

Method	cross-entropy optimization				CIDEr optimization			
	B@4	M	C	S	B@4	M	C	S
BUTD [2]	36.2	27.0	113.5	20.3	36.3	27.7	120.1	21.4
VLP [47]	36.5	28.4	117.7	21.3	39.5	29.3	129.3	23.2
AoANet [11]	37.2	28.4	119.8	21.3	38.9	29.2	129.8	22.4
OSCAR _B	36.5	30.3	123.7	23.1	40.5	29.7	137.6	22.8
OSCAR _L	37.4	30.7	127.8	23.5	41.7	30.6	140.0	24.5