

Inference Latency vs. Network Bandwidth (ViT-small)

