

Convolution-part for batch size 4, D=256, 18 layers, L=2¹⁶

