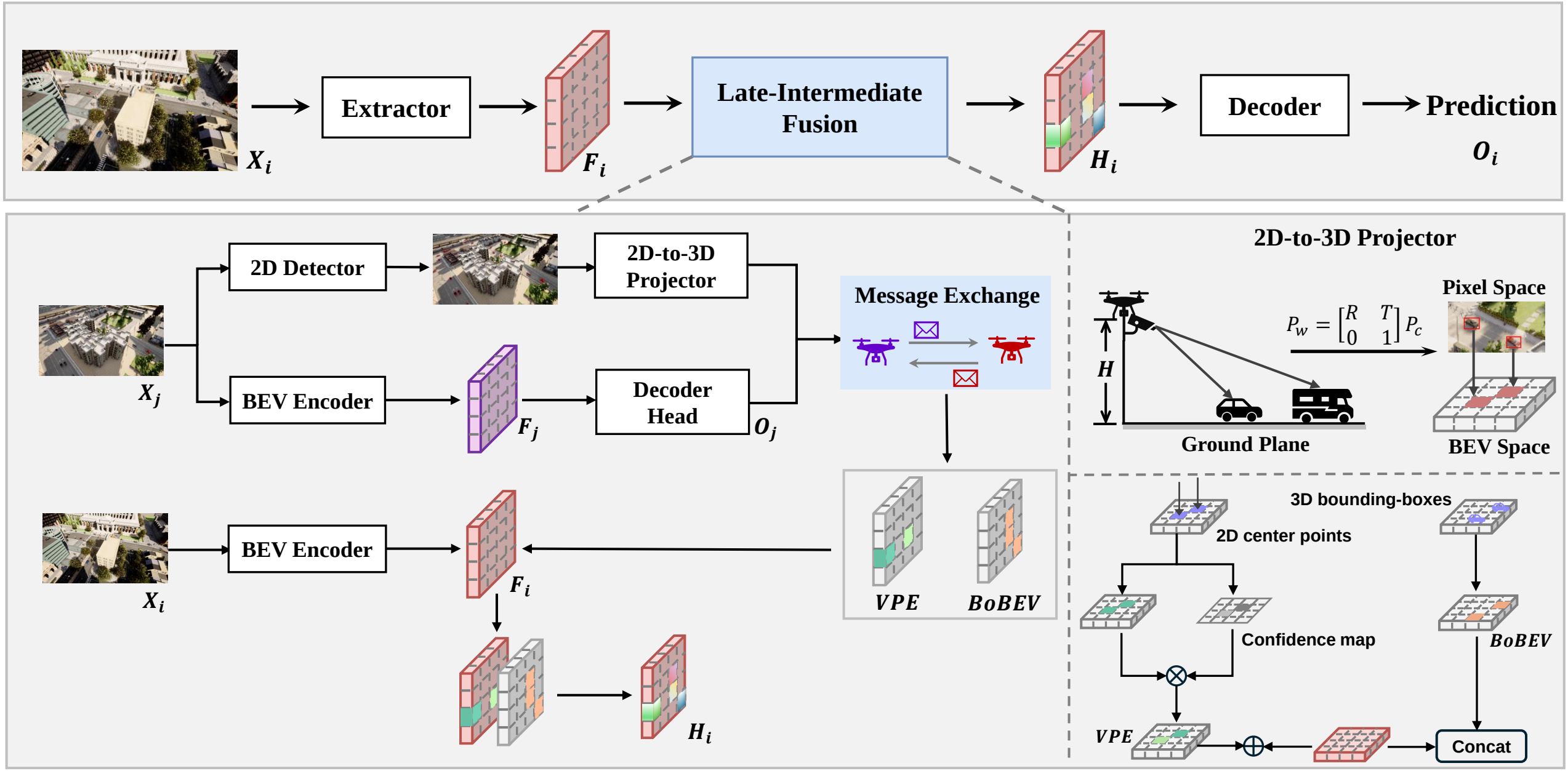

X_i Observation F_i BEV feature map H_i Fused feature map O_i Detection VPE Vision-guided positional embedding $BoBEV$ Augmented feature