

# Generalisation report

## Generalisation report produced by model-vs-human

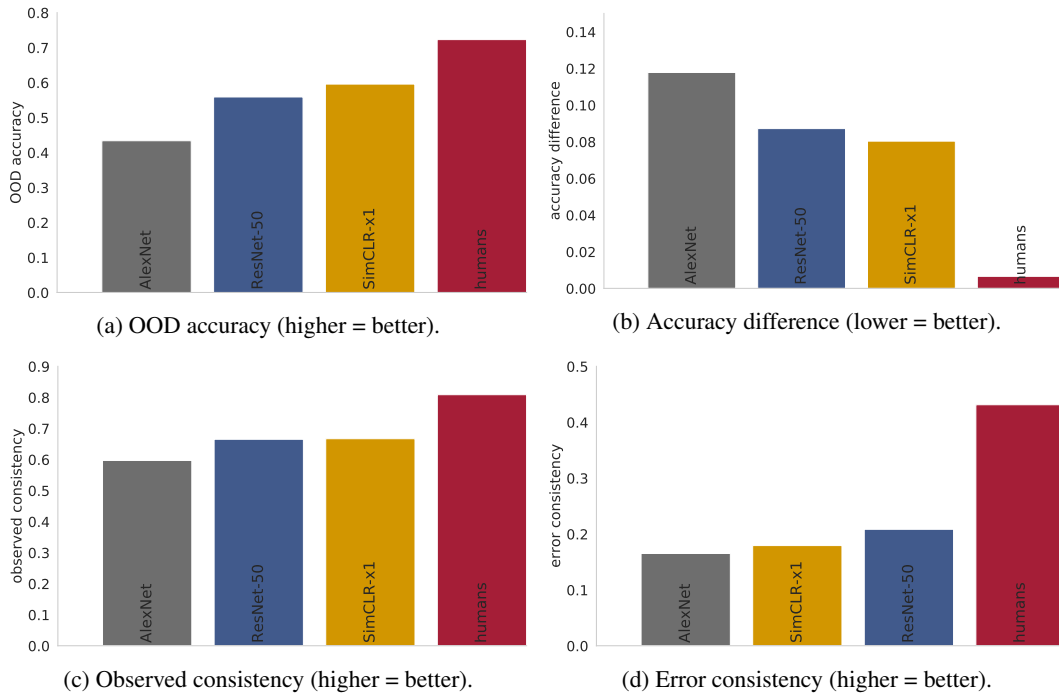


Figure 1: Benchmark results for different models, aggregated over datasets.

Table 1: Benchmark table of model results for most human-like behaviour. The three metrics “accuracy difference” “observed consistency” and “error consistency” (plotted in Figure 1) each produce a different model ranking. The mean rank of a model across those three metrics is used to rank the models on our benchmark.

model	accuracy diff. ↓	obs. consistency ↑	error consistency ↑	mean rank ↓
SimCLR-x1	<b>0.080</b>	<b>0.667</b>	0.179	<b>1.333</b>
ResNet-50	0.087	0.665	<b>0.208</b>	1.667
AlexNet	0.118	0.597	0.165	3.000

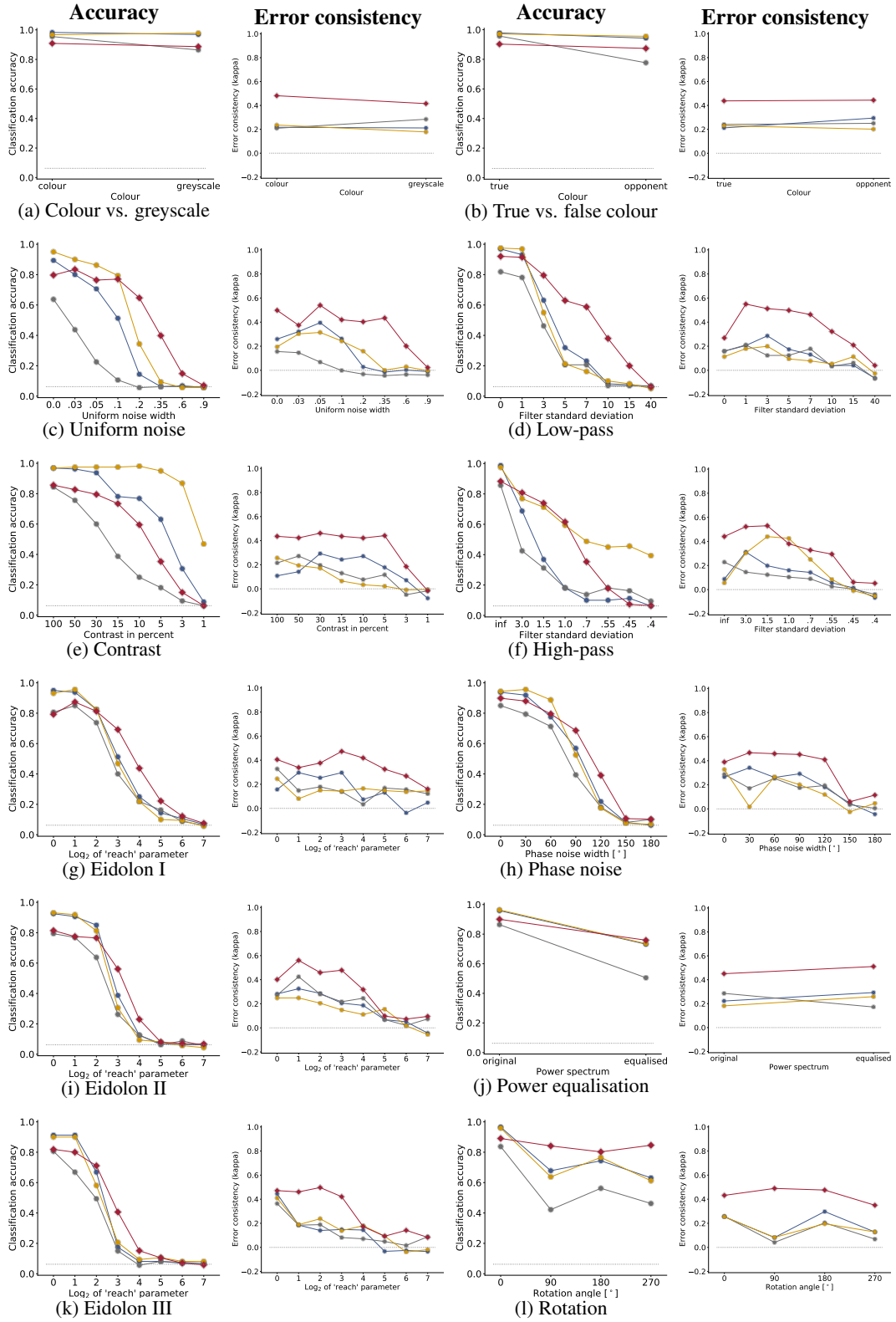


Figure 2: OOD accuracy and error consistency.

Table 2: Benchmark table of model results for highest out-of-distribution robustness.

model	OOD accuracy $\uparrow$	rank $\downarrow$
SimCLR-x1	<b>0.596</b>	<b>1.000</b>
ResNet-50	0.559	2.000
AlexNet	0.434	3.000

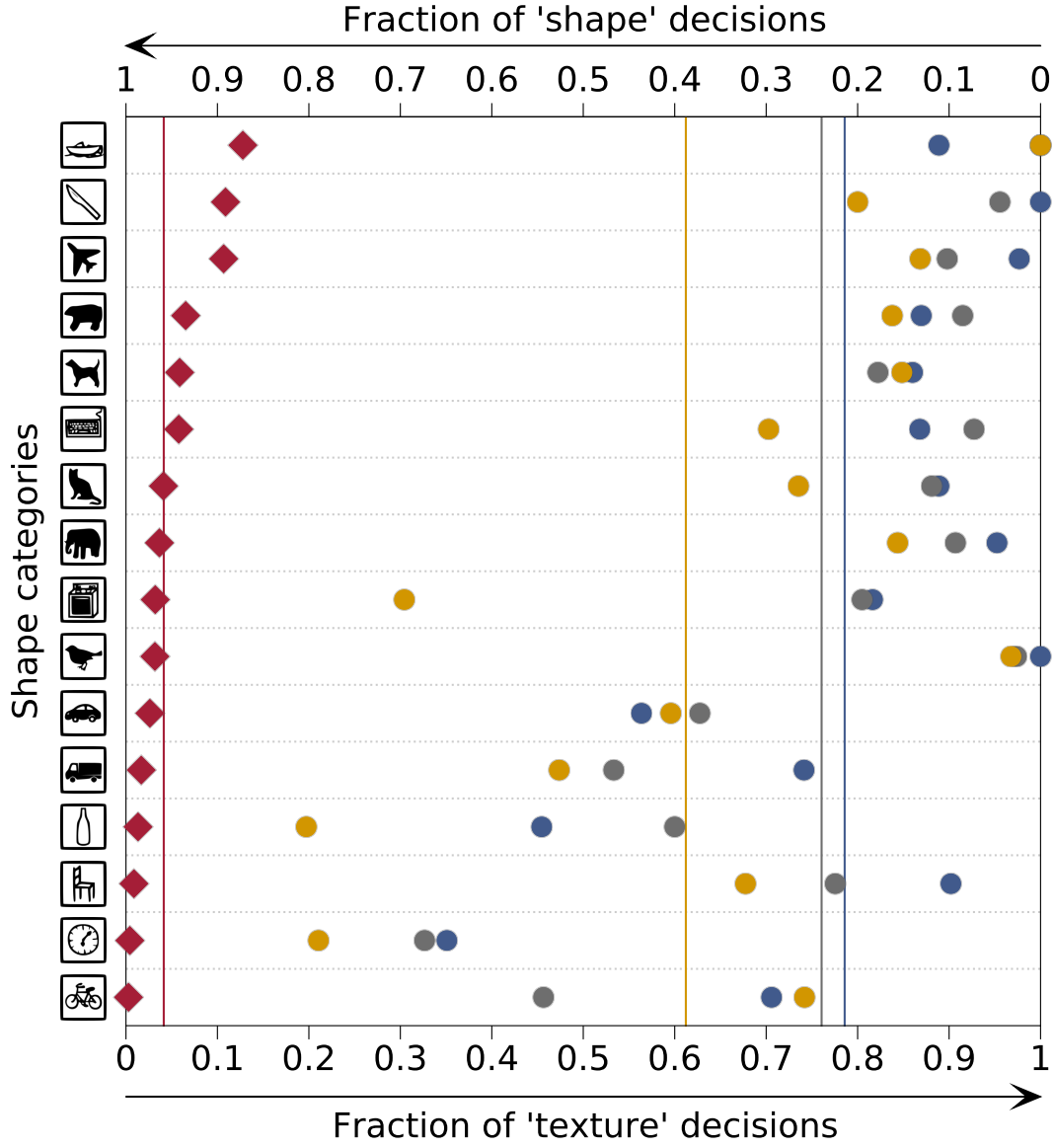


Figure 3: Shape vs. texture bias: category-level plot.

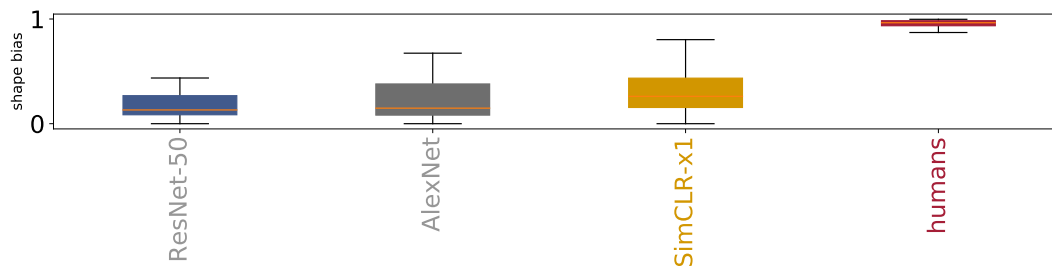


Figure 4: Shape vs. texture bias: boxplot.

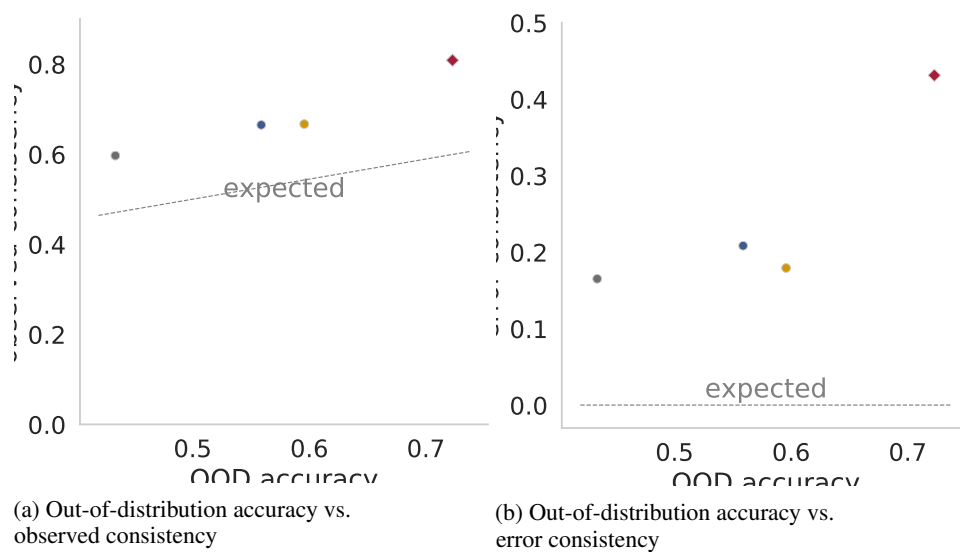


Figure 5: Observed consistency and error consistency between models and humans as a function of out-of-distribution (OOD) accuracy. Dotted lines indicate consistency expected by chance.

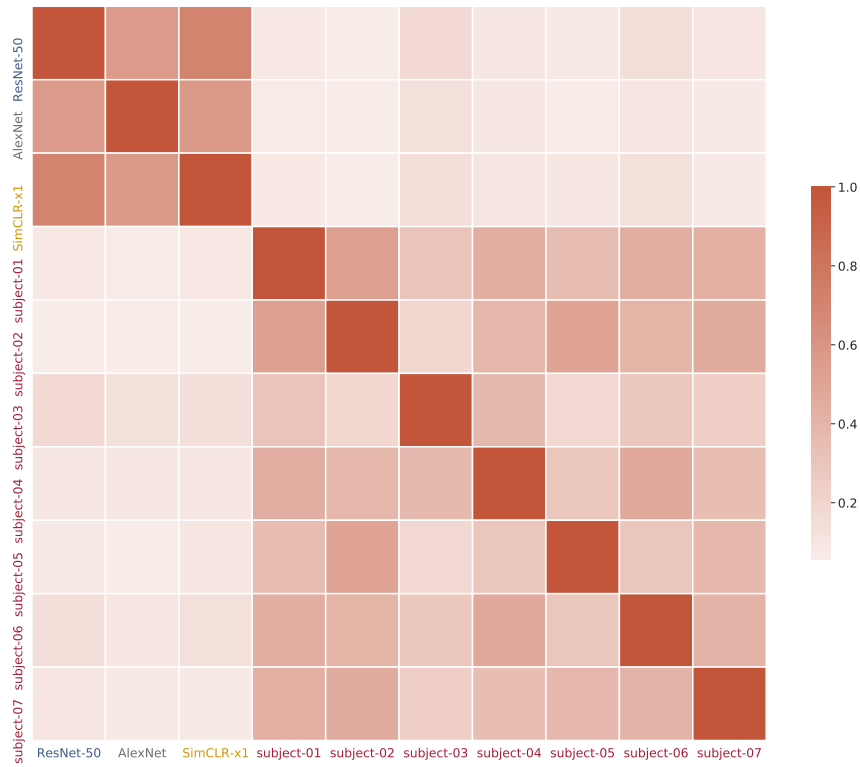


Figure 6: Error consistency for ‘sketch’ images.

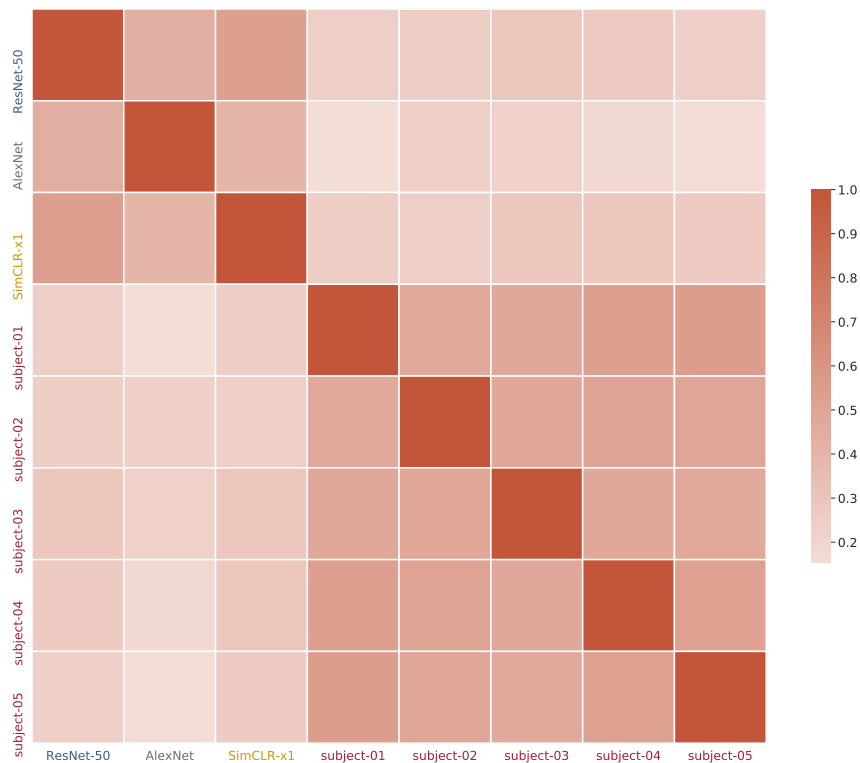


Figure 7: Error consistency for ‘stylized’ images.

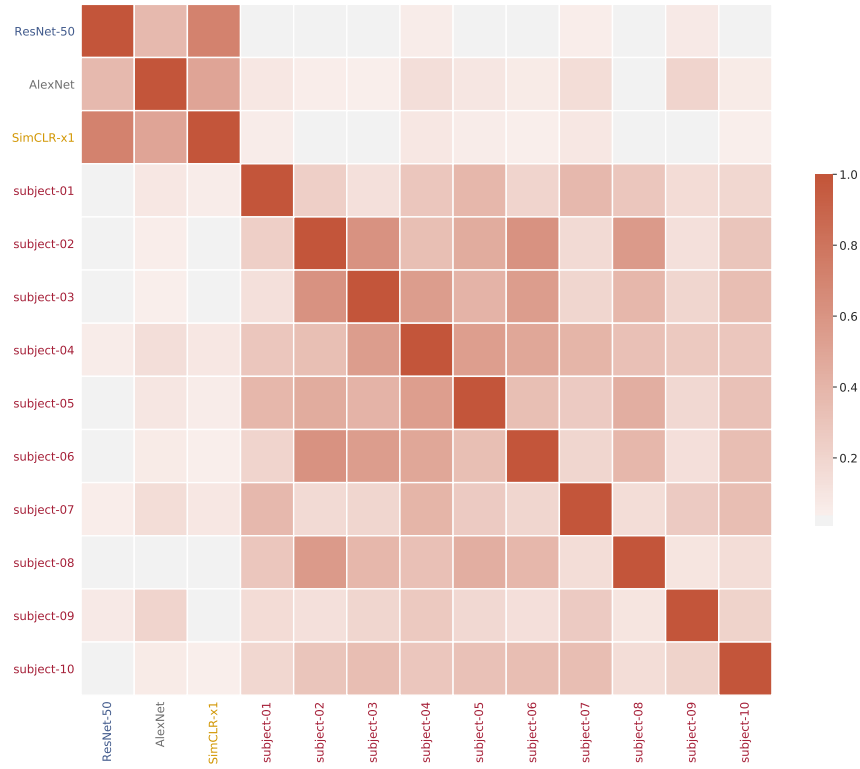


Figure 8: Error consistency for 'edge' images.

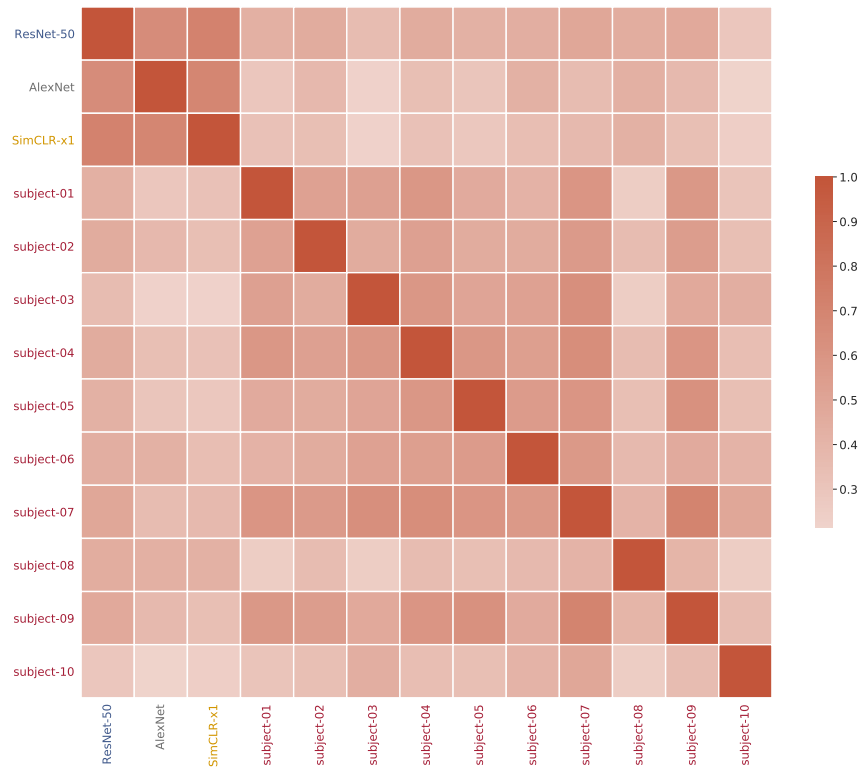


Figure 9: Error consistency for 'silhouette' images.

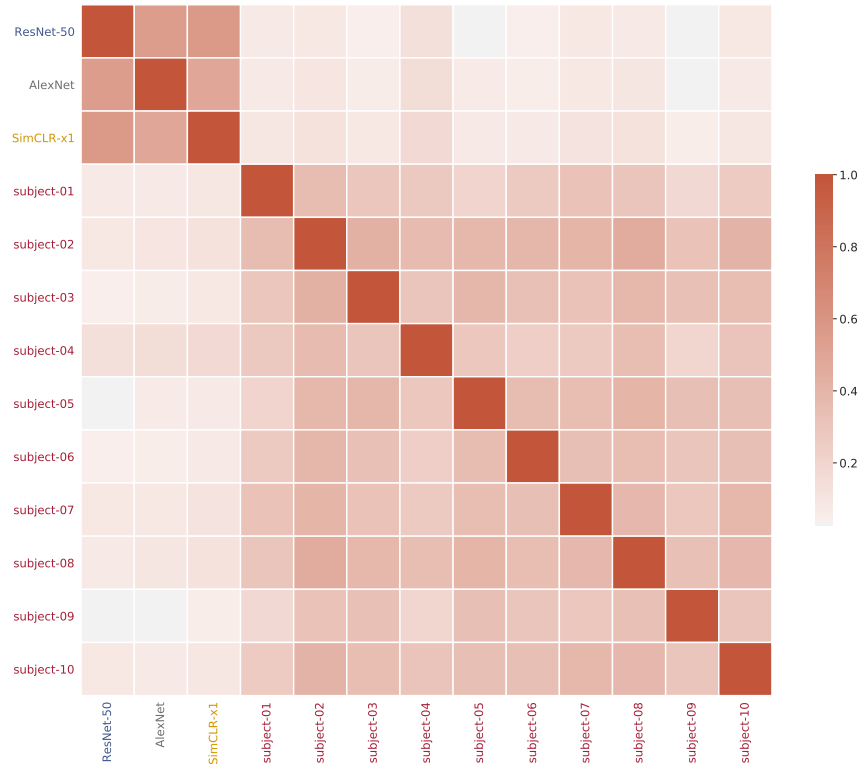


Figure 10: Error consistency for ‘cue conflict’ images.

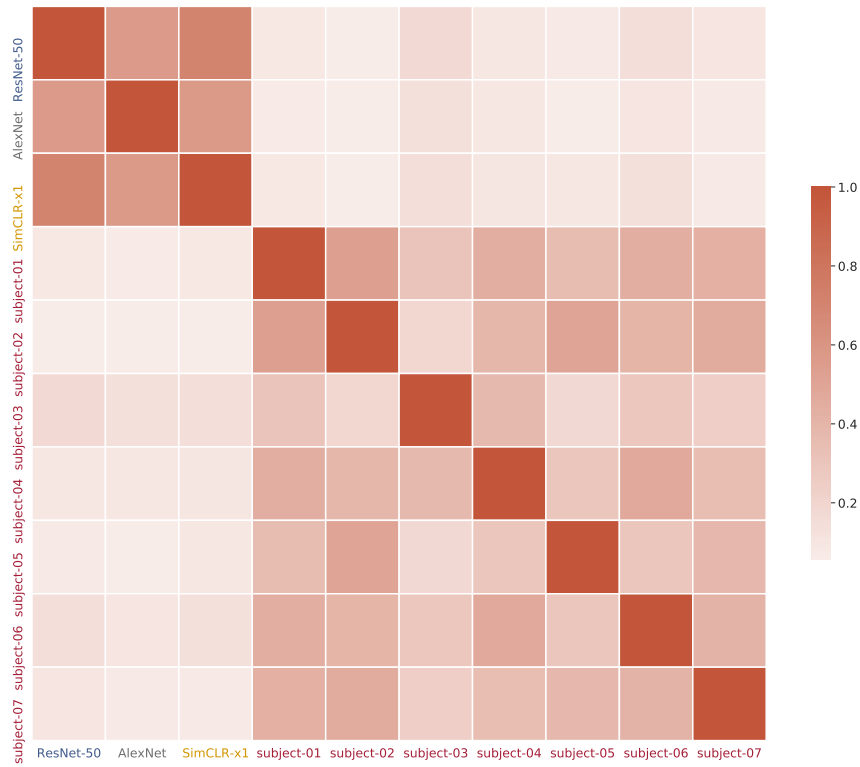


Figure 11: Error consistency for ‘sketch’ images.

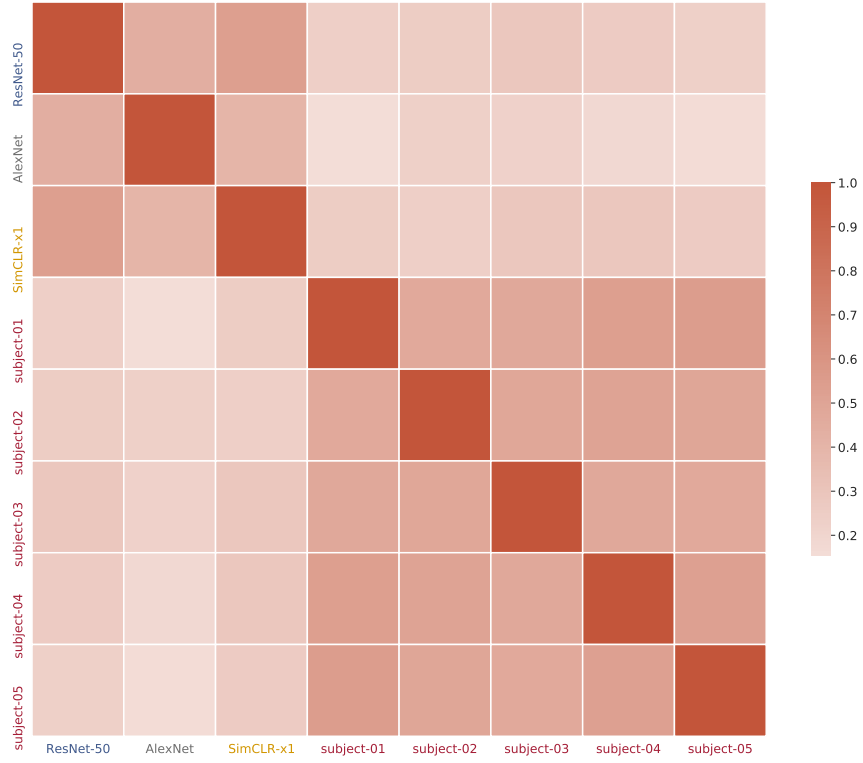


Figure 12: Error consistency for ‘stylized’ images.

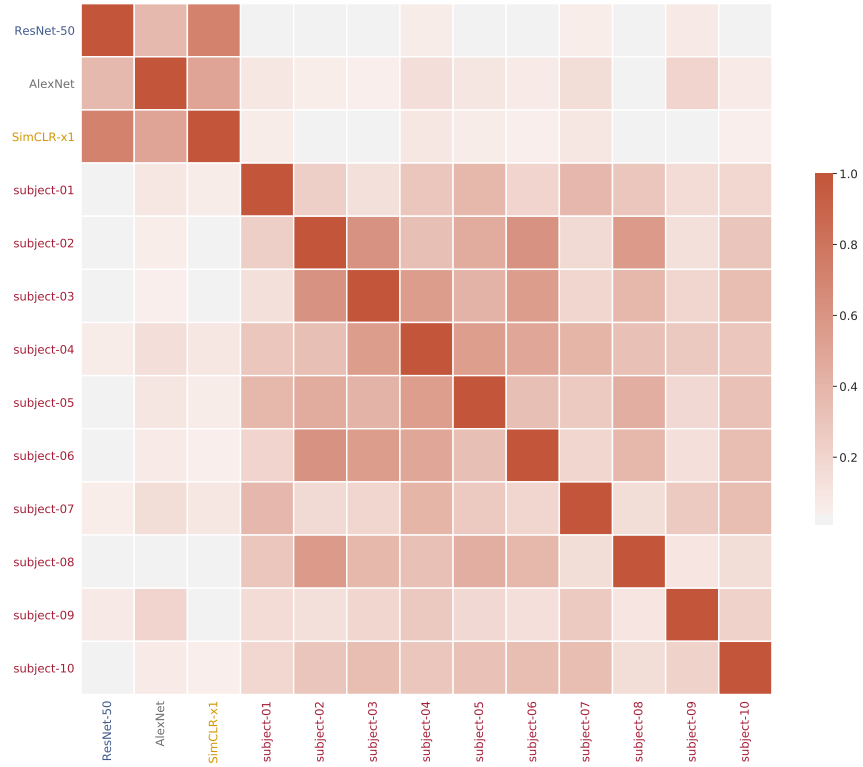


Figure 13: Error consistency for ‘edge’ images.



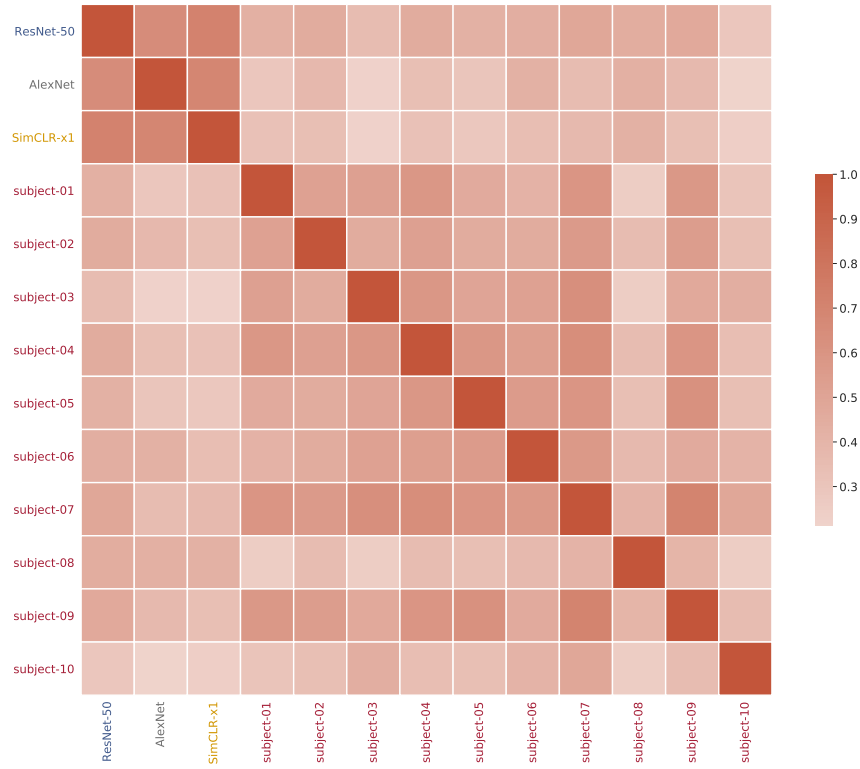


Figure 14: Error consistency for 'silhouette' images.

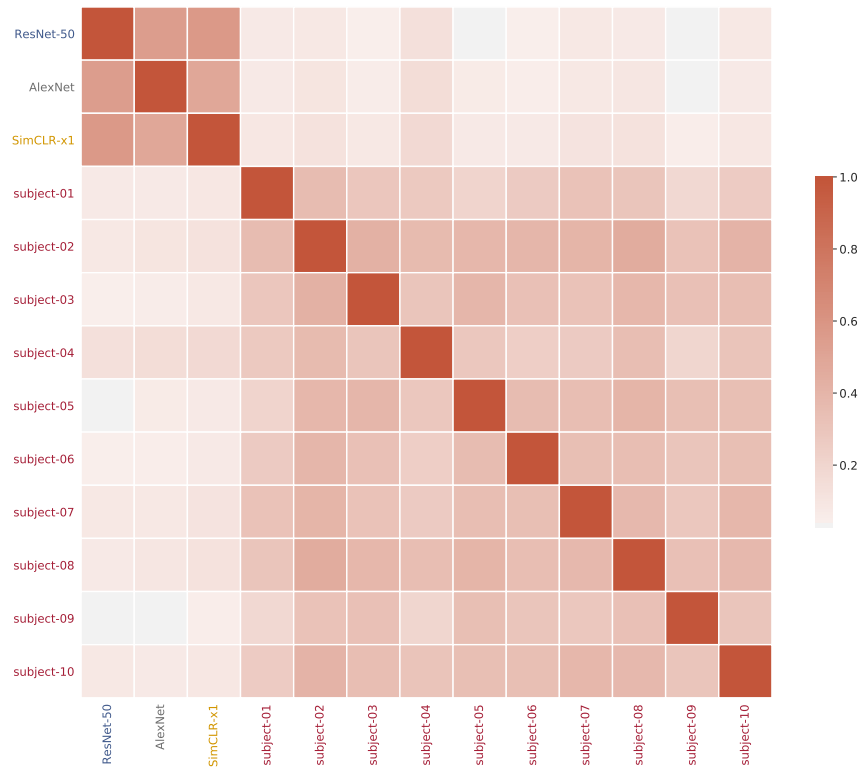


Figure 15: Error consistency for 'cue conflict' images.