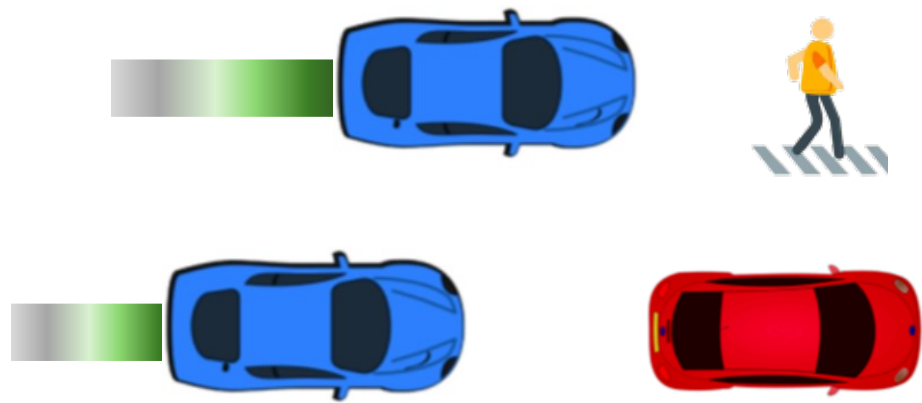
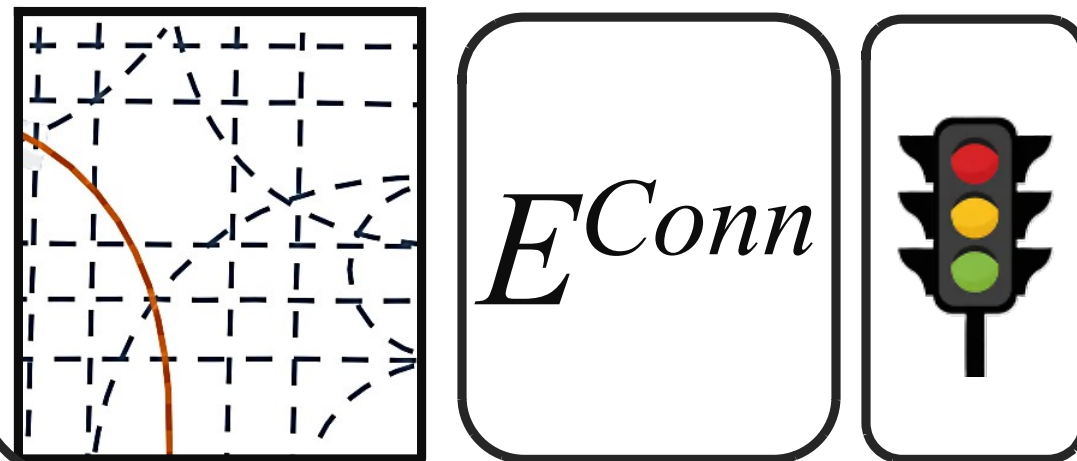


Agent elements



Lane elements



Motion-Aware Agent Encoder

Static branch

h_{st}

Gating $g = \sigma(\cdot)$

g

Motion branch

h_{mot}

$(1 - g) \odot h_{st}$

Fused embedding

$g \odot h_{mot}$

Lane Encoder

Lane embed

Lane conn embed

Learnable lane-count query

Factorized Scene Transformer

N_L

L2L

L2A

A2A

Lane Latent

Agent Latent

Gaussian Latent Heads

$$z = \mu(l) + \exp\left(\frac{1}{2} \log \sigma^2(l)\right) \odot \epsilon$$