Scaling Analysis: Architecture Comparison

(a)

(b)

Validation Loss - Irreducible (Log Scale)

Compute (FLOPs)