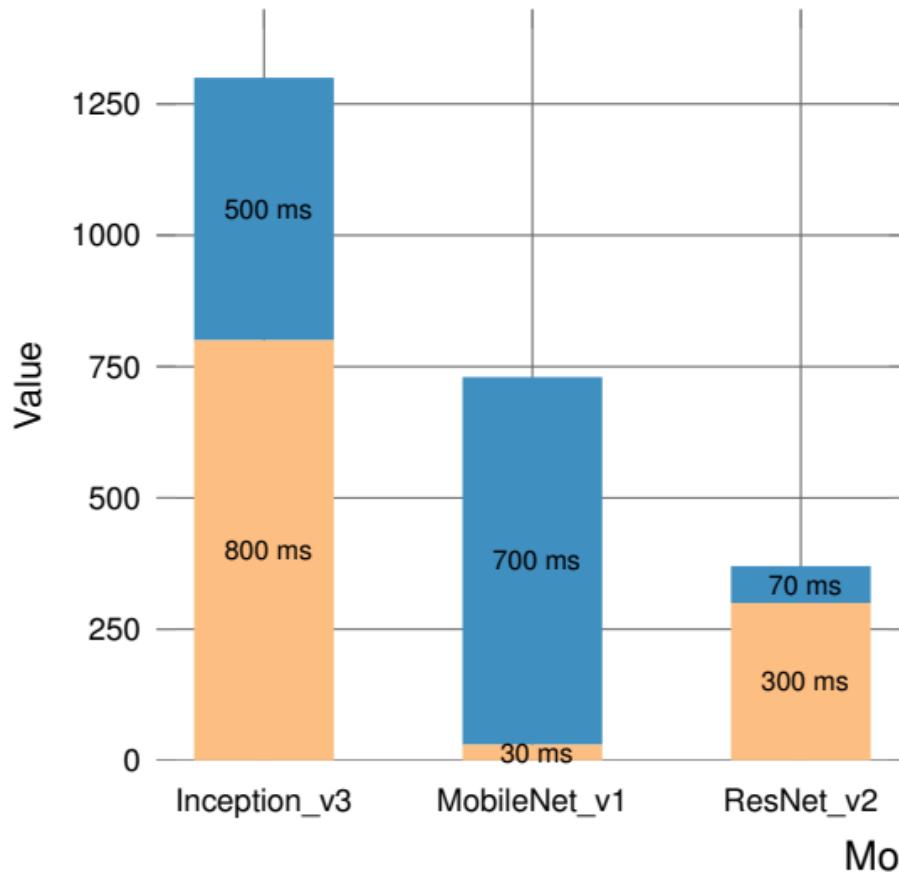
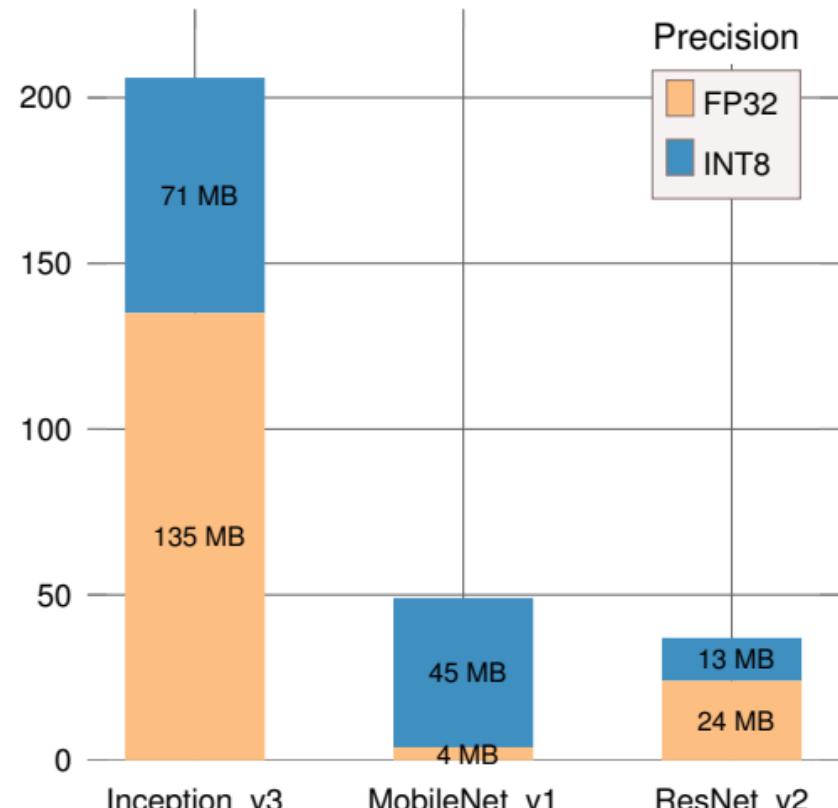


# Impact of Quantization on Inference Time and Model Size

Inference\_Time



Model\_Size



Model

Precision

- FP32
- INT8