

Models	#Trainable Params	#Total Params	VQAv2		OK-VQA	GQA
			val	test-dev	test	test-dev
VL-T5 _{no-vqa}	224M	269M	13.5	-	5.8	6.3
FewVLM (Jin et al., 2022)	740M	785M	47.7	-	16.5	29.3
Frozen (Tsimpoukelli et al., 2021)	40M	7.1B	29.6	-	5.9	-
VLKD (Dai et al., 2022)	406M	832M	42.6	44.5	13.3	-
Flamingo3B (Alayrac et al., 2022)	1.4B	3.2B	-	49.2	41.2	-
Flamingo9B (Alayrac et al., 2022)	1.8B	9.3B	-	51.8	44.7	-
Flamingo80B (Alayrac et al., 2022)	10.2B	80B	-	56.3	50.6	-
BLIP-2 ViT-L OPT _{2.7B}	104M	3.1B	50.1	49.7	30.2	33.9
BLIP-2 ViT-g OPT _{2.7B}	107M	3.8B	53.5	52.3	31.7	34.6
BLIP-2 ViT-g OPT _{6.7B}	108M	7.8B	54.3	52.6	36.4	36.4
BLIP-2 ViT-L FlanT5 _{XL}	103M	3.4B	62.6	62.3	39.4	<u>44.4</u>
BLIP-2 ViT-g FlanT5 _{XL}	107M	4.1B	<u>63.1</u>	<u>63.0</u>	40.7	44.2
BLIP-2 ViT-g FlanT5 _{XXL}	108M	12.1B	65.2	65.0	<u>45.9</u>	44.7