Lab S1-Basics and related work

Ekaterina Sysoykova k11842415
Marco Badici k11804695
Kevin Serrano k12215919
Dimitrios Koletsis k12216668
Lukas Renoth k12215903

Main problems/tasks to solve

Get familiar with properties of thermal images

- Pros and cons of RGB vs thermal images

Object detection on thermal imaging

- Related work on how to extract features which represent the target objects.

Multi-object tracking (MOT)

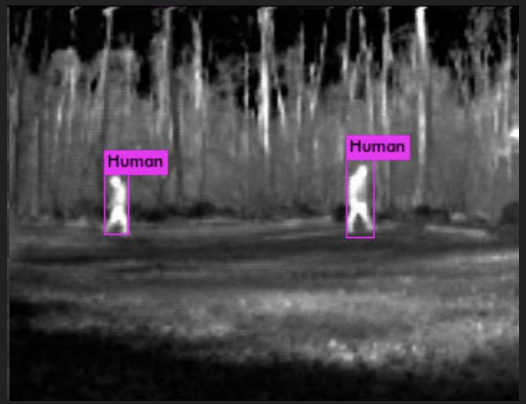
- Related work on how to perform accurate object association between video frames.

Pros and Cons of RGB vs thermal images

Camera type	Pros	Cons
Thermal	● Easier segmentation ● Independent of light conditions (ideal for nighttime applications)	● 1 channel proving thermal information ● Low resolution → no details, mostly silhouettes ● Low availability of open-source datasets
RGB	● 3 channels provide information on color and texture ● High resolution ● Wide variety of open-source datasets and trained CNN models	● Dependent on light conditions ● Shadows ● Harder segmentation

Object Detection in Thermal Imaging

YOLO (You Only Look Once) is a is a state-of-the-art, real-time CNN for object detection. This model, in combination with dataset augmentation, transfer learning and architectural modifications achieved high performance in object detection using thermal images as input [2][3].



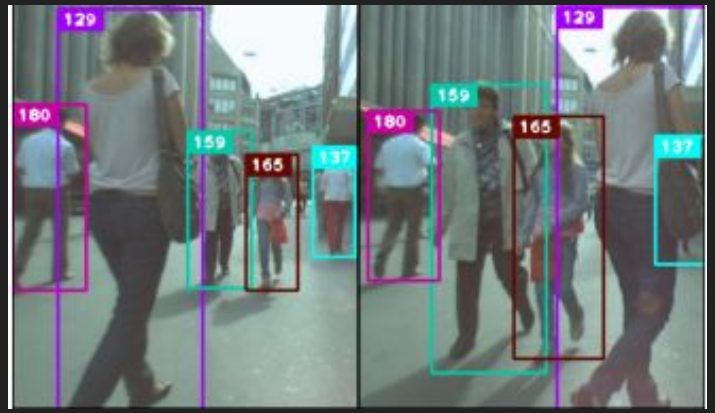
Multi-Object Tracking with DeepSORT

DeepSORT (*Deep* Sorting Online and Realtime Tracking):

- CNNs (*deep*) features per object → Kalman filter → Hungarian algorithm

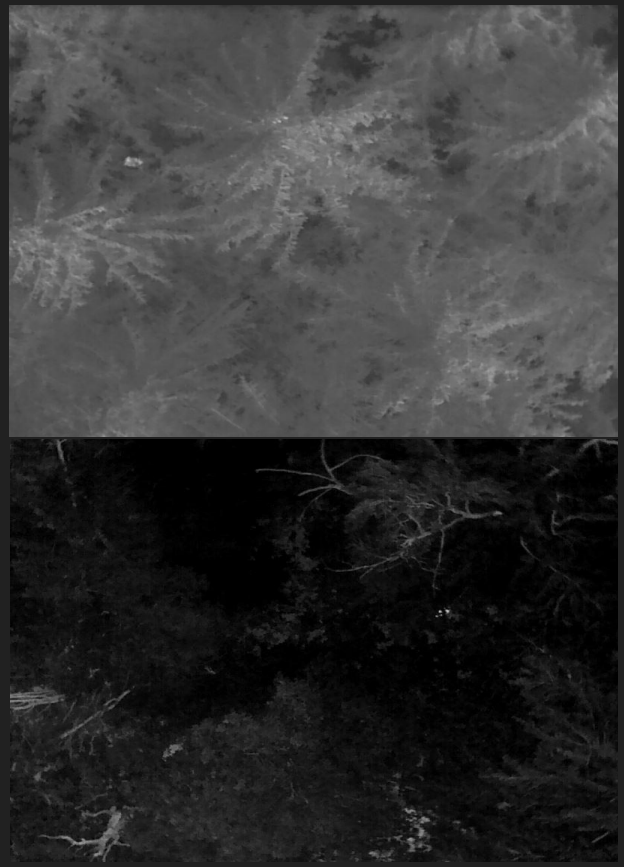
Advantages of *deep* association metric [4][5]:

- provides robustness against occlusions
- Reduces number of identity switches



Dataset insights

- 28 videos (1.27 GB)
- Duration: 18 to 80 seconds
- 30 FPS
- Target object is typically *hotter* than the background
- Small target size
- Trees in background are oscillating
- Contrast scene changes between videos
- Only one video contains two tracking targets



References

[1] Nielsen, Søren & Gade, Rikke & Moeslund, Thomas & Skov-Petersen, Hans. (2014). *Taking the Temperature of Pedestrian Movement in Public Spaces*. Transportation Research Procedia. 2. 10.1016/j.trpro.2014.09.071.

[2] M. Krišto, M. Ivasic-Kos and M. Pobar, *"Thermal Object Detection in Difficult Weather Conditions Using YOLO,"* in IEEE Access, vol. 8, pp. 125459-125476, 2020, doi: 10.1109/ACCESS.2020.3007481.

[3] Du, Shuangjiang & Zhang, Baofu & Zhang, Pin & Xiang, Peng & Xue, Hong. (2021). *FA-YOLO: An Improved YOLO Model for Infrared Occlusion Object Detection under Confusing Background*. Wireless Communications and Mobile Computing. 2021. 1-10. 10.1155/2021/1896029.

[4] N. Wojke, A. Bewley and D. Paulus, *"Simple online and realtime tracking with a deep association metric,"* 2017 IEEE International Conference on Image Processing (ICIP), 2017, pp. 3645-3649, doi: 10.1109/ICIP.2017.8296962.

[5] Wang, Yu-Hsiang. (2022). *SMILEtrack: SiMIlarity LEarning for Multiple Object Tracking*. 10.48550/arXiv.2211.08824.

[6] Xuerui Dai, Xue Yuan, and Xueye Wei. 2021. *TIRNet: Object detection in thermal infrared images for autonomous driving*. Applied Intelligence 51, 3 (Mar 2021), 1244–1261. https://doi.org/10.1007/s10489-020-01882-2

Additional resources

1ST PRESENTATION: BASICS AND RELATED WORK

- Identify the main problems/tasks that you need to solve for this project
- Find related scientific publications and present a quick overview
- Get familiar with the data (start programming and working with videos)
- Exemplary related work:
  - Amala Arokia Nathan, Rakesh John, Indrajit Kurmi, and Oliver Bimber. "Inverse Airborne Optical Sectioning." Drones 6.9 (2022): 231.
  - Amala Arokia Nathan, Rakesh John, et al. "Through-Foliage Tracking with Airborne Optical Sectioning." Journal of Remote Sensing 2022 (2022).
  - Jia, Jinlu, et al. "Aerial video trackers review." Entropy 22.12 (2020): 1358.
  - Bondi, Elizabeth, et al. "BIRDSAI: A dataset for detection and tracking in aerial thermal infrared videos." Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2020.

Teledyne FLIR Thermal Dataset

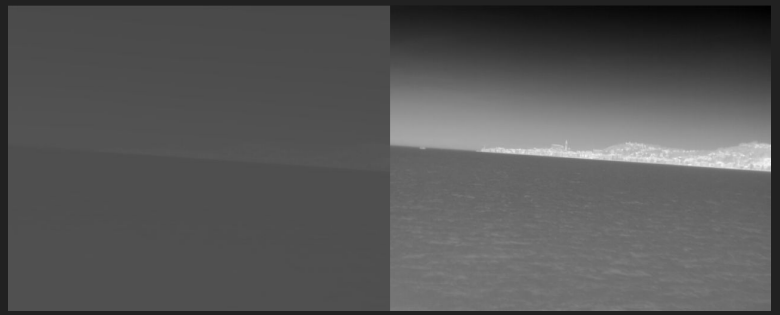
Data constructed to encourage research on visible + thermal sensor fusion algorithms ("RGBT") and empower the automotive community to create safer and more efficient ADAS and driverless vehicle systems.

https://www.flir.eu/oem/adas/adas-dataset-form/

Teledyne FLIR Thermal Dataset Specifications

26,442 fully annotated frames with **520,000** bounding box annotations across 15 different object categories.

- Thermal - 14-bit TIFF (no AGC)
- Thermal 8-bit JPEG (AGC applied)
- RGB - 8-bit JPEG
- MSCOCO formatted annotations (JSON)
- Conservator formatted annotations (JSON)



Popular RGB object detection dataset for comparison

- COCO (Microsoft Common Objects in Context)
  - 330K images
  - 1.5 million object instances
  - 80 object categories

https://www.flir.eu/oem/adas/adas-dataset-form/, https://cocodataset.org/#home