

LLM	FlanT5 <sub>XL</sub>	OPT <sub>2.7B</sub>	OPT <sub>6.7B</sub>
Fine-tuning epochs		5	
Warmup steps		1000	
Learning rate		1e-5	
Batch size		128	
AdamW $\beta$		(0.9,0.999)	
Weight decay		0.05	
Drop path		0	
Image resolution		490	
Prompt		“Question: {} Answer:”	
Inference beam size		5	
Layer-wise learning rate decay for ViT	0.95	0.95	0.9