



Vehicle Motion Forecasting using Prior Information and Semantic-assisted Occupancy Grid Maps

Rabbia Asghar, Manuel Diaz-Zapata, Lukas Rummelhard, Anne Spalanzani, Christian Laugier

Université Grenoble Alpes, Inria



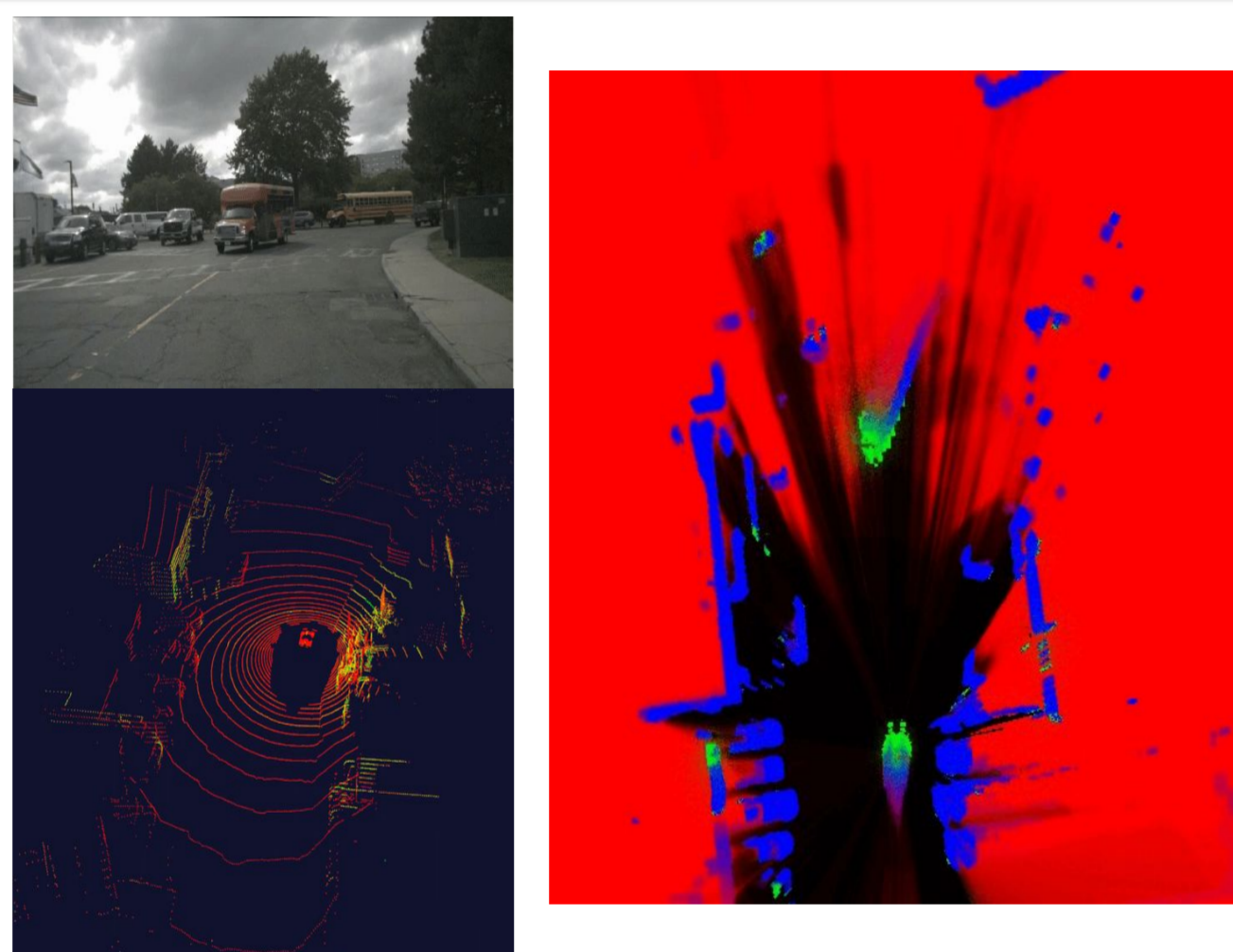
Motivation

Challenges in motion forecasting:

- Uncertainty in sensor data
- Multiple feasible futures
- Complex behavior of agents

What OGMs offer

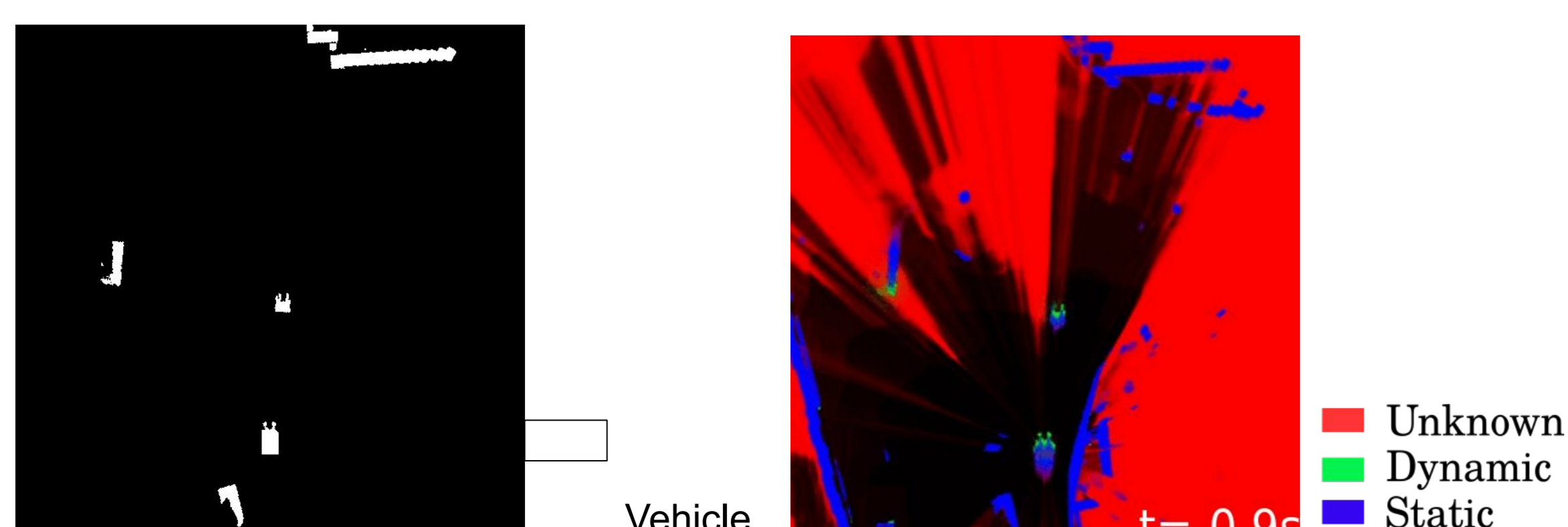
- Bird's-eye-view grid
- Efficient sensor fusion
- Real-time computation
- Models uncertainty
- No confidence thresholding



But, conventional Occupancy Grid Maps (OGMs) predictions struggle with retaining dynamic vehicles in the scene and suffer from subjective choice of evaluation (MSE, IS).

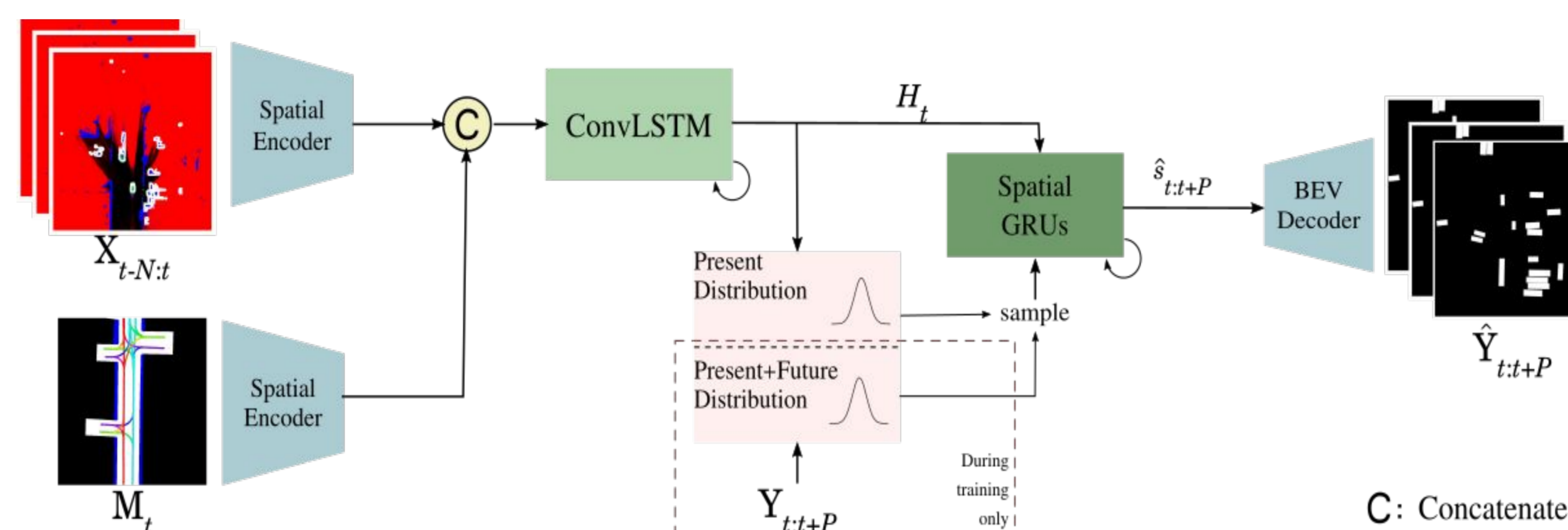
Contribution

- Incorporate contextual information to Dynamic Occupancy Grid Maps that relies on **vehicle semantics** and **map**
- **Predict vehicle semantic grids** that can be evaluated against ground truth annotations.



Vehicle Semantic labels Grid fixed to world frame

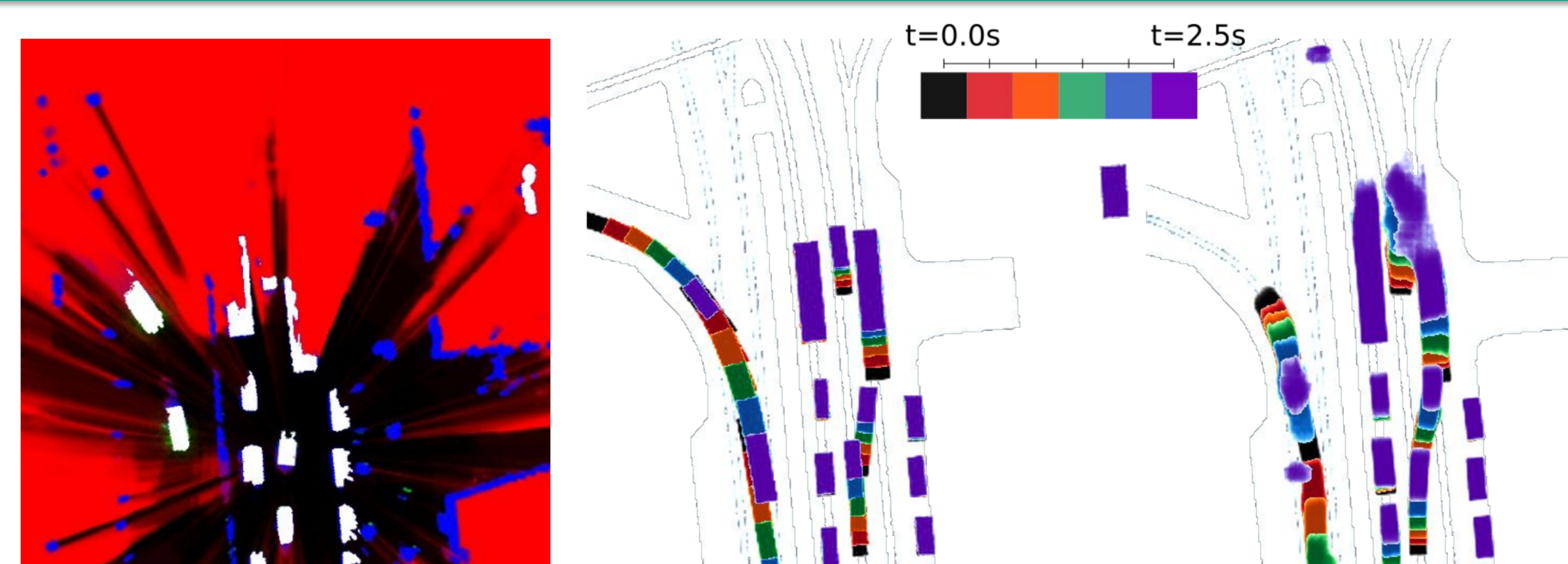
Methods



The prediction architecture takes inspiration from both **video prediction** and **conditional variational approach**.

DOGM with semantic labels and raster map features are fed to spatio-temporal and the probabilistic blocks to capture the evolution of the scene and predict a sequence of future vehicle semantic grids.

Qualitative Results

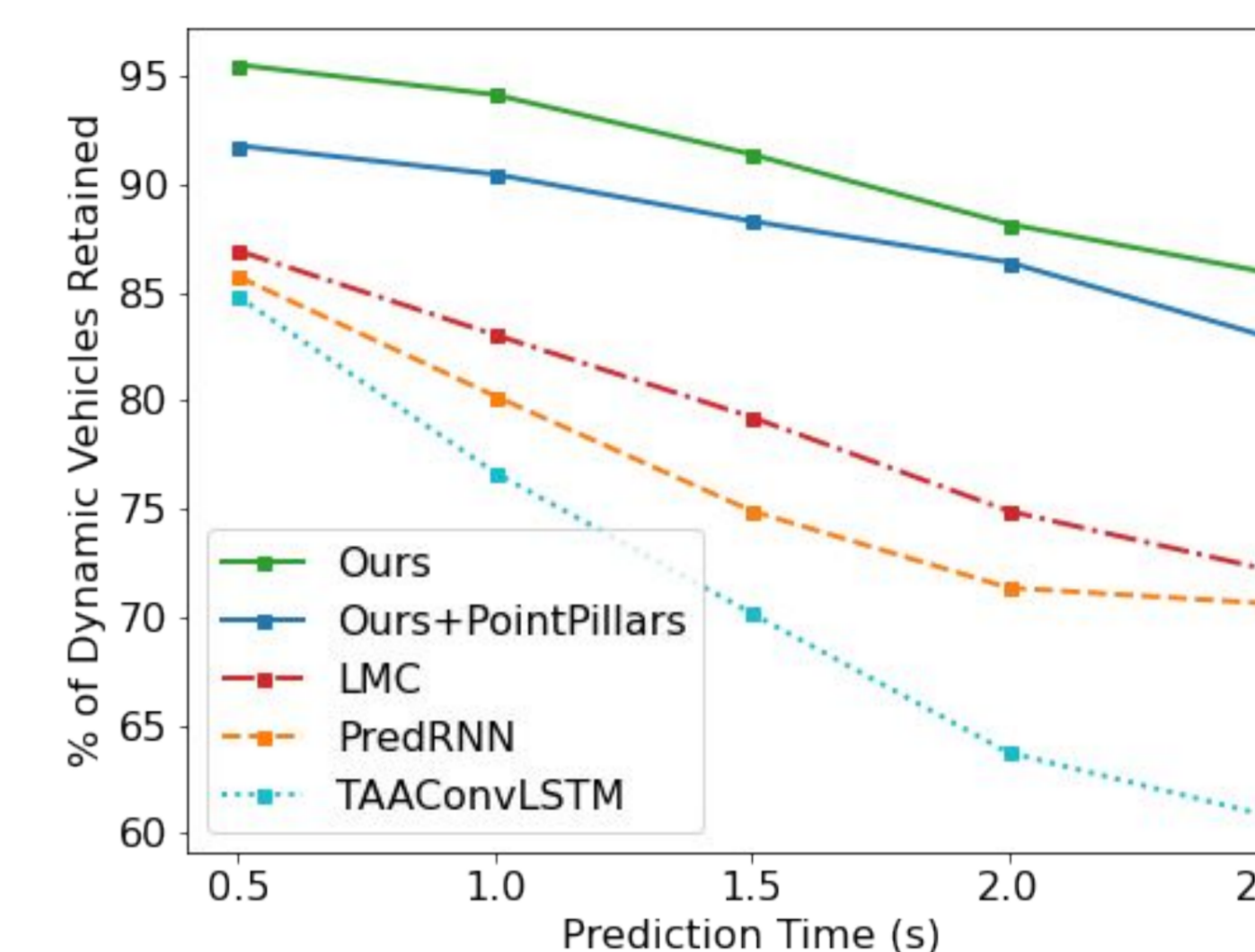


DOGM with semantic labels Input (1s) Prediction Ground Truth

Experiments are conducted on Nuscenes dataset with grid dimension of 60x60m. For training, the grid images are resized to 256x256 pixels. In total, dataset has 3,499 training and 750 validation sequences.



Quantitative Results



Retention of **dynamic vehicles** compared to conventional OGM approaches.

Ablation study on adding information to DOGM. We report mean values of IoU and AUC over 2.5s prediction horizon.

Input	IoU(↑)	AUC(↑)
DOGM	0.336	0.496
DOGM + Map	0.430	0.604
DOGM + Semantics	0.590	0.771
DOGM + Map + Semantics	0.605	0.781

Conclusions

- The addition of semantic labels plays a significant role in improving the prediction results.
- Superior ability to predict both static and dynamic vehicles in the scene.
- For future work, we envision fusion of semantic labels in the DOGM grid generation and prediction of other agents.

References

[1] L. Rummelhard, J.-A. David, A. G. Moreno, and C. Laugier, "A cross-prediction, hidden-state-augmented approach for dynamic occupancy grid filtering," in ICARCV, 2022.
 [2] A. Hu, Z. Murez, N. Mohan, S. Dudas, J. Hawke, V. Badrinarayanan, R. Cipolla, and A. Kendall, "FIERY: Future instance segmentation in bird's-eye view from surround monocular cameras," in ICCV, 2021.
 [3] S. Lee, H. G. Kim, D. H. Choi, H.-I. Kim, and Y. M. Ro, "Video prediction recalling long-term motion context via memory alignment learning," in CVPR, 2021.