

Convolution-part for batch size 8, D=768, 18 layers, L= 2^{14}

