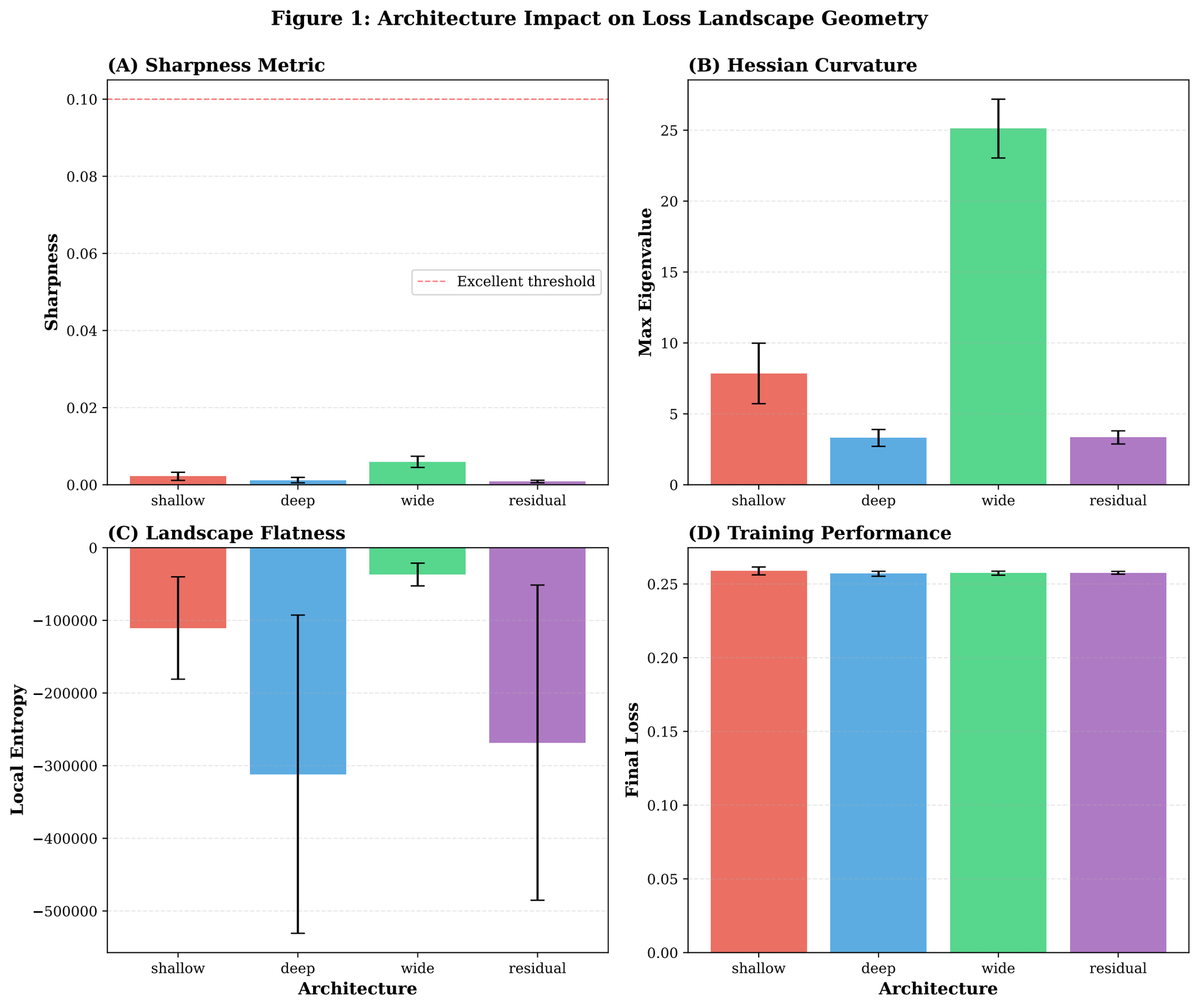
Figure 1: Architecture Impact on Loss Landscape Geometry
(A) Sharpness Metric
Sharpness
0.10
0.08
0.06
0.04
0.02
0.00
Excellent threshold
shallow
deep
wide
residual
(B) Hessian Curvature
Max Eigenvalue
25
20
15
10
5
0
shallow
deep
wide
residual
(C) Landscape Flatness
Local Entropy
0
−100000
−200000
−300000
−400000
−500000
shallow
deep
wide
residual
Architecture
(D) Training Performance
Final Loss
0.25
0.20
0.15
0.10
0.05
0.00
shallow
deep
wide
residual
Architecture