

Convolution-part for batch size 2, D=256, 18 layers, L=2¹⁷

