1.5 Method to Fuse Map Layer and Pedestrian State Space to Improve 3D Detector Output

Ravi Kothari, Aditya Kumar Agarwal, Federica Paolicelli
| AVL Software and Functions GmbH

Introduction
Due to the black box nature of AI models and the nature of their training strategy, the output does not inherently align with the real world. We propose a method to combine the location of the object (pedestrian) and its detected state space attributes to define a decision boundary. The updated detections are more reliable and improve performance in real-world applications.

Methodology
We have tested our approach with real world dataset (nuScenes [1]) and synthetic data generated under the KI Wissen project. For the baseline model, a LiDAR-based 3D-detector (PointPillars [2]) with a Hungarian [3] association-based tracker has been chosen. The method uses the following attributes to calculate the plausibility score:

- 3D-detector output (aspect ratio, occlusion, confidence score)
- Map layers (drive lane, parking zone, crossing, stop line and walkway)
- State space (velocity, acceleration, distance travelled)

The incorporation of map layers stems from the observation that pedestrian behavior varies with the context. For example, humans generally walk at a consistent velocity while crossing a road. Thus, our approach uses the map layer information to detect conformity.

Implementation
The above mentioned attributes are combined to form a plausibility score (0-1). As all the features do not belong to the same dimension, it is difficult to form a hard boundary with linear weights. Hence, we rely on the XGBoost regressor [4] to generate a fuzzy boundary.

The state space features are first normalized using Kernel Density Estimation (KDE) and the map layers are one-hot encoded.

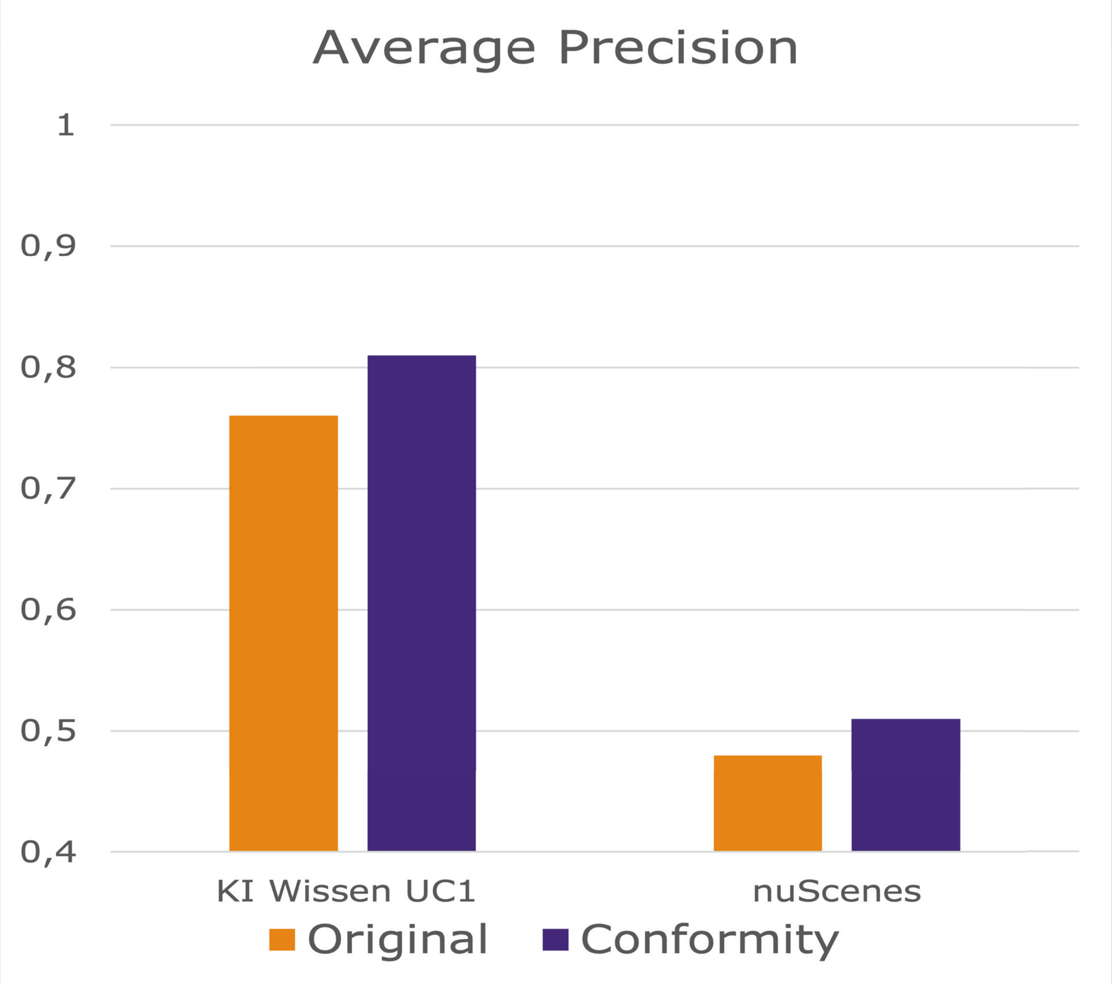

Figure 2: Average Precision comparison between baseline model and conformed model. For KI Wissen data, we use five scenes from UC1. (© AVL Software and Functions GmbH)

Results
The initial analysis shows positive improvements in mean average precision (mAP) values for pedestrians. We tested the method on two datasets and used 20% of the detection results to fine-tune the XG Boost. In nuScenes, the AP is improved by 2 points, as compared to a 5-point improvement in KI Wissen data. We believe the discrepancy is due to the higher data diversity in nuScenes than in synthetic data. Although the results are promising, further investigation is required to determine the impact of different detectors, datasets, and classes.

References
[1] Holger Caesar et al. “nuScenes: A multimodal dataset for autonomous driving”. In: arXiv preprint arXiv:1903.11027 (2019)
[2] Alex H. Lang et al. “PointPillars: Fast Encoders for Object Detection from PointClouds”.URL: http ://arxiv. org/abs/1812.05784.
[3] Harold W. Kuhn. “The Hungarian Method for the Assignment Problem”. In: Naval Research Logistics Quarterly 2.1–2 (Mar. 1955), pp. 83–97. DOI: 10.1002/nav.3800020109.
[4] Tianqi Chen et al. “XGBoost: A Scalable Tree Boosting System”. URL: http://doi.acm.org/10.1145/2939672.2939785

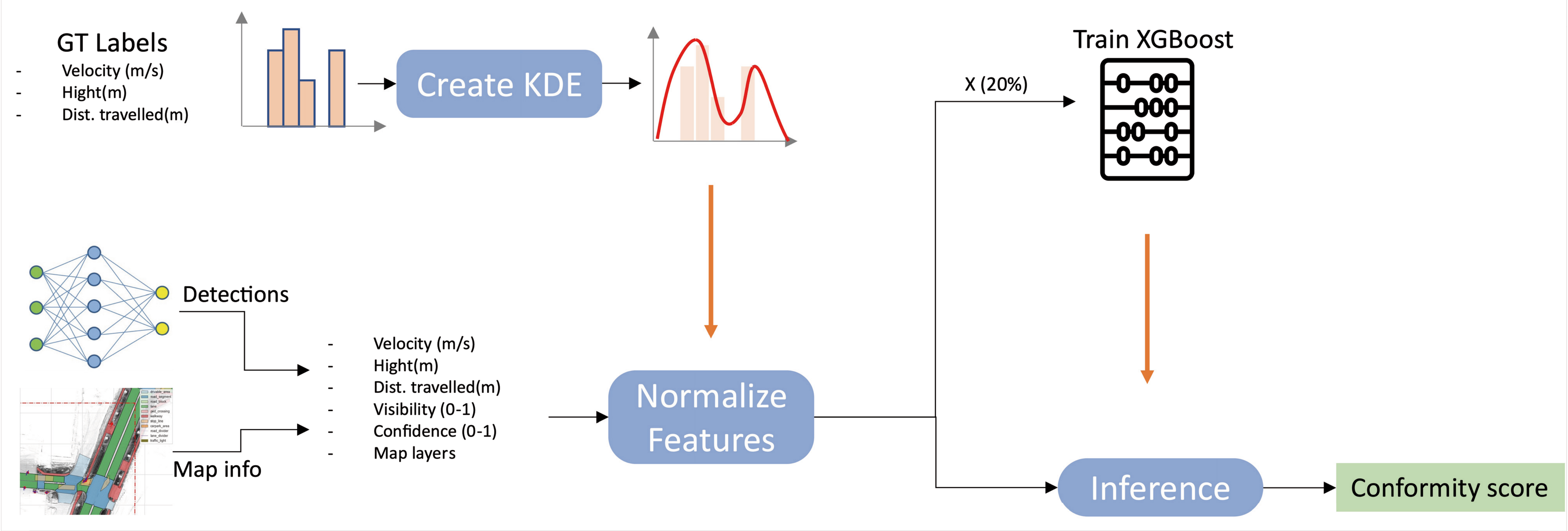

Figure 1: Overview of the Conformity Pipeline (© AVL Software and Functions GmbH)

Detection

Partners

External partners

For more information contact:
ravi.kothari@avl.com
federica.paolicelli@avl.com

KI Wissen is a project of the KI Familie. It was initiated and developed by the VDA Leitinitiative autonomous and connected driving and is funded by the Federal Ministry for Economic Affairs and Climate Action.

www.ki-wissen.de @KI_Familie KI Familie

Funded by the European Union
NextGenerationEU

Supported by:
Federal Ministry for Economic Affairs and Climate Action
on the basis of a decision by the German Bundestag