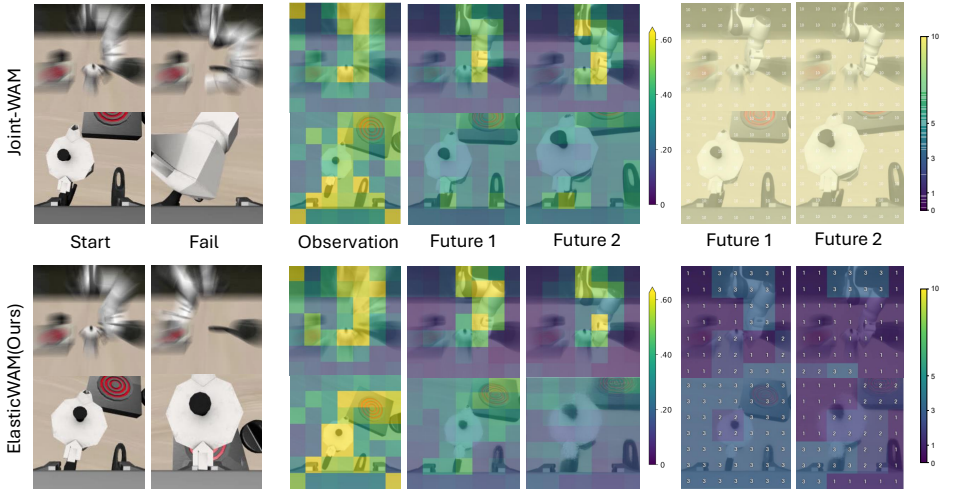

(a) Task Execution
(b) Action-to-Video Attention
(c) Video Token Updates
Joint-WAM
ElasticWAM(Ours)
Start
Fail
Observation
Future 1
Future 2
Future 1
Future 2
Start
Success
Observation
Future 1
Future 2
Future 1
Future 2
.60
.40
.20
0
10
5
3
1
0