

1 Multimodal Chooser

Multimodal Preprocess

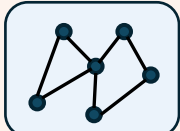
Multimodal Knowledge Graph



Triplet Example

<Zebra, take in, grass>

S-Mod



V-Mod



T-Mod

Zebra is a mammal native to Africa.

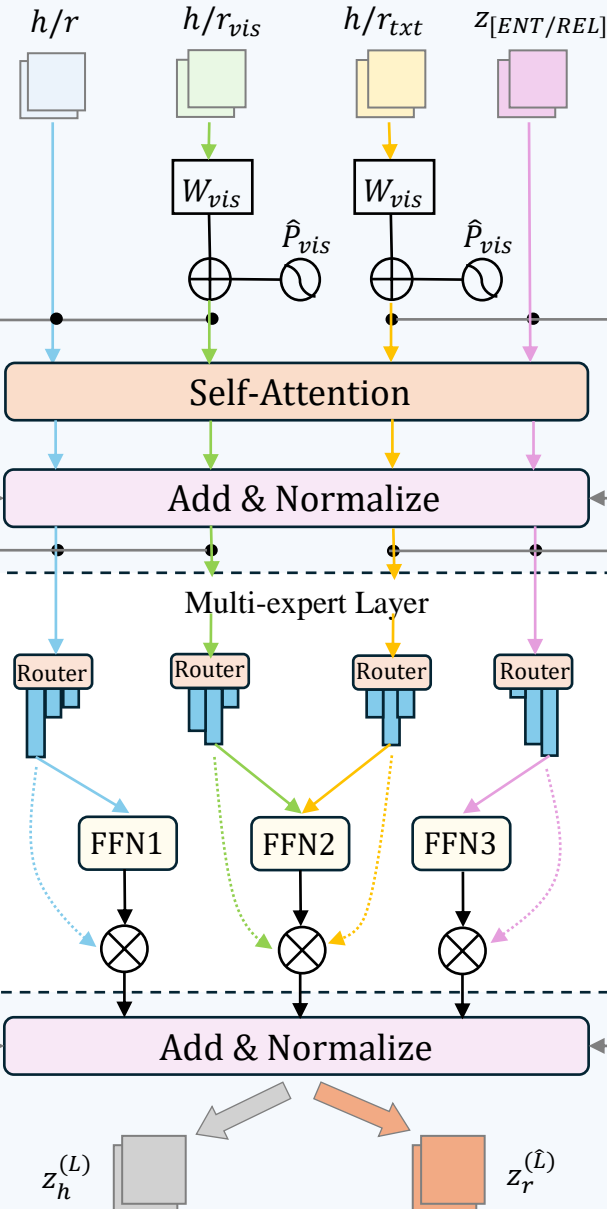
ViT_{BASE}

h/r_{vis}

$BERT_{BASE}$

h/r_{txt}

Multimodal Entity/Relation Encoder



2 Geometric Chooser

