

Models	Eval Size	#Params	#FLOPs	TPUv3-core-days	Top-1 Accuracy
ResNet + ViT-L/16	384 ²	330M	-	-	87.12
ViT-L/16	512 ²	307M	364B	0.68K	87.76
ViT-H/14	518 ²	632M	1021B	2.5K	88.55
NFNet-F4+	512 ²	527M	367B	1.86K	89.2
CoAtNet-3 [†]	384 ²	168M	114B	0.58K	88.52
CoAtNet-3 [†]	512 ²	168M	214B	0.58K	88.81
CoAtNet-4	512 ²	275M	361B	0.95K	89.11
CoAtNet-5	512 ²	688M	812B	1.82K	89.77
ViT-G/14	518 ²	1.84B	5160B	>30K [◇]	90.45
CoAtNet-6	512 ²	1.47B	1521B	6.6K	90.45
CoAtNet-7	512 ²	2.44B	2586B	20.1K	90.88

