detection. It achieves state-of-the-art performance on ImageNet VID dataset.