

Inference Throughput by Quantization Method

