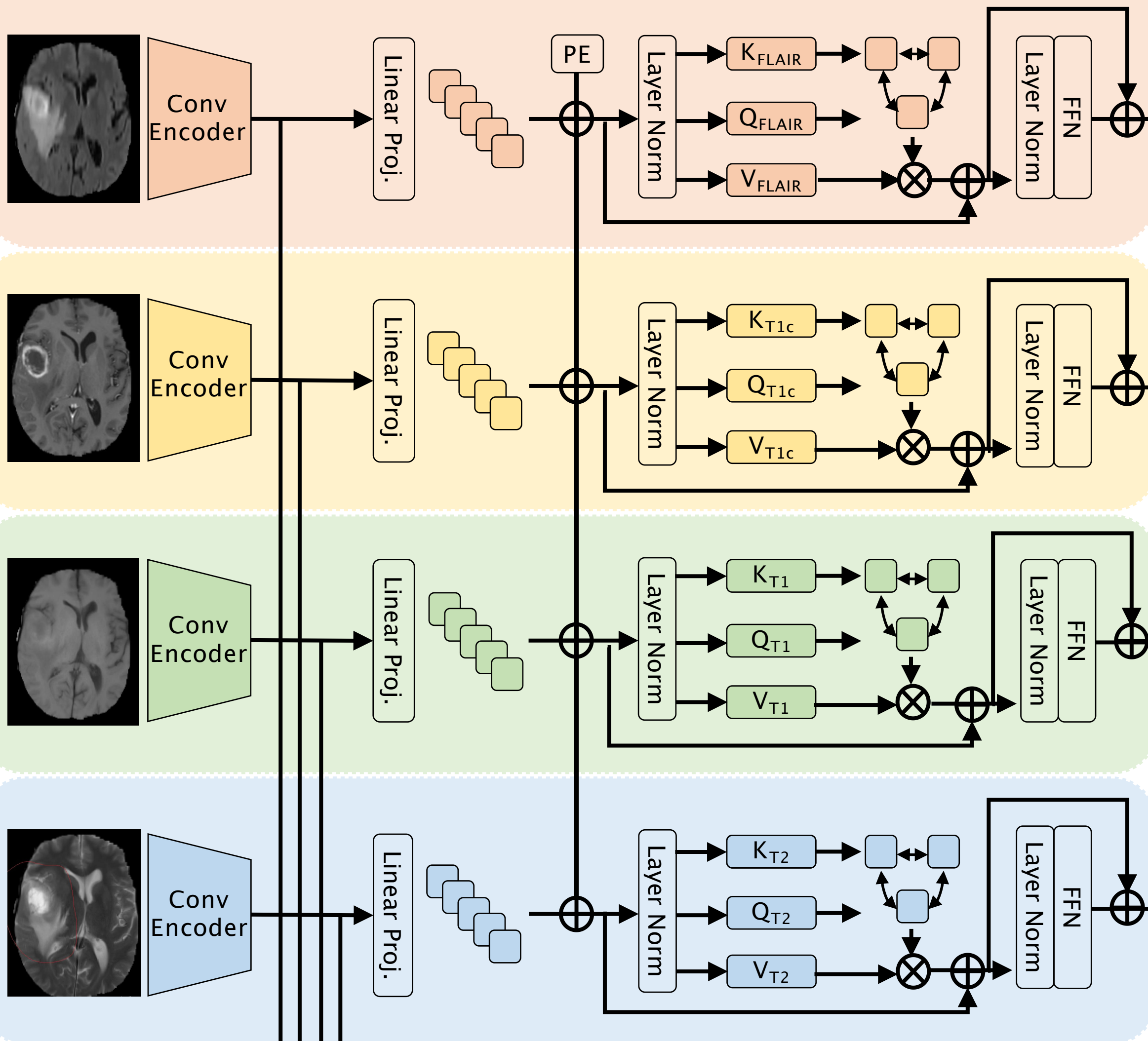
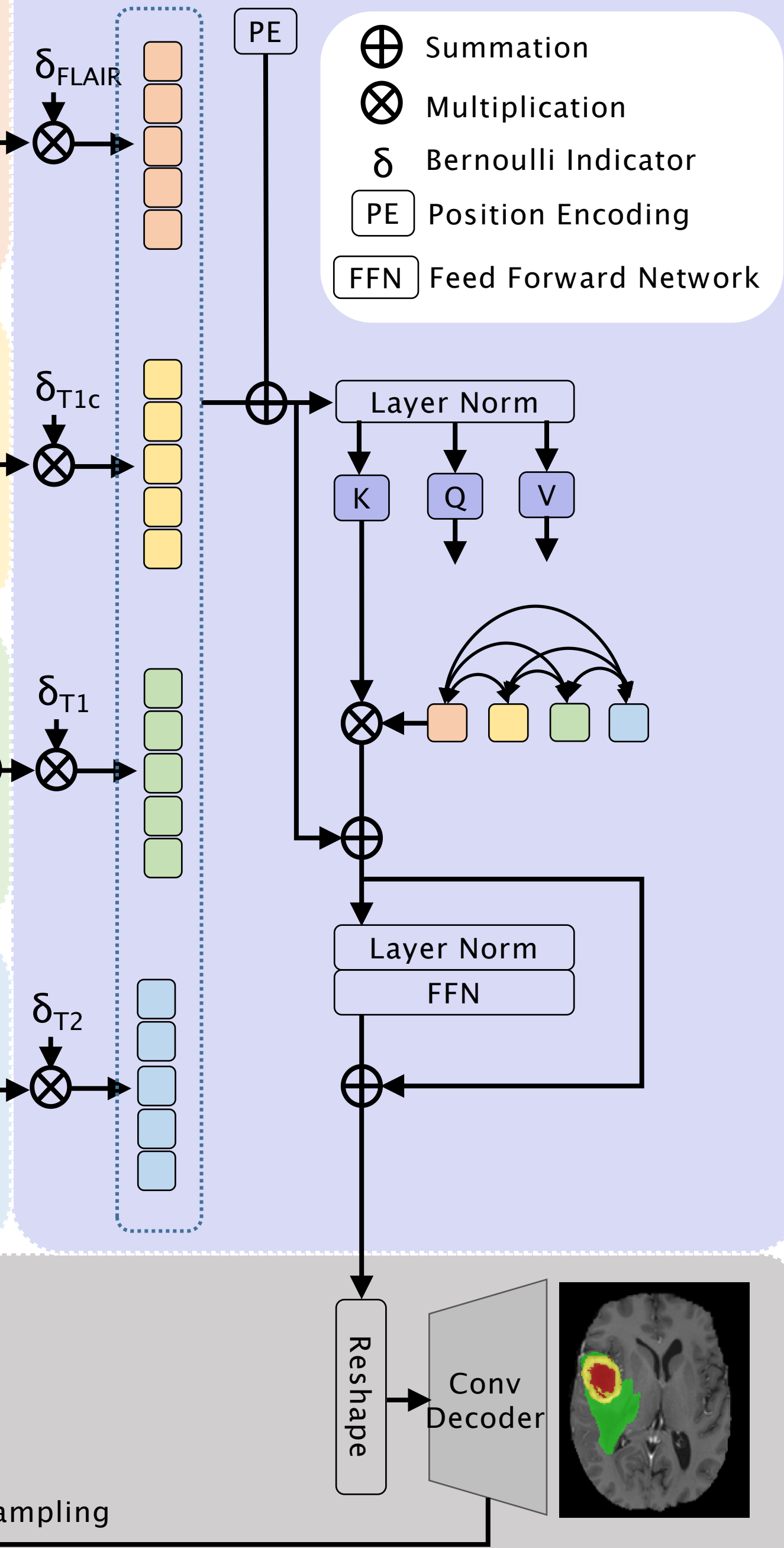


## Hybrid Modality-specific Encoder



## Modality-correlated Encoder



Upsampling

Auxiliary  
Regularizers

Upsampling

Conv  
Decoder