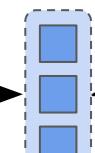
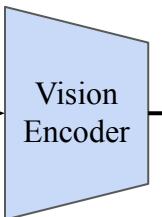
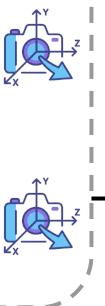




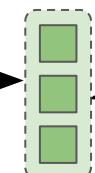
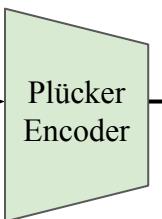
Question: If I am standing in front of the door. Is the trash bin on my right side?

Input views & Camera Poses

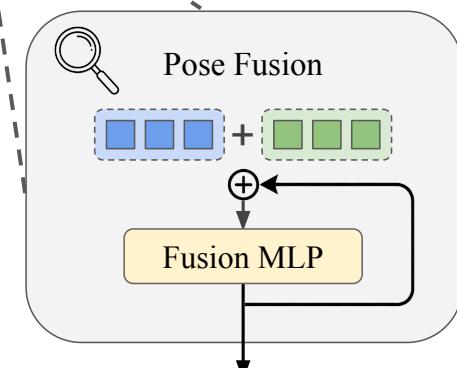


MLM

Image
Tokens



Camera
Tokens



Yes, the trash bin is on your right.

$\mathcal{L}_{\text{lang}}$

Ground Truth Answer



Pose Adapter

Predicted Camera Pose

$\mathcal{L}_{\text{pose}}$

Ground Truth Camera Pose

Image Decoder

Imagined Target View