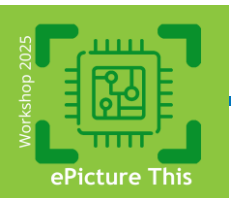


Multiview sensor fusion for improved traffic safety

Ljubomir Jovanov
imec-Ghent University

Eindhoven, the Netherlands

28 October 2025



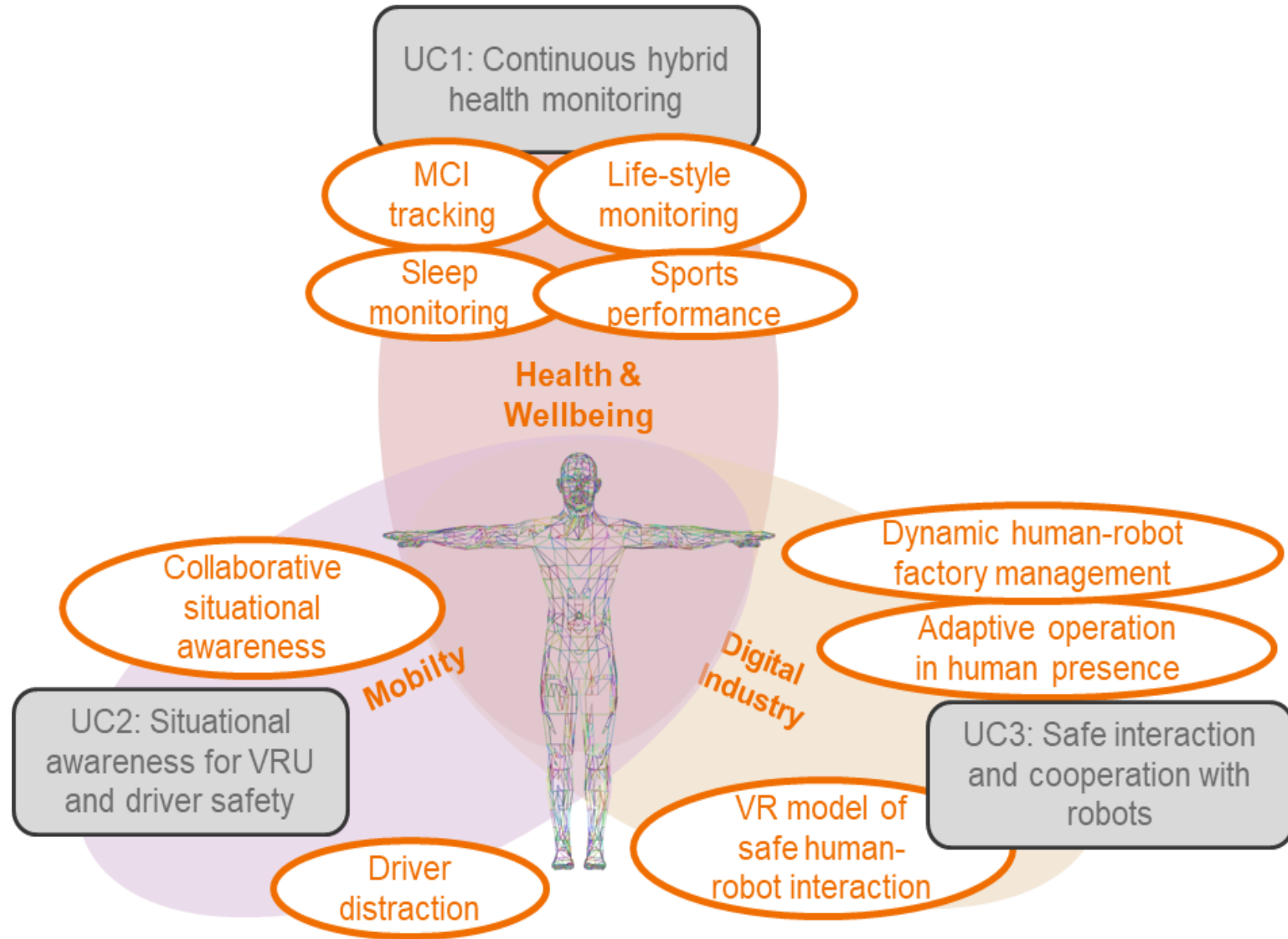
Organized by Penta projects:
2020005 Mantis Vision
2021004 Imagination



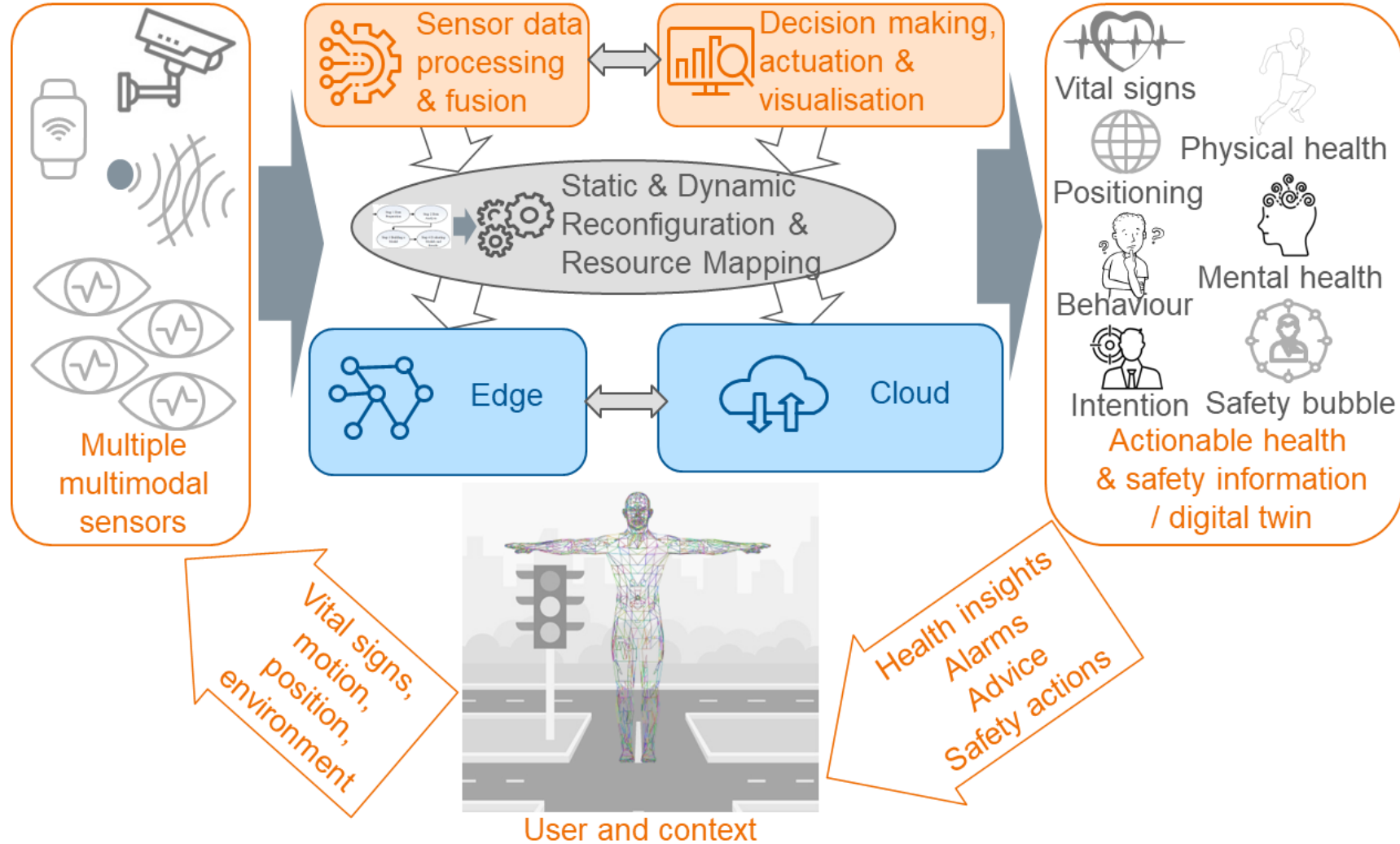
Agenda

- Introduction
- Thermal video fusion for VRU Detection
- Visible-thermal-radar fusion-based VRU detection
- Multi-node fusion
- Conclusions

USE CASES



CONCEPTUAL ARCHITECTURE



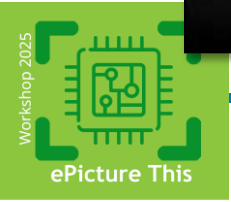
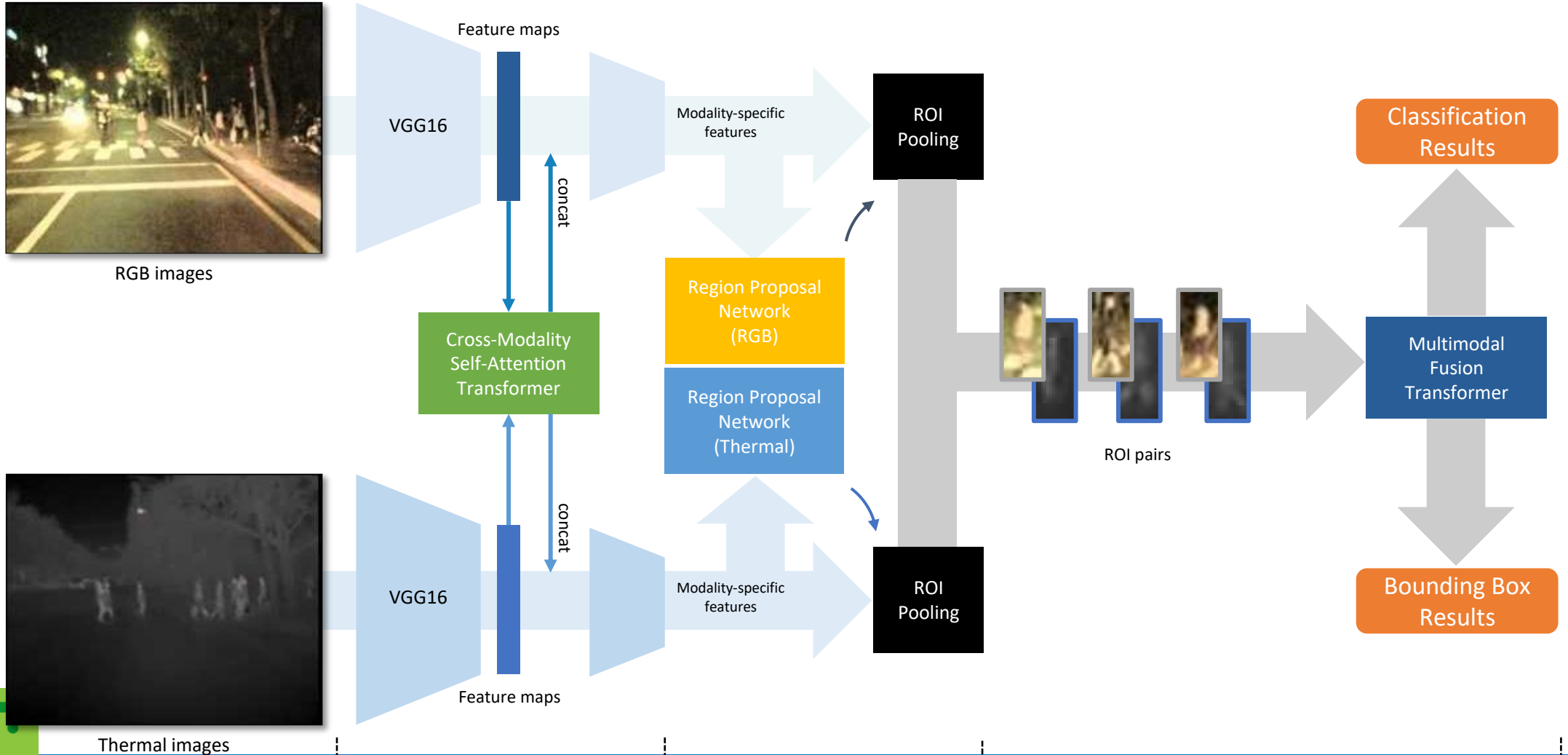
UC3: Safety and comfort at intersections

- Provide safety and comfort for all road users - including Vulnerable Road Users (VRUs) like pedestrians and bikers - at a road intersection. The use case will demonstrate the capability to detect the presence of traffic participants, determine their positions and track their motion and intent with high reliability. Final estimates of VRU trajectories are derived from multiple multi-sensor boxes



Thermal video fusion for VRU Detection

ARCHITECTURE



Two-stream feature extractor

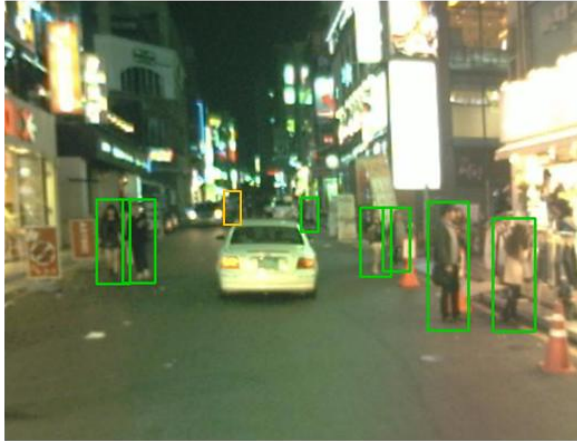
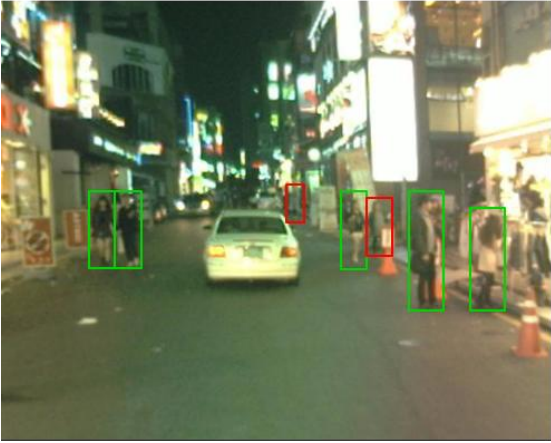
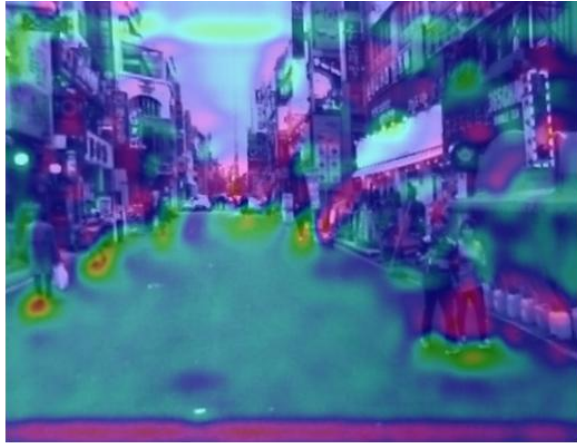
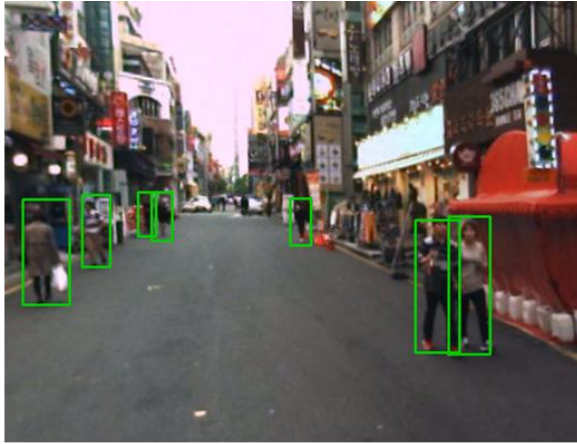
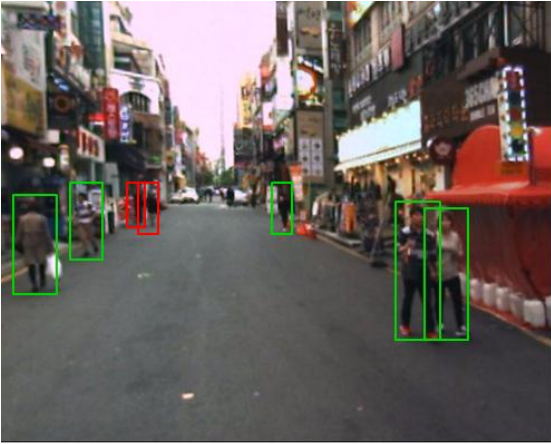
Modality-specific region proposals

Multimodal fusion



Experiment results

ABLATION STUDY

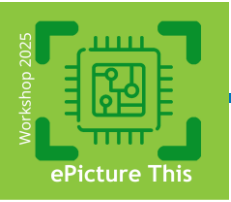


(a) Detection results w/o CAT

(b) Detection results w/ CAT

(c) Features w/o CAT

(d) Features w/ CAT



EXPERIMENTAL RESULTS

Method	Input	MR (partial occlusion)	MR (heavy occlusion)
Faster-RCNN	RGB	48.59	
Faster-RCNN	Thermal	47.59	
Faster-RCNN	RGB + Thermal	25.78	
Halfway fusion [1]	RGB + Thermal	25.75	
RPN + BDT [2]	RGB + Thermal	29.83	
IAF-RCNN [5]	RGB + Thermal	15.73	
IATDNN + IAMSS [9]	RGB + Thermal	14.95	
MSDS-RCNN [3]	RGB + Thermal	11.63	
GAFF [7]	RGB + Thermal	10.62	
AR-CNN [6]	RGB + Thermal	9.34	30.76
Ours	RGB + Thermal	8.02	28.74

[1] Liu, Jingjing, et al. "Multispectral deep neural networks for pedestrian detection." *BMVC*, 2016.

[2] Konig, Daniel, et al. "Fully convolutional region proposal networks for multispectral person detection." *CVPRW*, 2017.

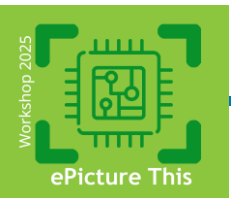
[3] Li, Chengyang, et al. "Multispectral pedestrian detection via simultaneous detection and segmentation." *BMVC*, 2018.

[5] Li, Chengyang, et al. "Illumination-aware faster R-CNN for robust multispectral pedestrian detection." *Pattern Recognition*, 2019.

[6] Zhang, Lu, et al. "Weakly aligned cross-modal learning for multispectral pedestrian detection." *ICCV*, 2019.

[7] Zhang, Heng, et al. "Guided attentive feature fusion for multispectral pedestrian detection." *WACV*, 2021.

[9] Guan, Dayan, et al. "Fusion of multispectral data through illumination-aware deep neural networks for pedestrian detection." *Information Fusion*, 2019.

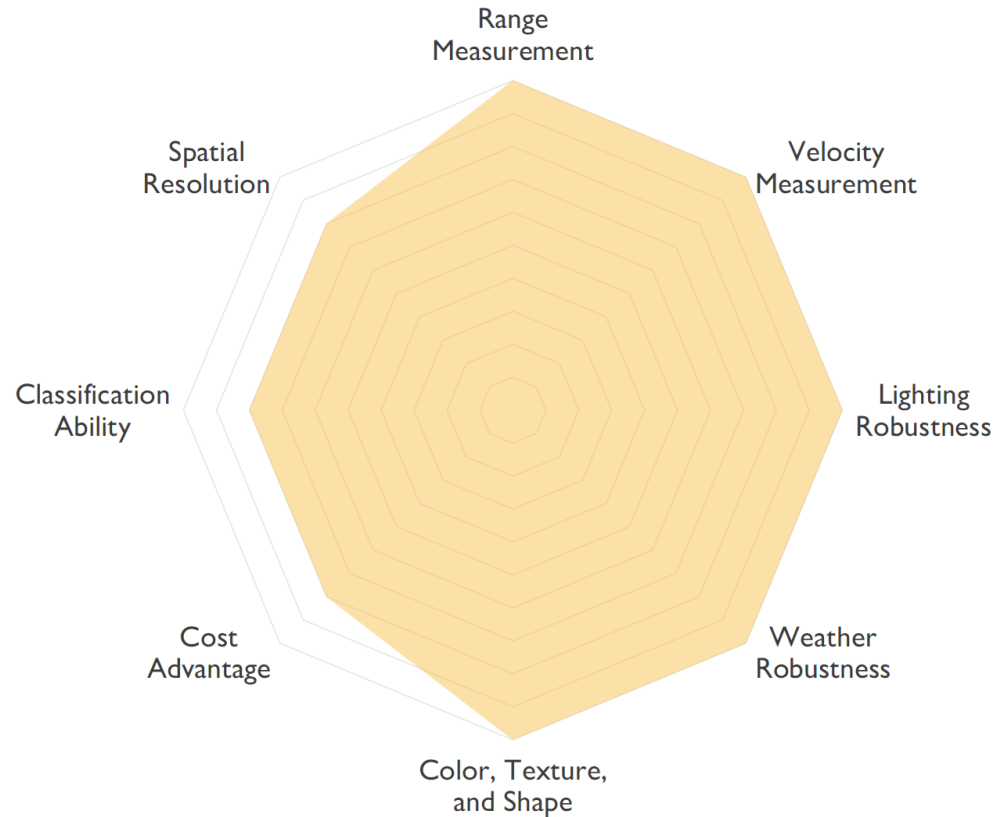


Visible-lidar-radar fusion-based VRU detection

VISIBLE-RADAR-THERMAL FUSION-BASED VRU DETECTION

- Radar: highly robust, low angular resolution
- Visible: high resolution, sensitive to illumination/weather
- Thermal: night vision, low contrast in warm weather
- Solution: cooperative fusion to improve accuracy and robustness

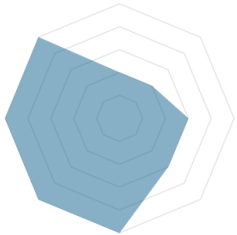
Radar + RGB + Thermal Sensors



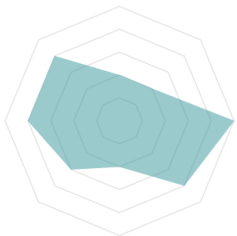
Radar Sensor



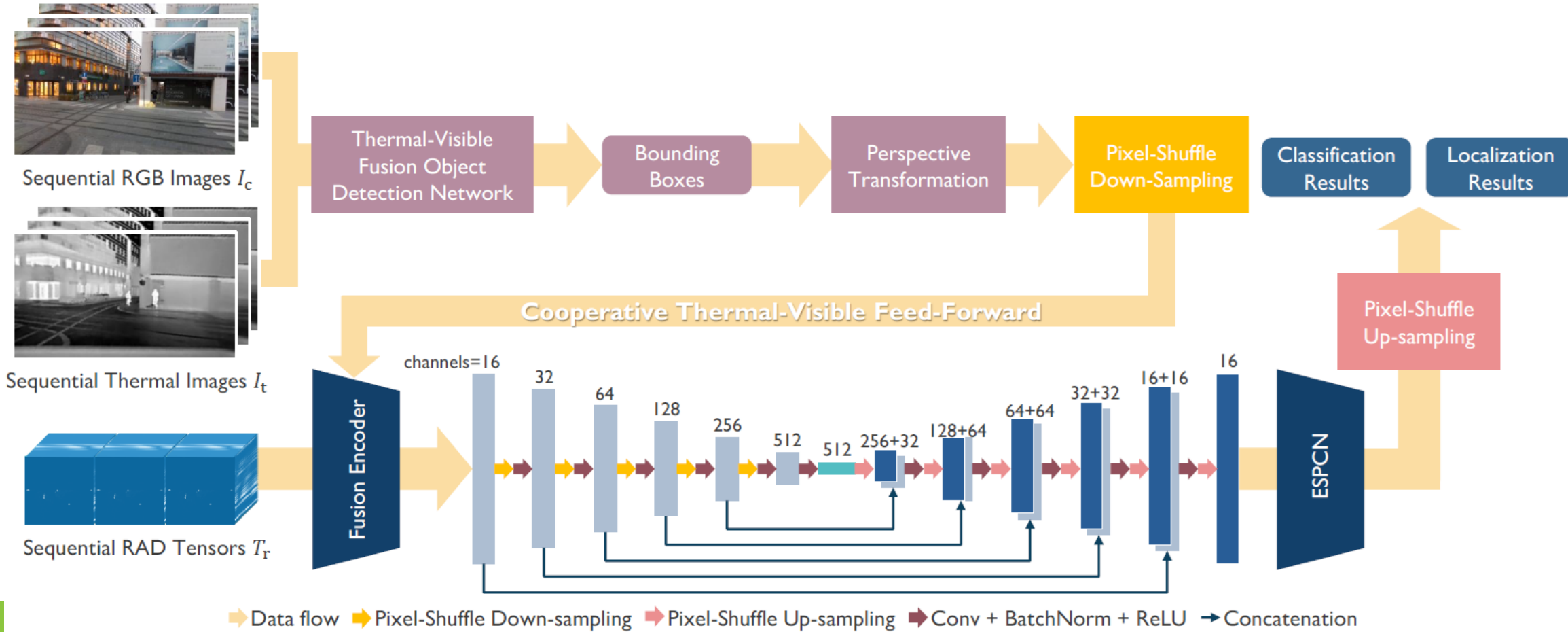
RGB Sensor



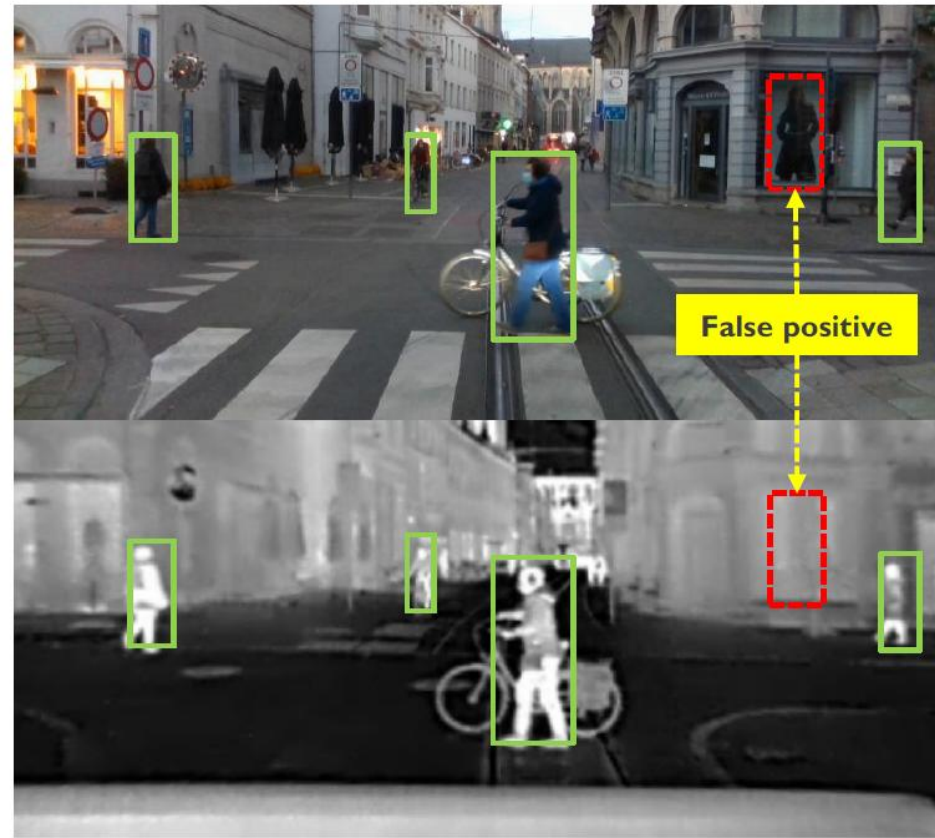
Thermal Sensor



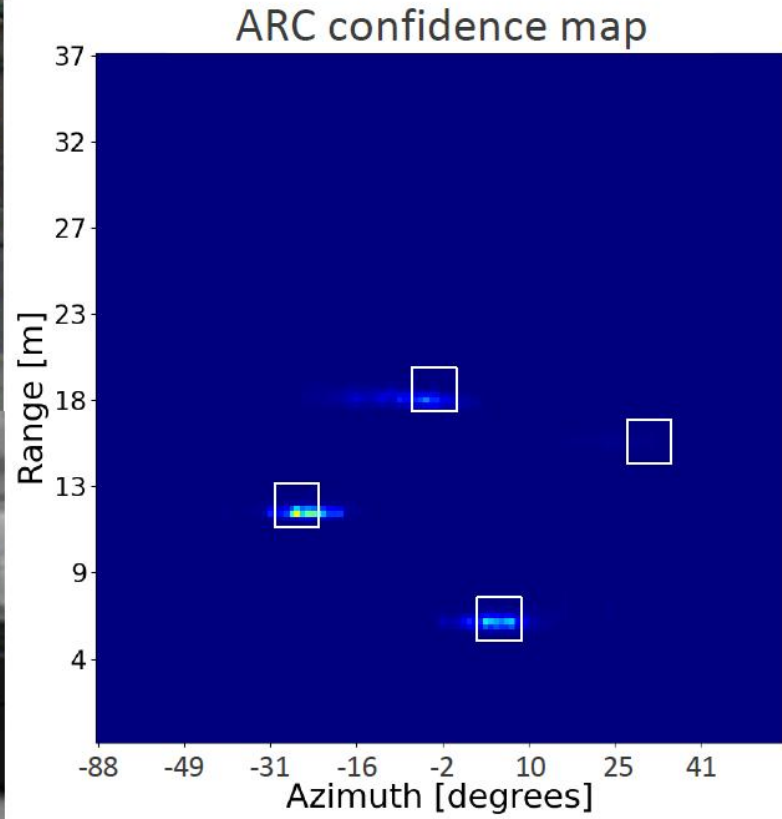
COOPERATIVE FUSION OF RGB, THERMAL AND LIDAR



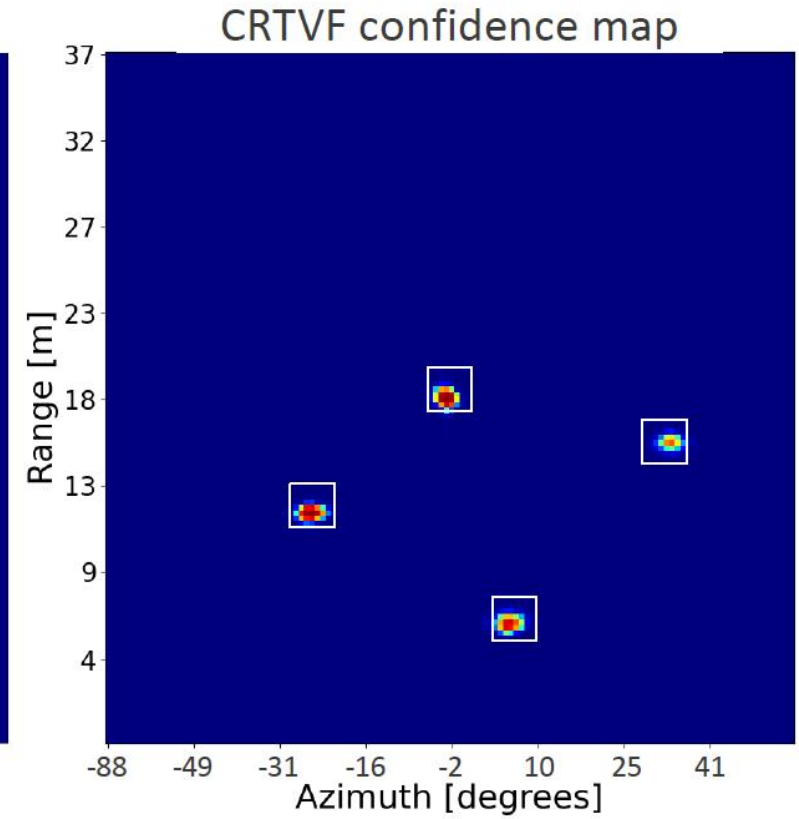
EXPERIMENTAL RESULTS



(a) Thermal-visible fusion network detections



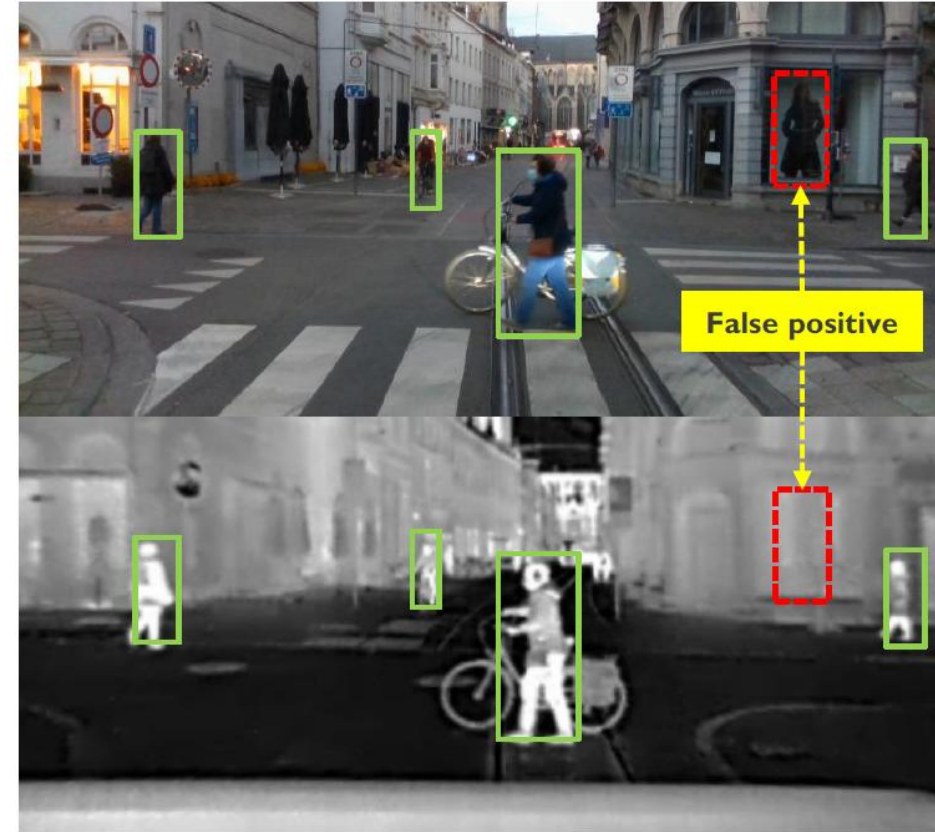
(b) Radar-based network confidence map



(b) Our confidence map

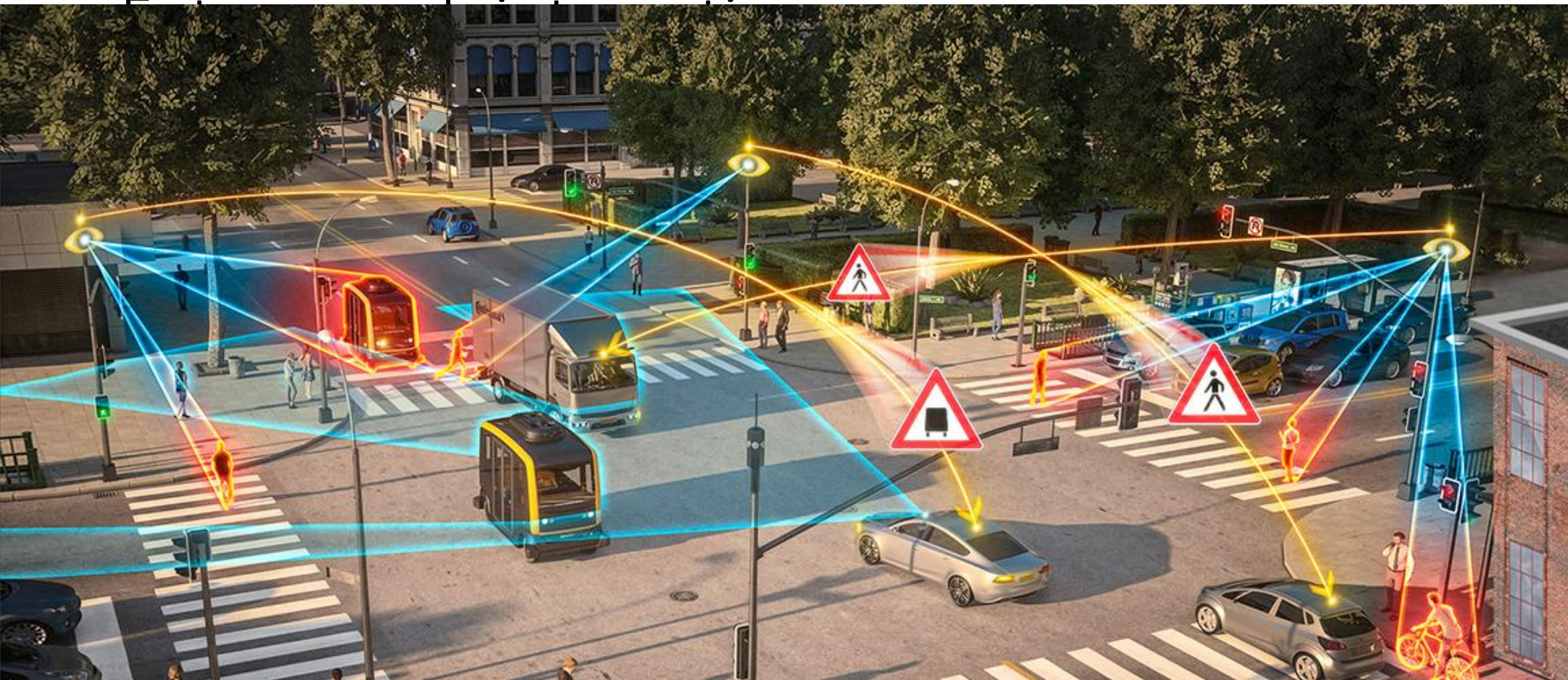
EXPERIMENTAL RESULTS

Method	Training Data	mAP	mean Dist. Error (m)	mean Azi. Error (rad)	mean Rng. Error (m)
ARC	Radar	0.51	0.64	0.08	0.39
CRVF-baseline	Radar	0.55	0.62	0.08	0.37
U-Net [114]	Radar+RGB	0.52	0.87	0.06	0.76
FCN [127]	Radar+RGB	0.35	0.94	0.06	0.82
Lim et al. [173]	Radar+RGB	0.67	0.63	0.05	0.51
CRVF	Radar+RGB	0.81	0.47	0.02	0.34
U-Net [114]	Radar+Thermal+RGB	0.57	0.84	0.06	0.71
FCN [127]	Radar+Thermal+RGB	0.41	0.91	0.06	0.80
Lim et al. [173]	Radar+Thermal+RGB	0.72	0.60	0.05	0.48
CRTVF	Radar+Thermal+RGB	0.89	0.45	0.02	0.32



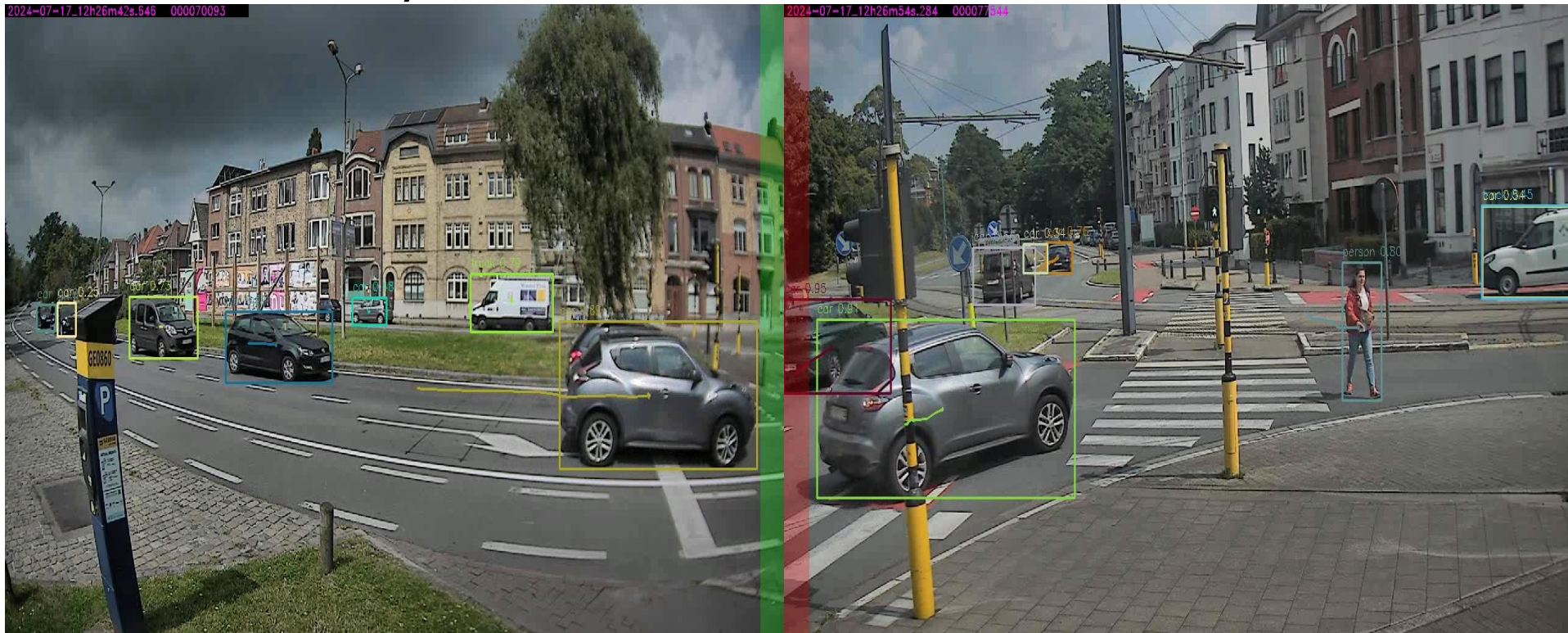
(a) Thermal-visible fusion network detections

Multi-sensor box tracking



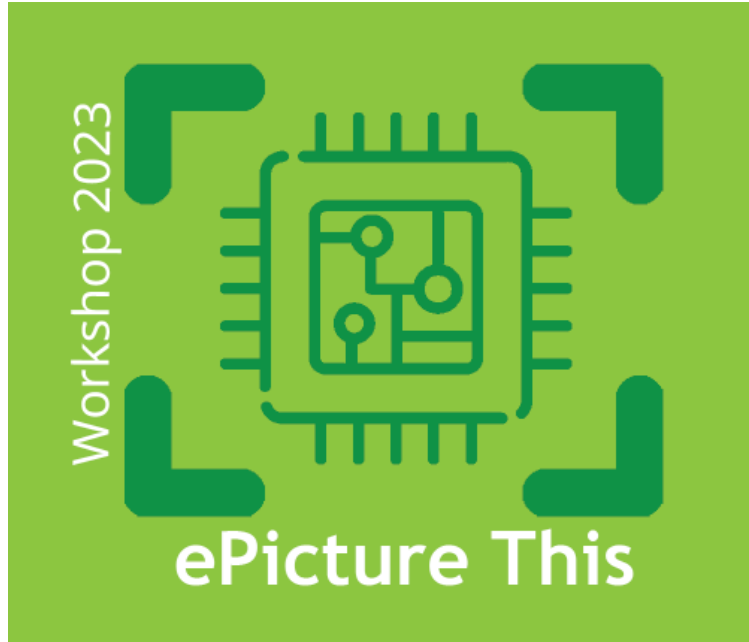
MULTI-SENSOR BOX TRACKING

- Tracking and prediction of VRU/vehicles and broadcasting alerts if a vehicle invisible from one camera has a predicted direction towards a road user that cannot see that vehicle otherwise
- Both static multi-modal sensors and sensors at the vehicle are included
- Tracking and prediction are distributed over multiple processes that are exchanging data and decisions
- Current version relies solely on visible streams



QUESTIONS





THANK YOU

an initiative by PENTA label projects
MANTIS and IMAGINATION with AENEAS support

