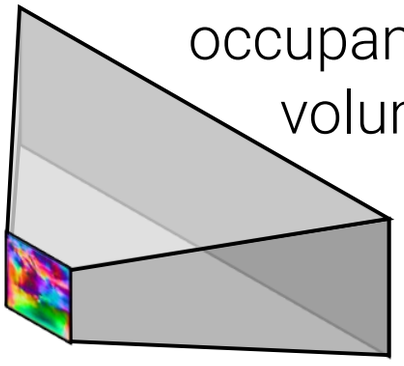
1. Bird's-Eye View inference

image

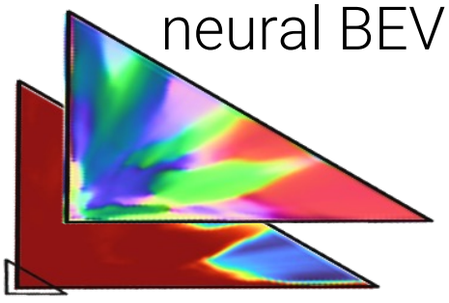
CNN

occupancy volume



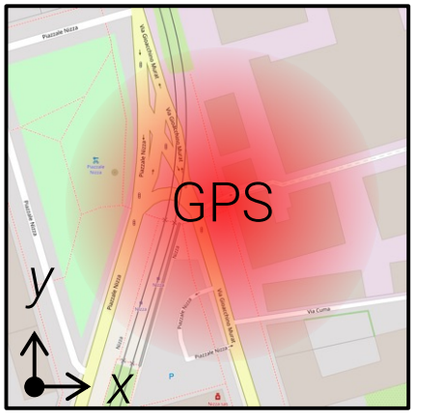
pool + CNN

neural BEV



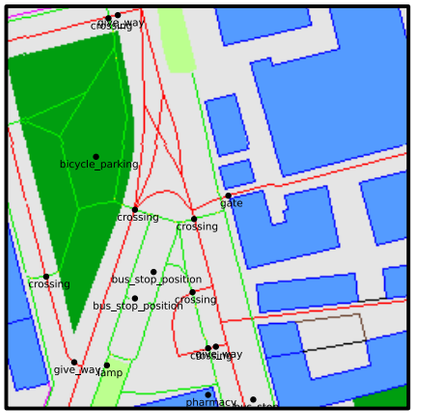
confidence

2. Map encoding



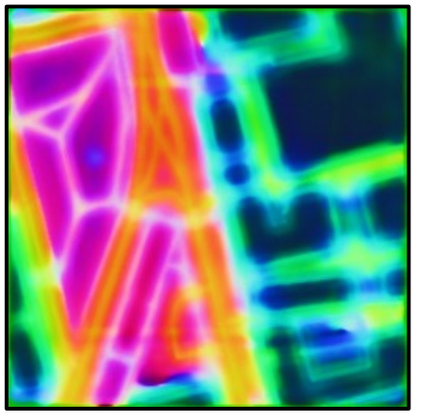
OpenStreetMap

Rasterize



raster map

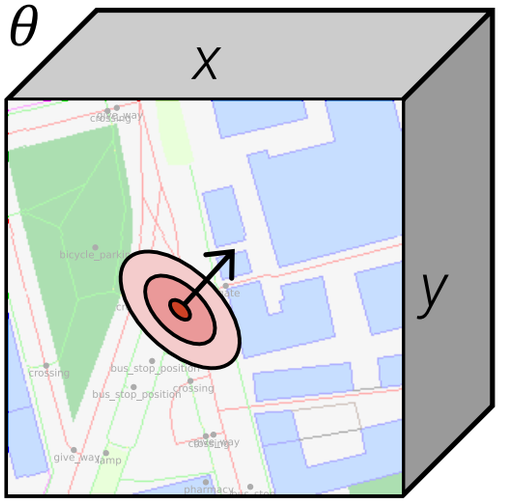
CNN



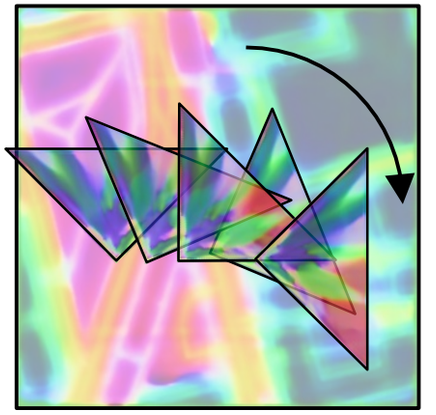
neural map

3-DoF pose (x, y, θ)

argmax



pose likelihood



rotate + correlate

3. BEV-map matching