

Convolution-part for batch size 2, D=768, 36 layers, L=2¹⁵

