

model	layer	TCAV Score	signif cant
balanced_5x50	fc1	$0.15 \pm 0.10$	yes
	fc4	$0.14 \pm 0.13$	yes
blue2_5x50	fc1	$1.00 \pm 0.00$	yes
	fc4	$1.00 \pm 0.00$	yes
balanced_3x50	fc1	$0.14 \pm 0.13$	yes
	fc2	$0.08 \pm 0.05$	yes
blue2_3x50	fc1	$0.80 \pm 0.08$	yes
	fc2	$0.78 \pm 0.11$	yes
balanced_3x20	fc1	$0.20 \pm 0.06$	yes
	fc2	$0.14 \pm 0.05$	yes
blue2_3x20	fc1	$1.00 \pm 0.01$	yes
	fc2	$0.98 \pm 0.01$	yes