

Figure 1: Processing pipeline from the initial sensor data, through image segmentation, to a textual scene description used for scene search.

Motivation

- A pipeline for automatically generating textual scene descriptions, eliminating the need to manually search for specific situations in large datasets to train supervised ML models.

Segmentation

- Objects are masked for a more accurate scene description.
- Zero-shot image segmentation from SAM used, utilizing bounding boxes as prompts for meaningful segments (Fig. 3).

Sensor Fusion (Cam, Lidar) for Auto-Labeling

- Multimodal sensor fusion is crucial for accurate 3D object detection.
- BEV-Fusion framework [1] explored for multi-sensor fusion on KI-DT dataset.
- Multi-sensor fusion is used to prompt image segmentation (Fig. 3).

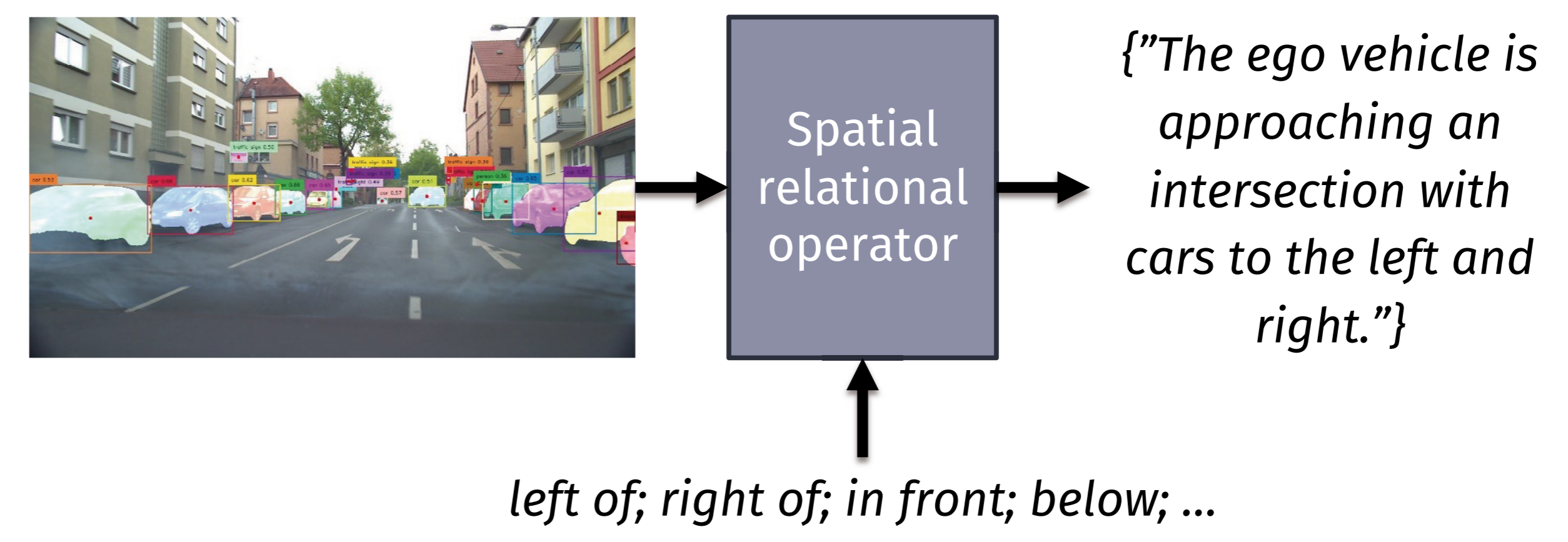


Figure 4: Establishing connections between objects through spatial information to convey their relationships.

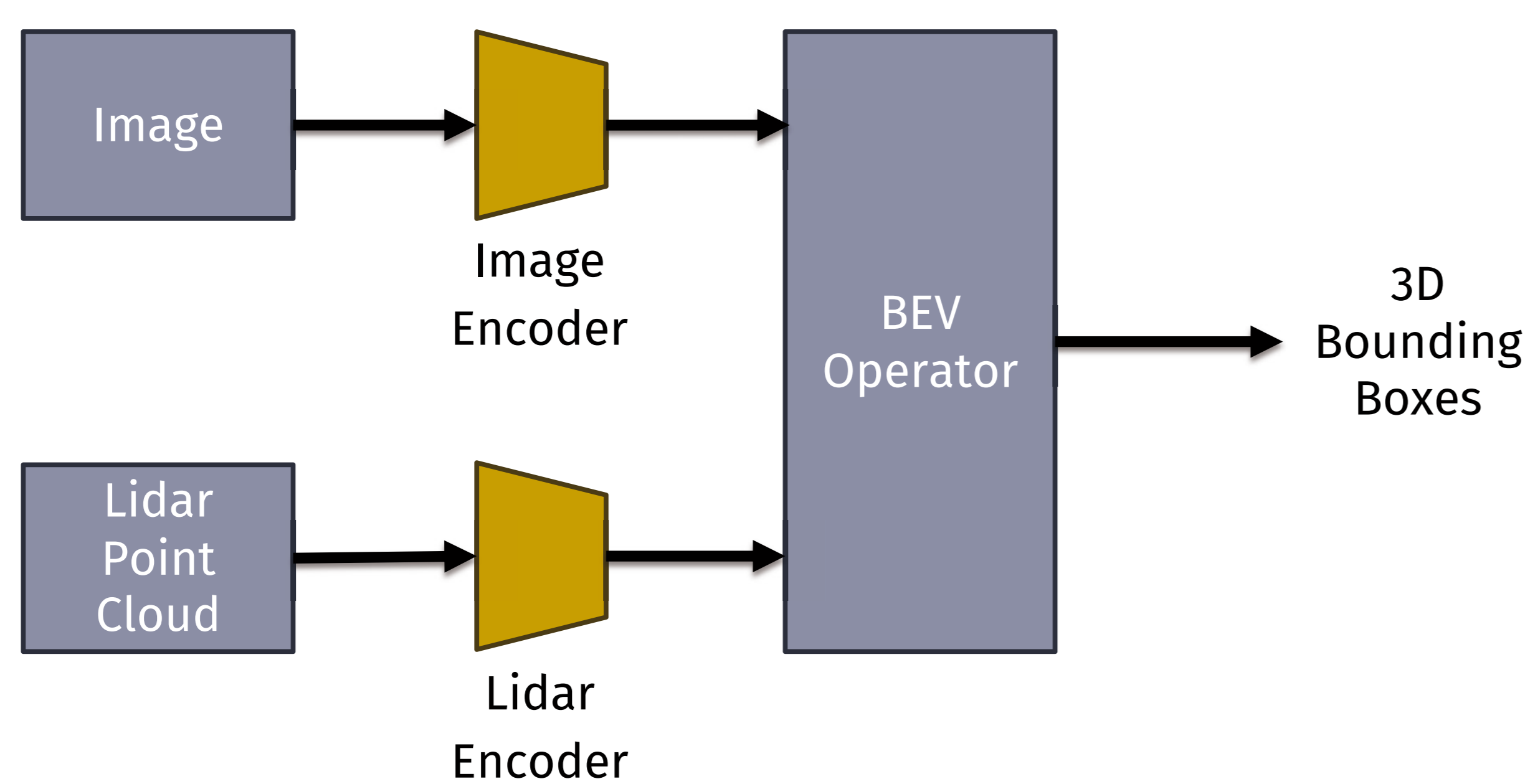


Figure 2: Block diagram showing the steps of sensor fusion using the BEVFusion framework [1].

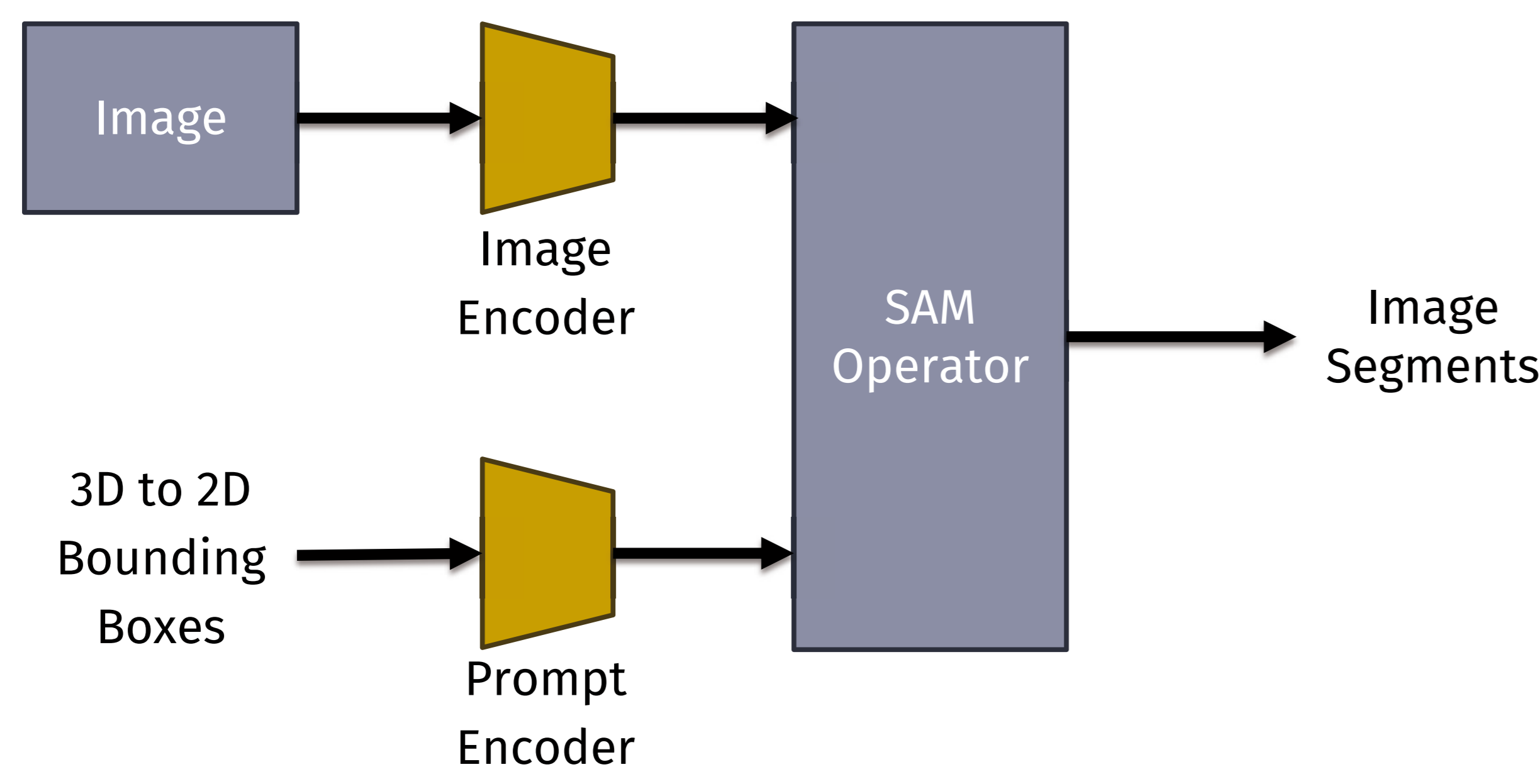


Figure 3: Segment Anything Model (SAM) [2] to mask objects. By prompting SAM with bounding boxes, we retrieve larger segments.

Scene Description with Spatial Awareness

- Method provides a caption describing the image scene by linking detected objects (Fig. 4).
- Preserves spatial information in the caption about objects in relation to the ego-vehicle and the number of objects.

References:

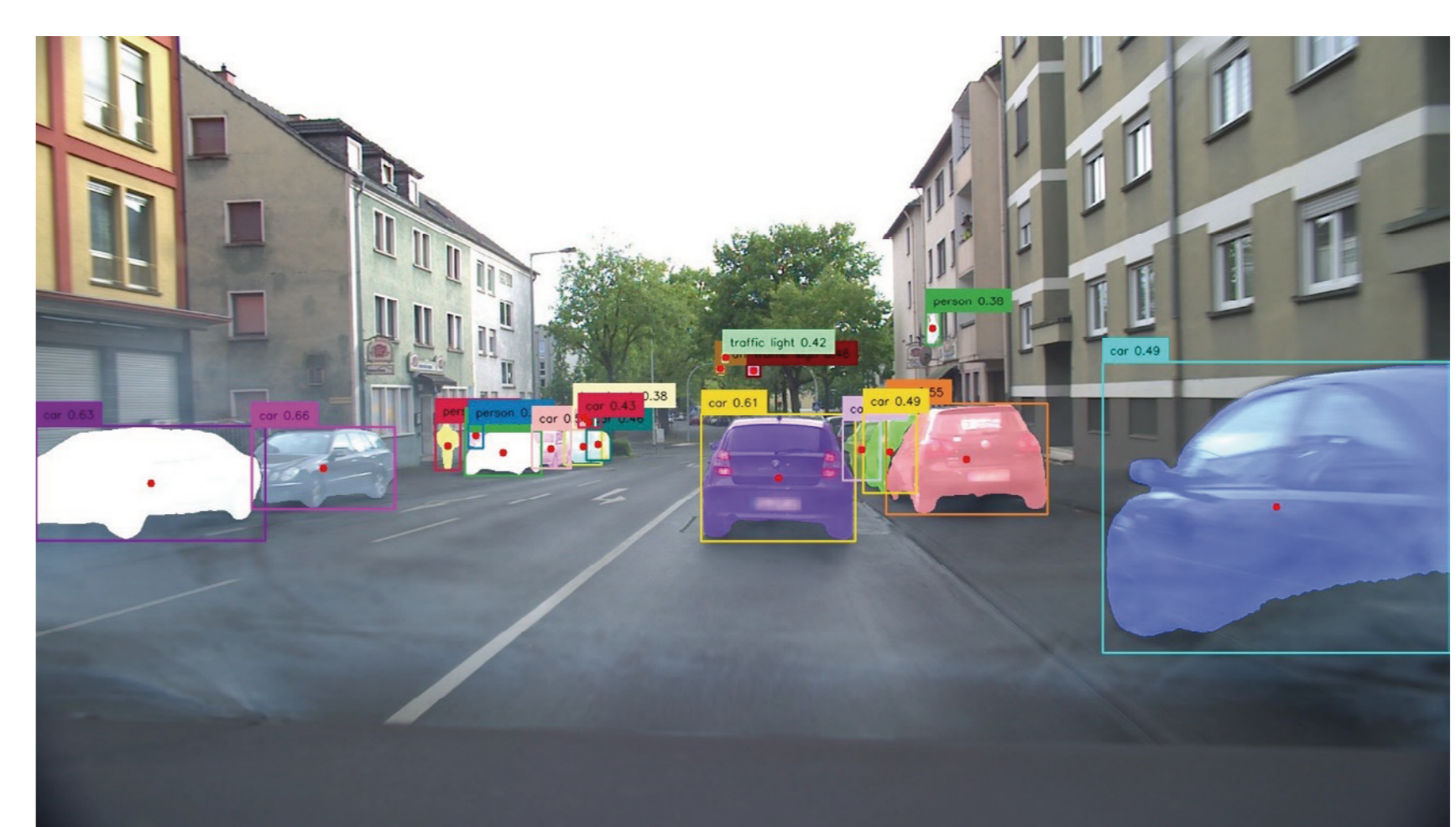
- [1] Liu, Zhijian, et al. "BEVFusion.", ICRA, 2023.
- [2] Kirillov, Alexander, et al. "Segment Anything." arXiv:2304.02643 (2023).
- [3] Liu, Shilong, et al. "Grounding dino." arXiv preprint arXiv:2303.05499 (2023).



["Pedestrians to the left of ego, car in front of Ego, cars on the right of ego, cars at Intersection"]



["Car in front of ego, cars on the right of ego"]



["Car in front of ego, cars on the right of ego, cars on the left of ego"]

Partners



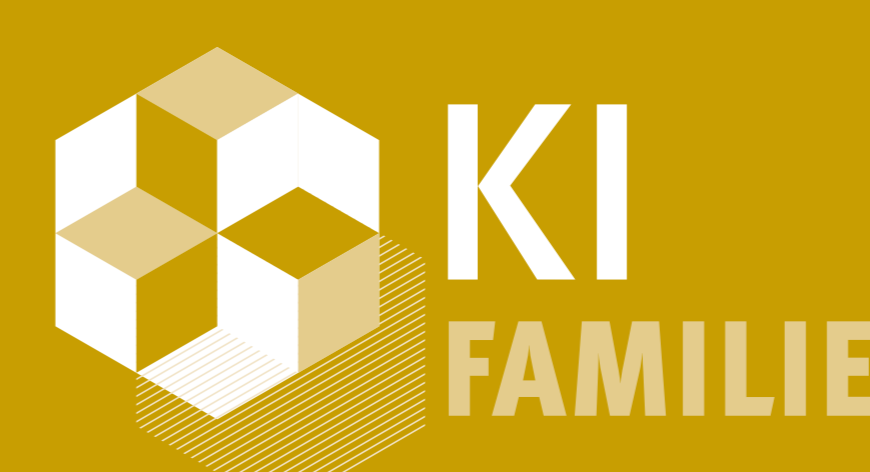
External partners



For more information contact:

rappensperger@fortiss.org
valle@fortiss.org
Naveen-Shankar.Nagaraja@bmw.de

KI Data Tooling 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.



Supported by:



on the basis of a decision by the German Bundestag