Speedup of GEMM on GTX3090 (INT8)

