Model Comparison Analysis

Per-Sample Performance Comparison

InteractiveRL Dice Score vs U-Net Dice Score

Better InteractiveRL

Better U-Net

Dice Difference (InteractiveRL - U-Net)

Legend:
- Equal Performance
- U-Net Better
- InteractiveRL Better
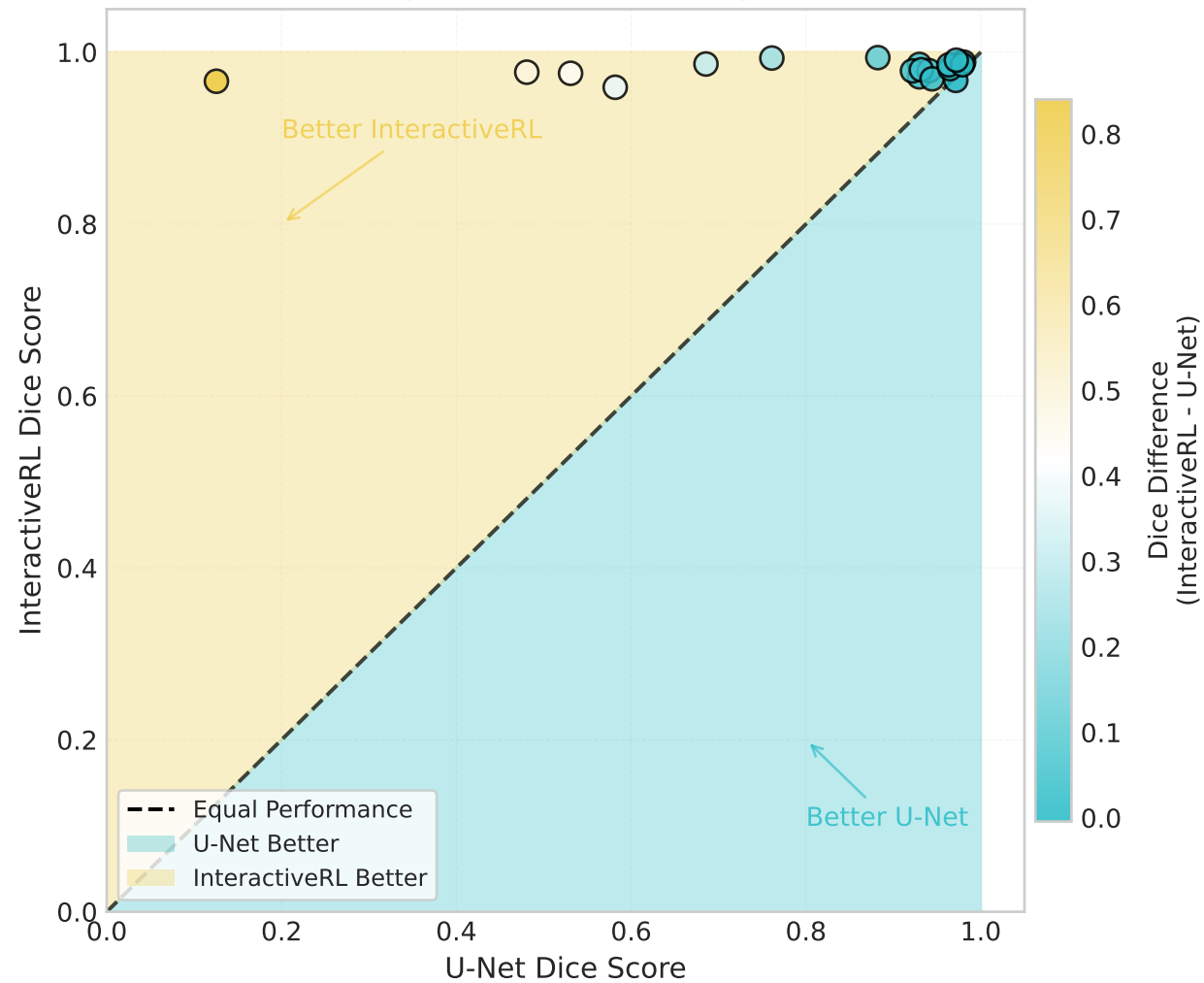
Distribution of Performance Differences

U-Net better: 0 (0.0%)
Similar: 6 (30.0%)
InteractiveRL better: 14 (70.0%)

Number of Samples vs InteractiveRL - U-Net Dice Difference