

Main Model

$$g_{EncTrans}(\bar{e}_\eta, \psi_{cat})$$

"Turn yourself so that you are going with the flow of traffic. There should be a purple theater banner on your left. Go forward on this street until you come to the first traffic light. Make a right at the light. You should see silver gates on your left. Go straight and when you come to the next light, make a right. Very shortly after you make the turn, look to the left and you should see a silver light post with a bike leaning against it."



Cross-modal encoding

$$g_{Attn}(e_l, e'_v)$$



Turn yourself so that you are going with the flow of traffic. There should be a purple theater banner on your left.

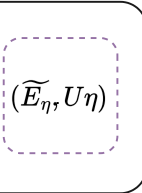
...



Very shortly after you make the turn, look to the left and you should see a silver light post with a bike leaning against it.

Cross-modal attention

$$g_{CFNs}(\tilde{E}_\eta, U_\eta)$$

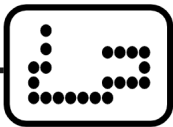


Combine outputs and select action

$$g_{Clas_{max}}(y_j, z')$$

{Forward,
Left,
Right,
Stop}

$$g_{PMTTP}(tr)$$



$$Normal(\mu_j, \sigma_j, x_{left}, x_{right})$$

Trajectory estimation



Start



...



End

Turn yourself so that you are going with the flow of traffic. There should be a purple theater banner on your left. Go forward on this street until you come to the first traffic light. Make a right at the light. You should see silver gates on your left. Go straight and when you come to the next light, make a right. Very shortly after you make the turn, look to the left and you should see a silver light post with a bike leaning against it.

$$g_{PMF}(v'_t, \psi_t)$$



"You should see silver gates on your left. Go straight and when you come to the next light, make a right. Very shortly after you make the turn, look to the left and you should see a silver light post with a bike leaning against it."



You should see silver gates on your left. Go straight and when you come to the next light, make a right. Very shortly after you make the turn, look to the left and you should see a silver light post with a bike leaning against it."

Feature-level prioritisation

PM-VLN

FL_{PM} Framework

