

Patch Embeddings + Positional Encodings



Multi-Head Self-Attention



Add & Norm



FFN (MLP: GELU)



Add & Norm