

Convolution-part for batch size 8, D=256, 18 layers, L=2¹⁵

