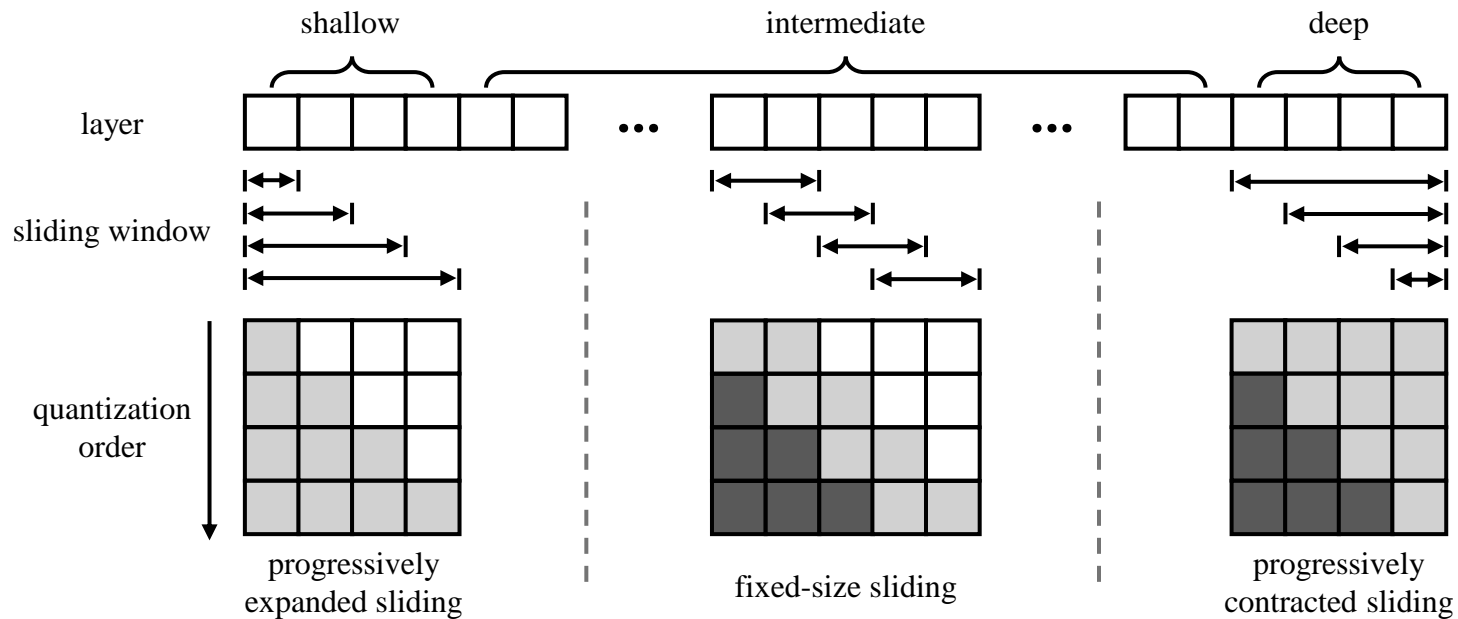


Inter-layer sliding quantization



Quantized



Quantizing



Unquantized



Weights/Activations

Intra-layer sliding quantization

