

Models	#Trainable Params	VQAv2	
		test-dev	test-std
<i>Open-ended generation models</i>			
ALBEF (Li et al., 2021)	314M	75.84	76.04
BLIP (Li et al., 2022)	385M	78.25	78.32
OFA (Wang et al., 2022a)	930M	82.00	82.00
Flamingo80B (Alayrac et al., 2022)	10.6B	82.00	82.10
<b>BLIP-2</b> ViT-g FlanT5 <sub>XL</sub>	1.2B	81.55	81.66
<b>BLIP-2</b> ViT-g OPT <sub>2.7B</sub>	1.2B	81.59	81.74
<b>BLIP-2</b> ViT-g OPT <sub>6.7B</sub>	1.2B	<b>82.19</b>	<b>82.30</b>
<i>Closed-ended classification models</i>			
VinVL	345M	76.52	76.60
SimVLM (Wang et al., 2021b)	~1.4B	80.03	80.34
CoCa (Yu et al., 2022)	2.1B	82.30	82.30
BEIT-3 (Wang et al., 2022b)	1.9B	<b>84.19</b>	<b>84.03</b>