

**Table 4: Comparison results with state-of-the-art methods on multi-paragraph video grounding in SynopGround.**

Method	mIoU	IoU@0.3	IoU@0.5	IoU@0.7
Human	85.1	97.3	92.5	85.0
Random	7.3	8.3	3.2	0.8
DepNet [5]	30.7	47.2	28.7	12.8
PRVG [55]	34.7	52.7	29.3	10.5
LGMR (Ours)	<b>44.4</b>	<b>67.9</b>	<b>46.7</b>	<b>21.8</b>

**Table 6: Effect of the local-level modeling, global-level modeling, and cross-level reasoning in the iterative query decoder.**

Local	Global	Cross	mIoU	IoU@0.3	IoU@0.5	IoU@0.7
✓	×	×	34.5	53.1	32.4	13.5
✓	✓	×	42.8	66.6	44.2	19.0
✓	✓	✓	<b>44.4</b>	<b>67.9</b>	<b>46.7</b>	<b>21.8</b>