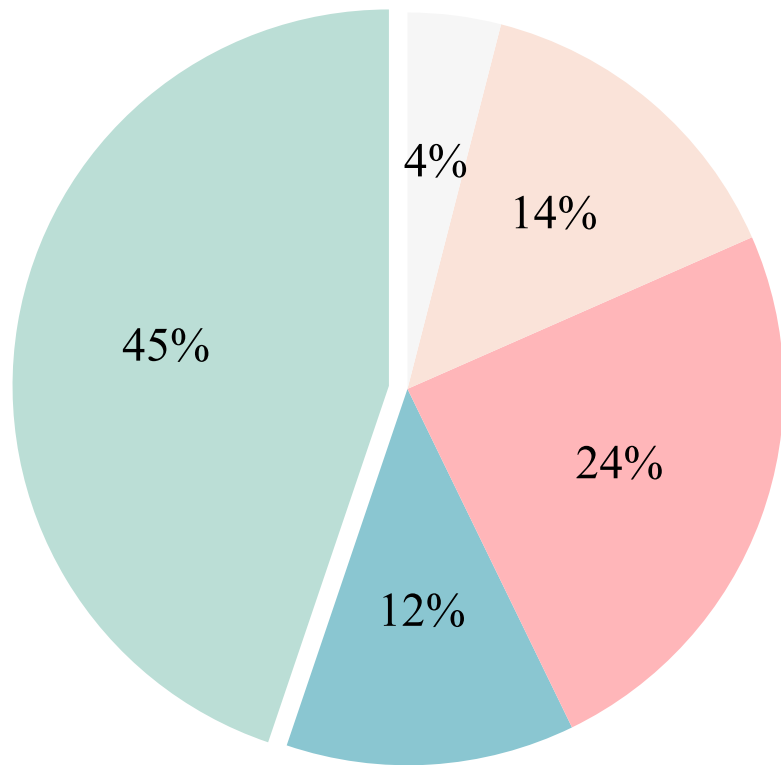
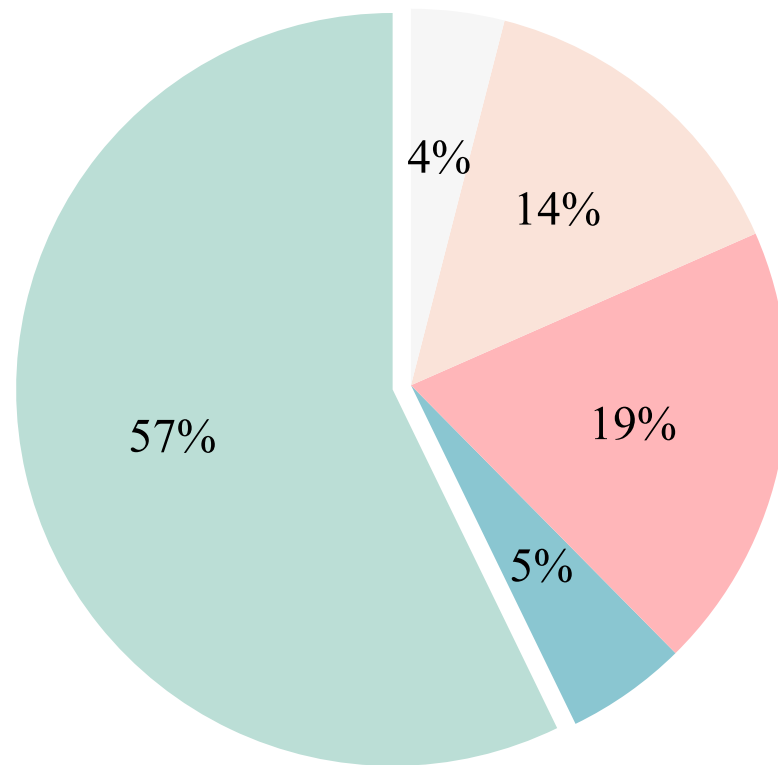


Ours (3D only)



Ours (3D+2D)



Correct Semantic Spatial Detection Other