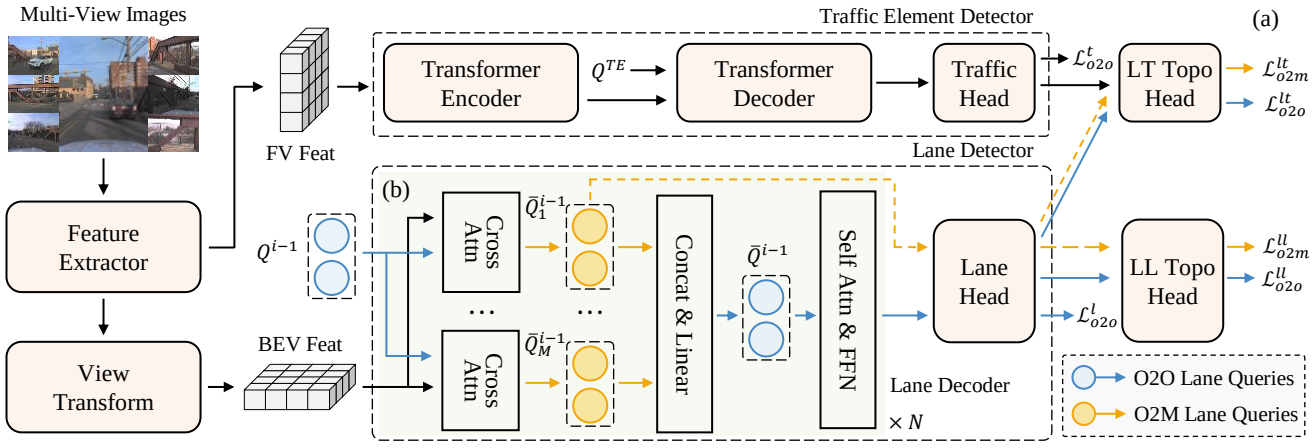
Multi-View Images
Feature Extractor
View Transform
FV Feat
BEV Feat
Traffic Element Detector
(a)
Transformer Encoder
Q^{TE}
Transformer Decoder
Traffic Head
$\mathcal{L}^{t}_{o2o}$
LT Topo Head
$\mathcal{L}^{lt}_{o2m}$
$\mathcal{L}^{lt}_{o2o}$
Lane Detector
(b)
Q^{i-1}
Cross Attn
$\bar{Q}^{i-1}_{1}$
$\bar{Q}^{i-1}_{M}$
Concat & Linear
$\bar{Q}^{i-1}$
Self Attn & FFN
Lane Head
$\mathcal{L}^{l}_{o2o}$
LL Topo Head
$\mathcal{L}^{ll}_{o2m}$
$\mathcal{L}^{ll}_{o2o}$
Lane Decoder
$\times N$
O2O Lane Queries
O2M Lane Queries