

Convolution-part for batch size 4, D=768, 36 layers, L=2¹⁴

