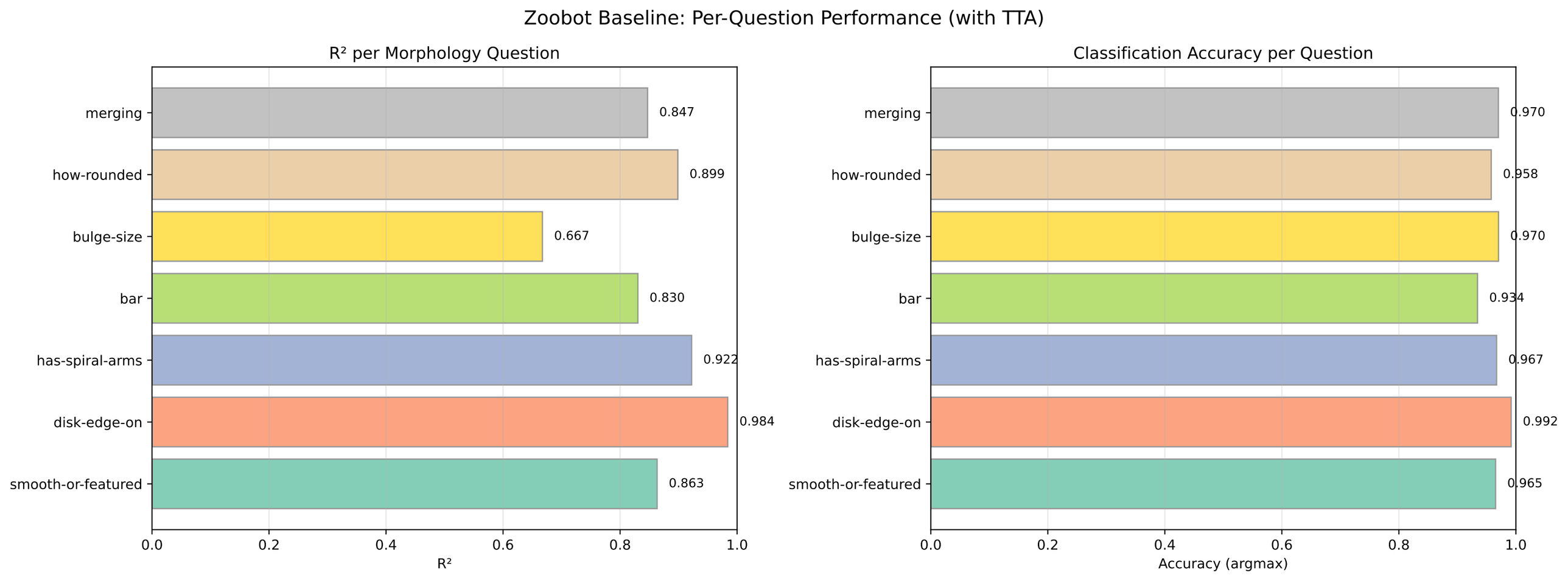
Zoobot Baseline: Per-Question Performance (with TTA)
R² per Morphology Question
merging
how-rounded
bulge-size
bar
has-spiral-arms
disk-edge-on
smooth-or-featured
0.847
0.899
0.667
0.830
0.922
0.984
0.863
0.0
0.2
0.4
0.6
0.8
1.0
R²
Classification Accuracy per Question
merging
how-rounded
bulge-size
bar
has-spiral-arms
disk-edge-on
smooth-or-featured
0.970
0.958
0.970
0.934
0.967
0.992
0.965
0.0
0.2
0.4
0.6
0.8
1.0
Accuracy (argmax)