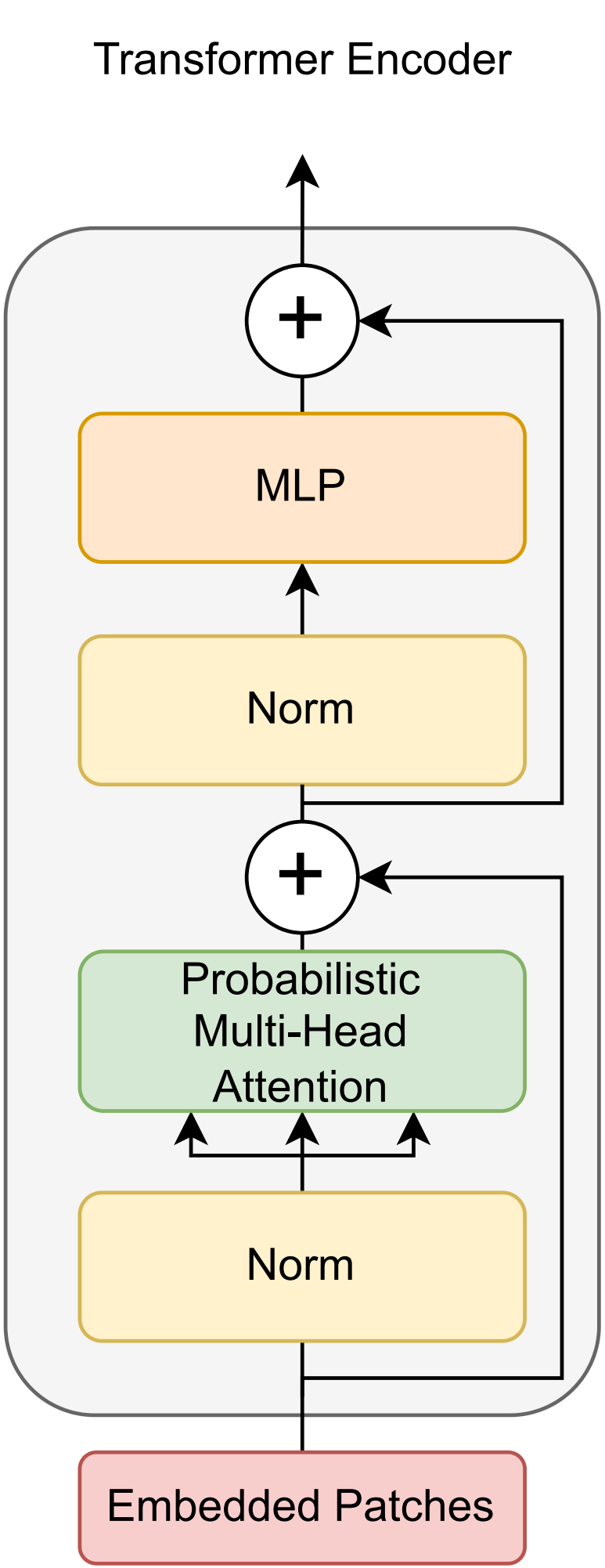
Transformer Encoder
MLP
Norm
Probabilistic
Multi-Head
Attention
Norm
Embedded Patches