

LayerNorm

MHA
($h=8, d_{model} = 768$)

\oplus

LayerNorm

FFN
($d_{ff} = 3072$)

\oplus