Table 3. Comparison with representative methods based on Qwen2.5-VL-7B. † denotes single-prefill variants.

Method	HR-Bench 4K	HR-Bench 8K	V^*			CV-Bench
			Attribute	Spatial	Overall	
Multi-turn						
Thyme-RL [47]	77.0	72.0	84.4	81.6	83.3	78.2
DeepEyes [48]	73.9	71.0	88.7	79.0	84.8	77.8
Two-stage						
API [45]	71.5	68.5	80.9	75.0	78.5	78.1
FGVP-Mask [44]	71.9	66.9	82.6	79.0	81.2	77.2
FGVP-RBM [44]	41.1	42.6	27.8	50.0	36.7	52.3
SoM [43]	51.6	41.4	33.0	65.8	46.1	50.1
ViCrop [14]	72.1	67.9	83.5	72.4	79.1	76.4
Single-pass						
Qwen2.5-VL-7B [4]	70.9	66.9	81.7	77.6	80.1	77.7
Vision-R1 [10]	72.5	70.9	80.0	80.3	80.1	75.4
VLM-R1 [28]	64.5	61.6	80.9	61.8	73.3	64.2
LVR [16]	67.5	64.6	83.5	76.3	80.6	77.7
HyLaR [8]	73.1	70.4	83.5	75.0	80.1	75.7
SkiLa [32]	72.3	64.6	84.4	81.6	83.3	76.9
Monet [35]	69.0	65.4	83.5	82.9	83.3	66.3
PAPO [38]	72.0	69.1	81.7	79.0	80.6	78.2
PixelReasoner [30]	72.3	65.9	80.9	79.0	80.1	77.9
Thyme-RL† [47]	72.3	67.8	81.7	80.3	81.2	77.5
DeepEyes† [48]	71.0	68.3	80.9	75.0	78.5	76.3
ReAccess (Ours)	**73.3**	**71.8**	**84.4**	**82.9**	**83.8**	**78.8**