Hypernetwork 🔥

1. Record Memory hidden states at each layer

Layer N Output	m(1)	m(2)	......	m(M)
......	......	......	......	......
Layer 3 Output	m(1)	m(2)	......	m(M)
Layer 2 Output	m(1)	m(2)	......	m(M)
Layer 1 Output	m(1)	m(2)	......	m(M)

2. Add Positional Embedding

Memory: (B, L, M, H)

ADD

Layer Positional Embedding: (L, 1, H)
Token Positional Embedding: (1, M, H)

Repeat

m(1)	m(2)	......	m(M)
......	......	......	......
m(1)	m(2)	......	m(M)
m(1)	m(2)	......	m(M)
m(1)	m(2)	......	m(M)

3. Bidirectional Transformer

Memory: (B, i, M, H)
Layer i Generated LoRA

Layer N Generated LoRA
......
Layer 3 Generated LoRA
Layer 2 Generated LoRA
Layer 1 Generated LoRA

4. Reshape Into LoRA Weights