

Convolution-part for batch size 4, D=768, 18 layers, L=2¹⁵

