

Convolution-part for batch size 1, D=768, 36 layers, L= 2^{16}

