



Institute of Automation

IROS 2018

10th Workshop on Planning, Perception and Navigation for Intelligent Vehicles (PPNIV)

DisNet: A novel method for distance estimation from monocular camera

Muhammad Abdul Haseeb, Jianju Guan, Danijela Ristic-Durrant, Axel Gräser

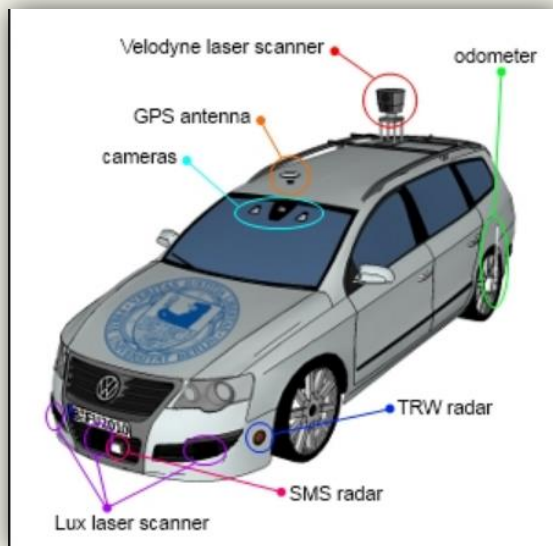
Madrid, 01.10.2018



Universität Bremen

Motivation

- ❖ reliable and accurate detection of obstacles is one of the challenges of safe autonomous driving
- ❖ the main principle of obstacle detection in front of a vehicle from the automotive sector can be applied to railway applications
- ❖ one of the key challenges is **long-range obstacle detection**



[1]



[2]

[1] AutoNOMOS Labs, <http://autonomos.inf.fu-berlin.de>

[2] Accurate 3D-vision-based obstacle detection for an autonomous train, J. Weichselbaum, C. Zinner, O. Gebauer, W. Pree, Computers in Industry Volume 64, Issue 9, December 2013, Pages 1209-1220

SMART - obstacle detection system

- ❖ a novel fully integrated multi-sensor on-board system
- ❖ for mid (up to 200 m) and long range (up to 1000 m) obstacle detection (approx. required braking distance for freight trains ~ 700 meters at 80 km/h)
- ❖ can operate in day and night conditions as well as in poor visibility conditions

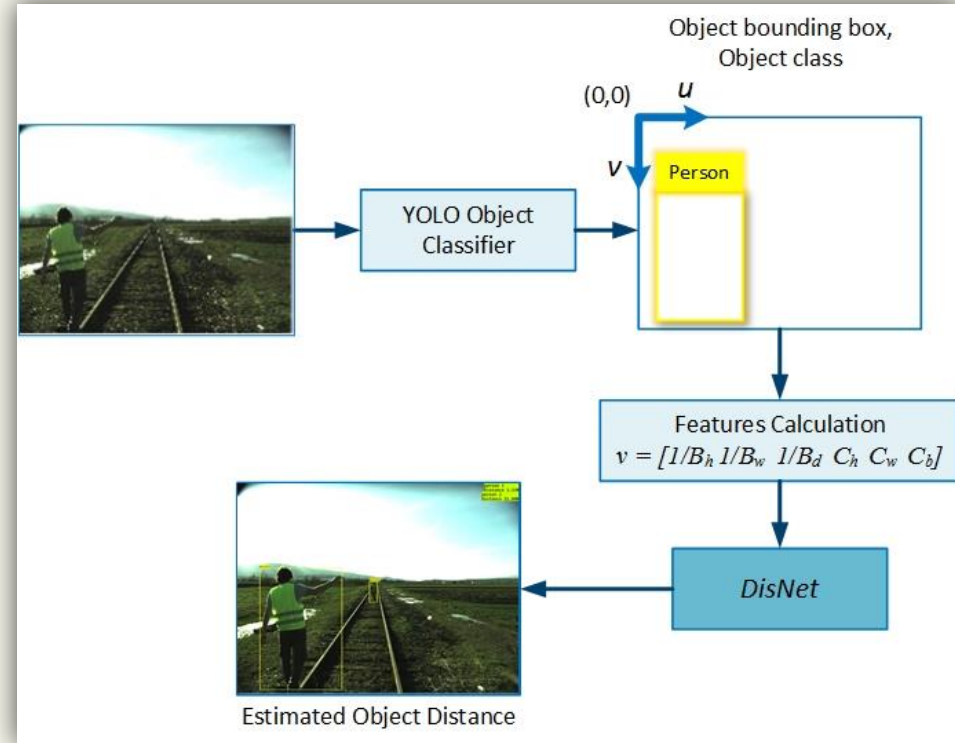


Proposed Solution

A novel **machine learning-based method** for long-range obstacle detection and distance estimation from a single monocular camera

❖ Working principle

Object distance estimation is based on learning the change in object appearance in an image (in terms of size) due to the change of the object distance with respect to the camera viewing the object.



Features and DisNet Structure

❖ **Dataset** - manually extracted bounding boxes of different objects from the images labelled with real distance measured using LiDAR.

❖ **Features** – Feature vector: $v = [1/B_h \ 1/B_w \ 1/B_d \ C_h \ C_w \ C_b]$

B_h =(height of the object bounding box in pixels/image height in pixels)

B_w =(width of the object bounding box in pixels/image width in pixels)

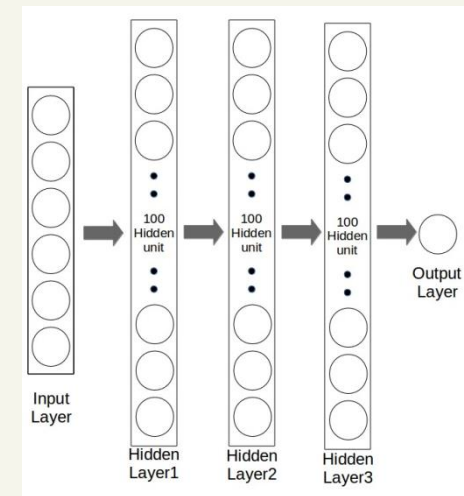
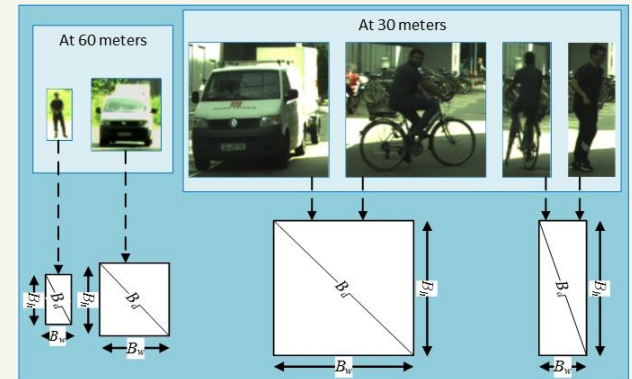
B_d =(diagonal of the object bounding box in pixels/image diagonal in pixels)

C_h, C_w, C_b are the values of average height, width and breadth of an object of the particular class.

❖ **DisNet Structure** - A Neural Network with 3 hidden layers with 100 hidden neurons per layer.

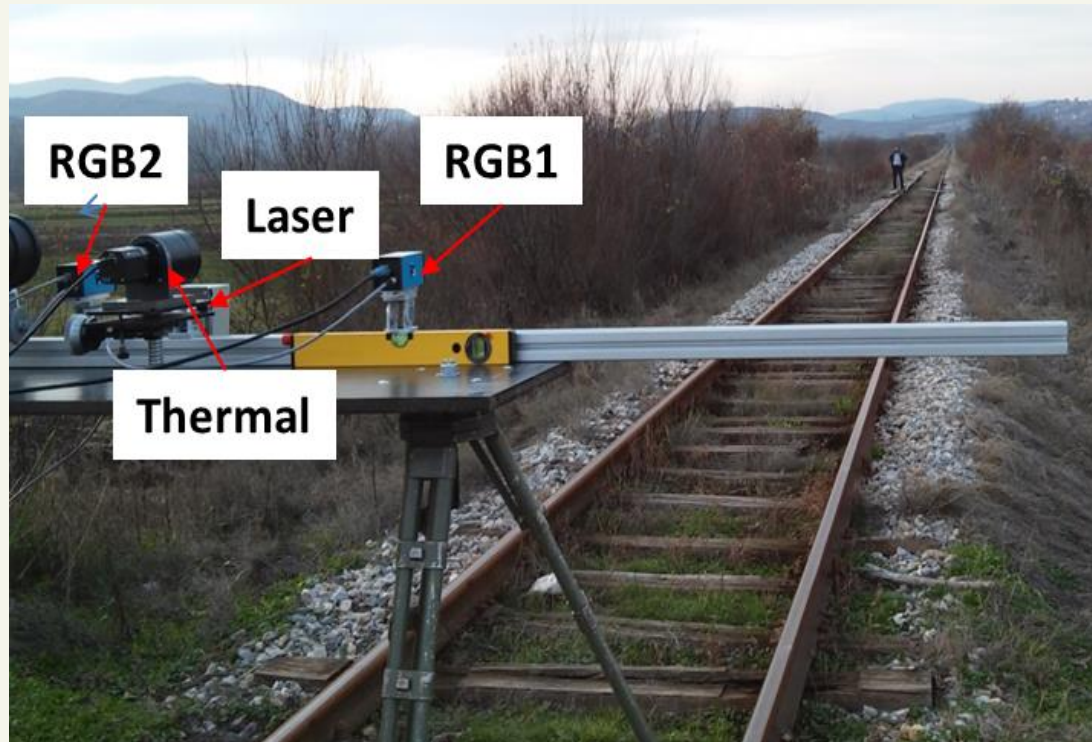
Supervised learning for the training of DisNet:

- training set of 2000 features;
- ground truth - LiDAR distance measurement



Static experiment

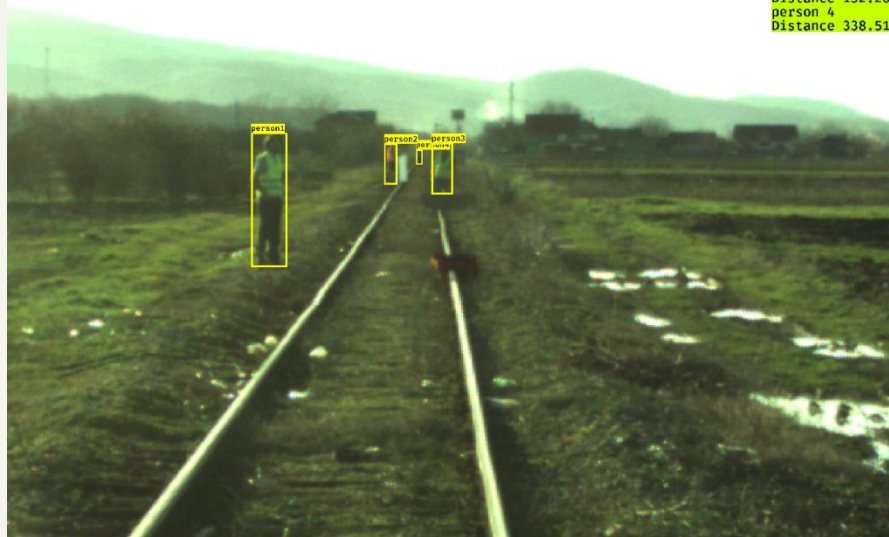
- Setup for recording the scene



Test-stand with the cameras and laser scanner viewing the rail tracks and an object

Results (Static experiment)

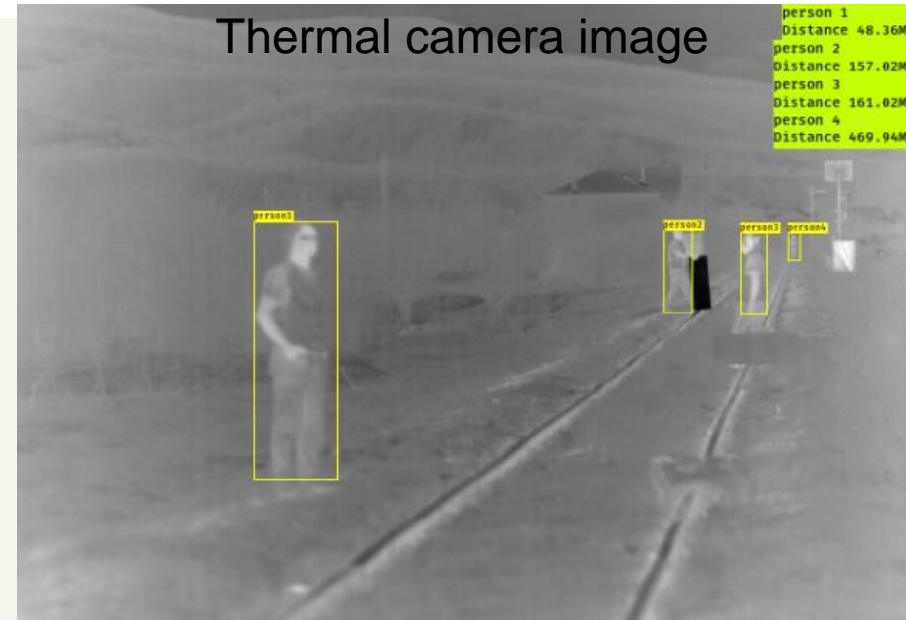
RGB camera image



```

person 1
Distance 54.26M
person 2
Distance 167.59M
person 3
Distance 132.26M
person 4
Distance 338.51M
    
```

Thermal camera image



```

person 1
Distance 48.36M
person 2
Distance 157.02M
person 3
Distance 161.02M
person 4
Distance 469.94M
    
```

Object distance with respect to Obstacle Detection System

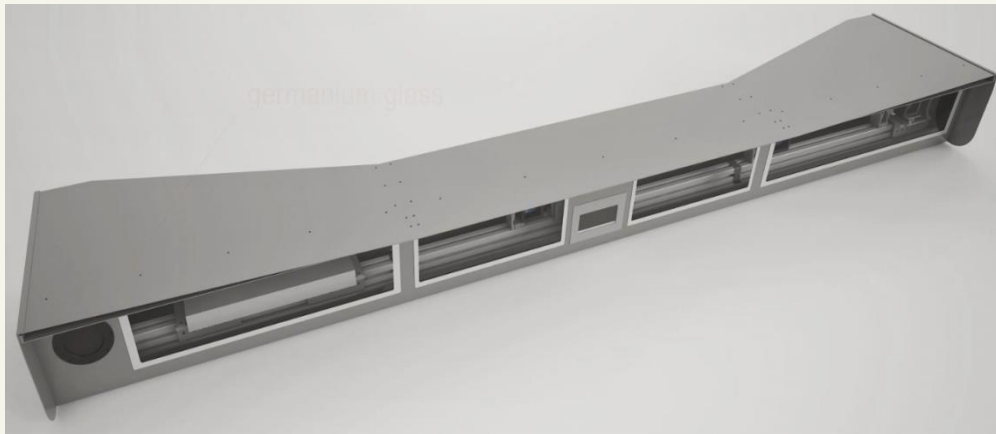
Ground Truth		RGB Camera	Thermal Camera
Person 1	50 m	54.26 m	48.36 m
Person 2	100 m	132.26 m	161.02 m
Person 3	150 m	167.59 m	157.02 m
Person 4	300 m	338.51 m	not-visible
Person 5	500 m	not-visible	469.94

Dynamic experiment

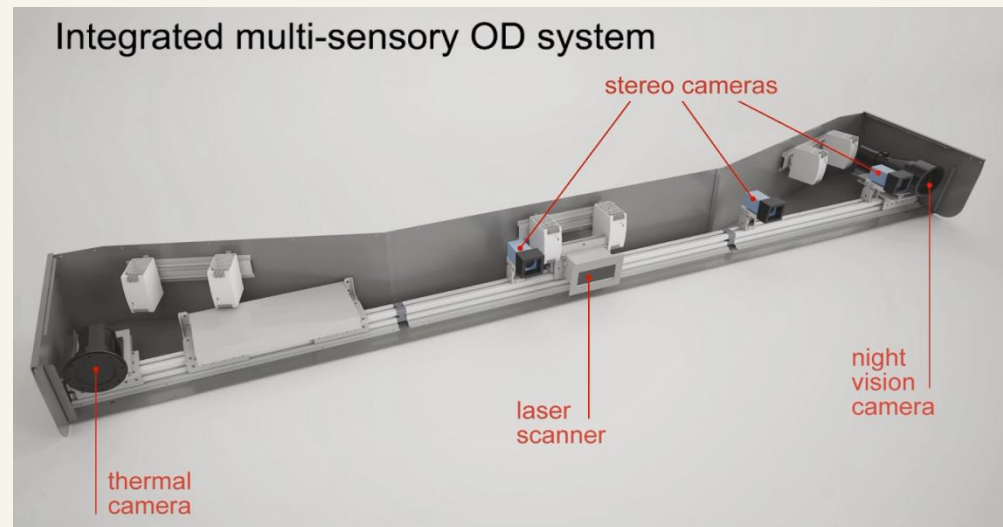


Mounted on Serbian locomotive

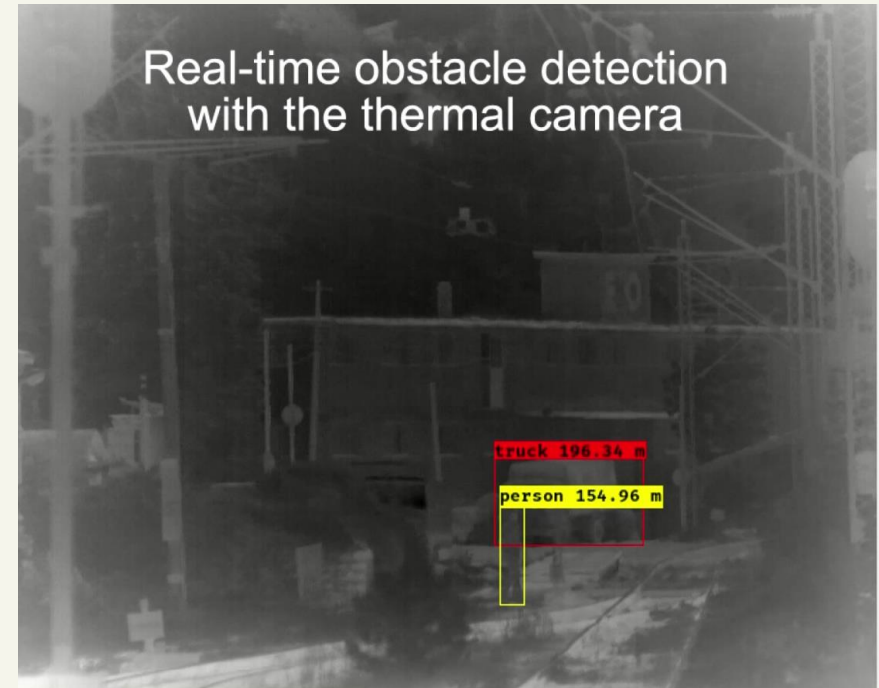
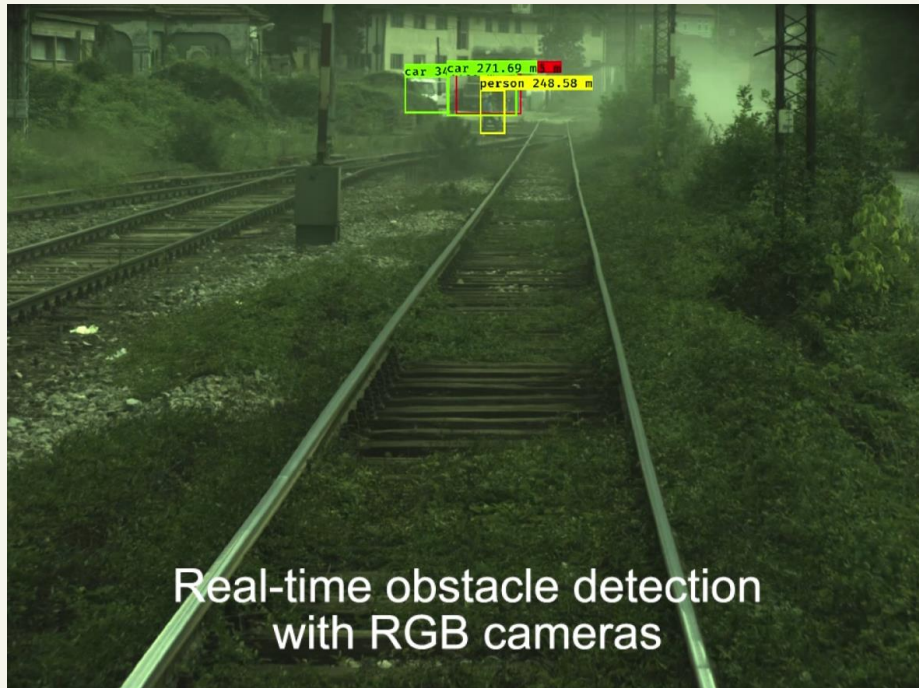
Dynamic experiment



SMART - multi-sensor obstacle detection system



Results (dynamic experiments)



Results (Road Scene)



Thank you very much
for your attention!

Muhammad Abdul Haseeb
Institute of Automation
University of Bremen
Bremen, Germany
haseeb@iat.uni-bremen.de