

Dataset	Context	Videos	Description source	# Clips	# Sentences
TACoS	cooking	123	crowd	7,206	18,227
TACoS multilevel	cooking	273	crowd	14,105	52,593
MSVD	various	-	crowd	1,970	70,028
movie description	various	72	Script + DVS	54,076	54,076
movie description (DVS part)	various	46	DVS	30,680	30,680
M-VAD (ours)	various	92	DVS	48,986	55,904