NEAT: Neural Attention Fields for End-to-End Autonomous Driving

Kashyap Chitta*1,2, Aditya Prakash*1, Andreas Geiger1,2

1Max Planck Institute for Intelligent Systems, Tübingen 2University of Tübingen

Bird’s Eye View
- Orthographic projection of 3D world
- No occlusions
- Better correlated with vehicle kinematics

Our Method
Implicit Bird’s Eye View Semantics from cameras

- Compact
- Arbitrary spatial and temporal resolution
- Fixed memory footprint
- Can use sparse supervision

However, associating 2D to BEV semantics requires geometry, scene motion, ego motion, agent intentions and agent interactions

Interpolating Predictions

Attention Map Visualizations

Learned attention for highlighted BEV locations (white circle on semantic map). We consistently attend to the object of interest (top left to bottom right: bicyclist, green light, vehicle, red light)

Safe Driving on CARLA Leaderboard

Check out our results video: http://bit.ly/neat_video
kashyap.chitta@tue.mpg.de

(from https://leaderboard.carla.org/ July 2021)