

# Real-Time Analysis through computer vision on dashcams and triggers in High Energy Physics

ESR 11: Henrique Piñeiro Monteagudo

SMARTHEP Annual meeting, Lund November 2023

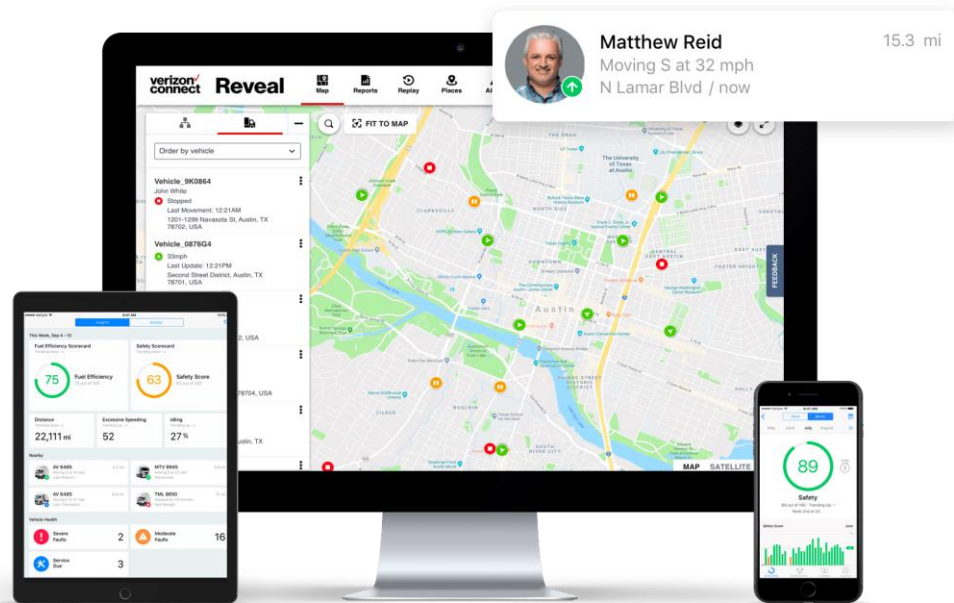


SMARTHEP is funded by the European Union's Horizon 2020 research and innovation programme, call H2020-MSCA-ITN-2020, under Grant Agreement n. 956086

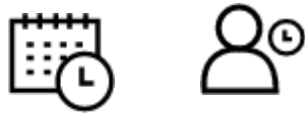
# About me

- PhD student in Computer Science and Engineering at the University of Bologna and Verizon Connect
  - Supervised by Francesco Sambo, Leonardo Taccari and Samuele Salti
- Background
  - Worked in European research projects on Computer Vision applications
  - MSc in Computer Vision (Universities of Santiago, Coruña, Vigo and Porto)
  - BSc in Automation Engineering (University of Vigo)
- Born and raised in Santiago de Compostela (Spain)

# Verizon Connect



## Field management



## Compliance Management



## Integrated Video



# Road scene understanding with dashcams

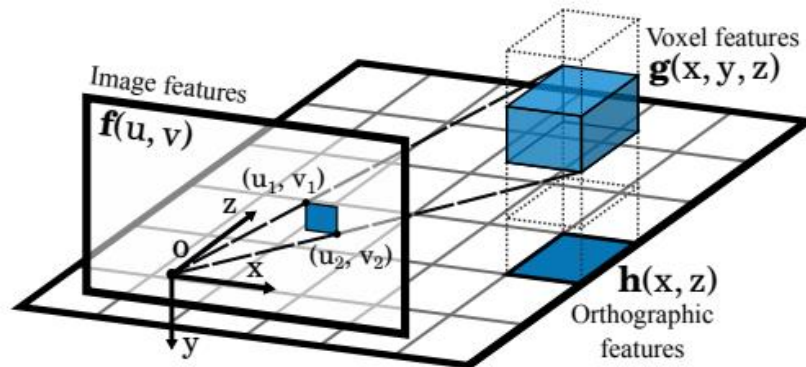


# What does this have to do with SMARTHEP?

- **Real time analysis**
- Tasks of interest
  - Accident anticipation
  - Collision warnings
  - Trajectory forecasting
- These tasks are difficult to solve from monocular images
  - Loss of scene structure from the 3D → 2D projection in image formation

# Bird's Eye View Perception

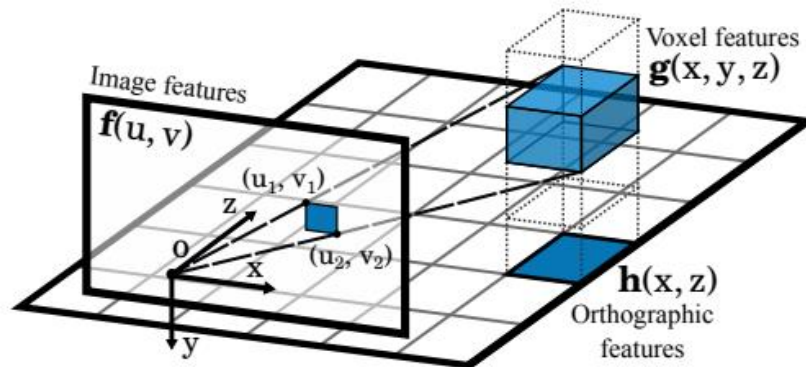
- Bird's Eye View: a 2D orthographic projection of the world along the direction of gravity



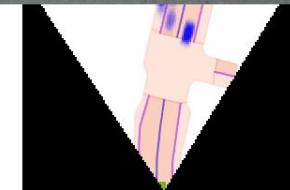
- Desirable representation for road scenes
  - Road agents' movement mostly restricted to ground plane
  - More compact than an explicit 3D representation
  - Easy to integrate additional cameras/sensors

# Bird's Eye View Perception

- Bird's Eye View: a 2D orthographic projection of the world along the direction of gravity



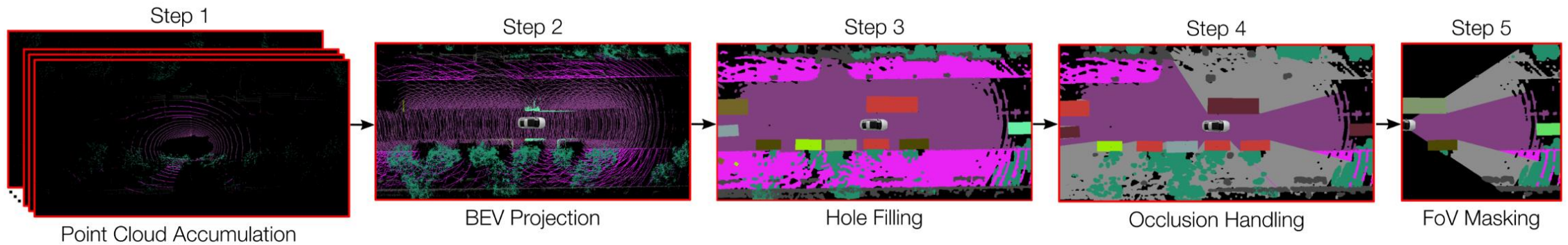
- Desirable representation for road scenes
  - Road agents' movement mostly restricted to ground plane
  - More compact than an explicit 3D representation
  - Easy to integrate additional cameras/sensors



Vehicle segmentation in BEV with *SimpleBEV*, Harley et al., ICRA 2023  
nuScenes dataset, underlying map is ground truth

# Data problem

- Current work focuses on supervised approach.
- But **how** do you generate BEV ground truth labels?

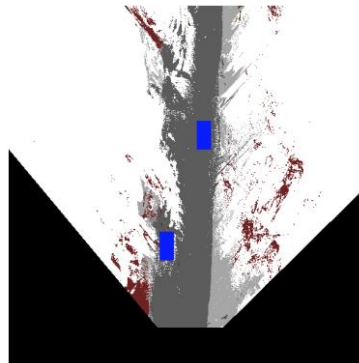


Pipeline to generate BEV panoptic segmentation labels from annotated point clouds in the KITTI360 dataset  
in *PanopticBEV*, Gosala and Valada, *Robotics and Automation Letters* 2022



# Pseudolabels

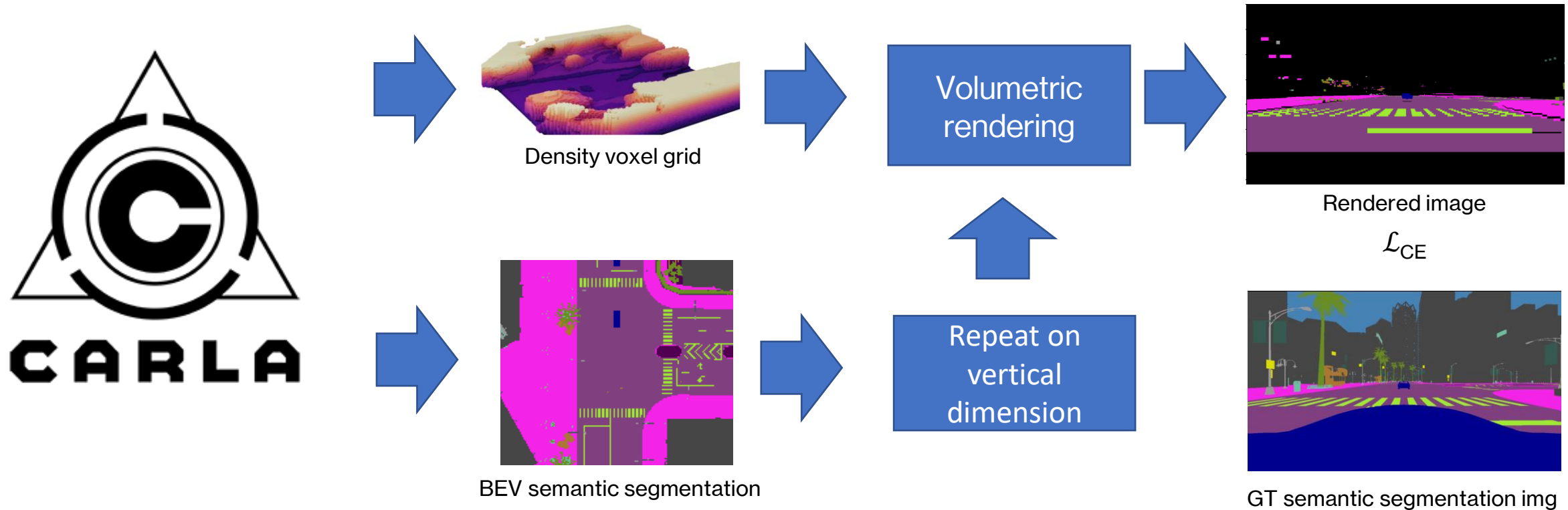
- Generating BEV segmentation maps with monocular camera, gps and accelerometer data
  - Inspired by *SkyEye, Gosala et al, CVPR 2023* and existing internal work



Internal data

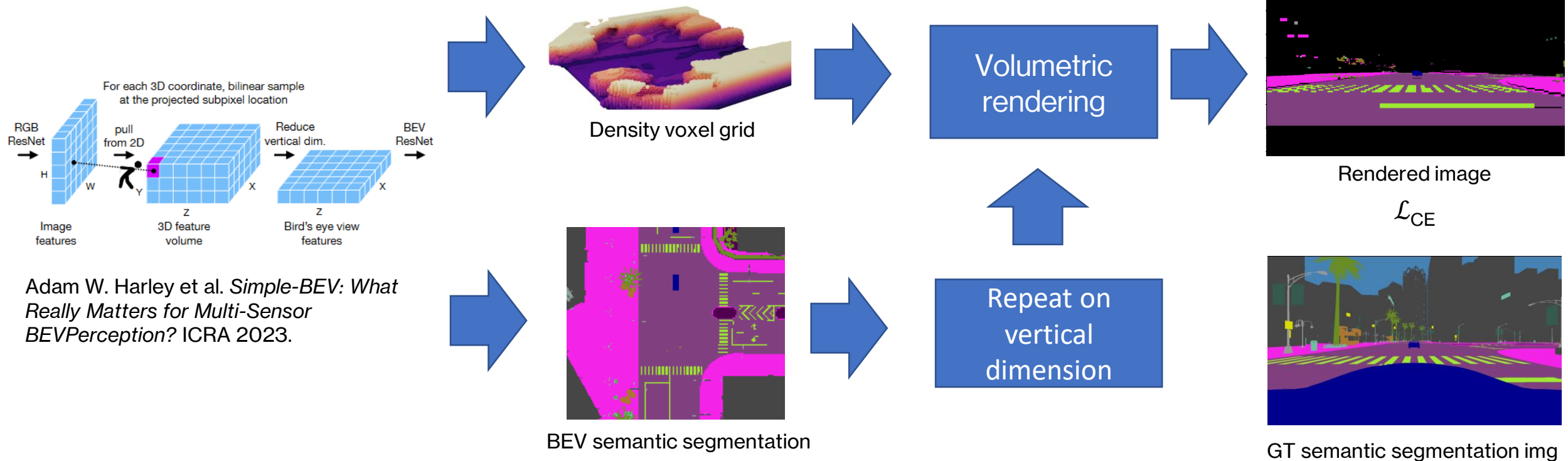
# Self-supervision

- Can we reconstruct the semantic seg of the input from a density volume and the semantic BEV?



# Self-supervision

- Can we reconstruct the semantic seg of the input from a density volume and the semantic BEV?



Adam W. Harley et al. *Simple-BEV: What Really Matters for Multi-Sensor BEV Perception?* ICRA 2023.

# Challenges and current work

- The described approach is too unconstrained
  - Scale problem
  - Many possible wrong 3D geometries and BEV semantic segmentation can generate a correct reconstructed input image
- How to resolve them?
  - Further constrain the model
    - Using more than a singular frame and forcing consistency among frames
    - Supervising with a loss on the RGB values, not only on the semantics

# Other Activities

## Training

- Different PhD courses from the University of Bologna
- ACDL summer school on Data Science and Machine Learning

## Tutoring

- Supervising a master student doing his internship and master's thesis at Verizon Connect

## Conferences

- Attended the IEEE Intelligent Transportation Systems Conference at Bilbao to present a paper stemming from an innovation project in Verizon Connect

*“An object detection approach for lane change and overtake detection from motion profiles”*

## Secondment

- Starting next week with prof. Caterina Doglioni on anomaly detection on HEP data

# Conclusion

- Road scene understanding from dashcams in real time
  - Certain tasks like **accident anticipation** and **trajectory forecasting** are difficult to tackle in the **monocular image** space
  - We would like to obtain a more appropriate representation for these tasks
- Obtaining a **Bird's Eye View representation** without 3D data
  - Explored using pseudolabels with unsatisfactory results
  - Working on a **self-supervised approach with volumetric rendering**

Thank you for your attention and feel free to reach me via email!