

Model	B@4	M	C
SA-GoogleNet+3D-CNN	41.9	29.6	-
LSTM-YT	33.3	29.1	-
S2VT	-	29.8	-
LSTM-E	45.3	31.0	-
HRNE	46.7	33.9	-
Boundary-aware encoder	42.5	32.4	63.5