Elektrotechnisches Kolloquium

der Bergischen Universität Wuppertal

Die Fakultät für Elektrotechnik, Informationstechnik und Medientechnik lädt zur Teilnahme an folgender Vortragsveranstaltung mit anschließender Diskussion ein:

Es spricht
M.Sc. Harsh Yadav

Lehrstuhl für
Prof. Dr. Tobias Meisen

über das Thema

Open- and Closed-Loop Multimodal Trajectory Prediction from BEV Perspective for Autonomous Driving

Inhalt:

Trajectory prediction forecasts the multimodal future motion of traffic participants and is a central component of autonomous driving systems. Current prediction models are developed and evaluated on open-loop benchmarks, whereas deployment requires them to be inspectable, to scale to crowded scenes, and to remain reliable when their own predictions are executed under receding-horizon replanning. This thesis addresses these three requirements for both formats of the bird's-eye-view (BEV) representation: dense rasterized grids and sparse vectorized tokens.

First, the thesis introduces localized, anchor-conditioned attention that makes the focus of the model controllable and inspectable. On dense BEV grids, CASPFormer grounds deformable cross-attention in explicitly controlled reference points, so that features are sampled only from a local neighborhood of each predicted pose. On vectorized inputs, LMFormer discretizes lanes into fine-grained tokens and caps the range of relative position embeddings, achieving the same spatial restriction at a granularity that identifies which part of a lane influences a prediction.

Second, the thesis establishes $SE(2)$ invariance across the representation spectrum. The algebraic route was already available for vectorized tokens; CASPFormer provides its geometric counterpart for dense grids by rotating the deformable sampling offsets into each agent's local frame, so that a single scene encoding is shared across all agents. LMFormer further replaces two-stage refinement with supervision on intermediate decoder layers, producing refined trajectories in a single pass.

Third, the thesis proposes a closed-loop training paradigm in which the model is rolled out on its own predictions and each drifted prediction is supervised against the recorded future. Mode-consistent on-policy rollouts retain multimodal prediction accuracy, and detaching gradients across state transitions reduces the coupling between the model and the replanning rate it was trained at. At the highest replanning rate, collision rates decrease by 22% on nuScenes, 35% on DeepScenario, and 17% on Argoverse~2, while single-shot accuracy is largely retained. The paradigm transfers unchanged to QCNet, and a decoder-only variant of LMFormer with reduced parameters by 45% and floating-point operations by 62% to make high-rate rollouts affordable.

A critical discussion delimits these results. Attention supports controllability and inspection but not causal explanation, and closed-loop robustness is demonstrated against non-reactive background traffic, since goal-conditioned scene prediction did not yield surrounding agents that react to the ego vehicle. Reactive simulation, planner-in-the-loop evaluation, and integration with learned perception are identified as the main directions for future work.

T e r m i n:	**18.11.2026, 14:00 Uhr**
O r t:	Bergische Universität Wuppertal Campus Freudenberg, Seminarraum FG 1.01