Fig. 1. Comparison of self-supervised/unsupervised VOS methods in terms of segmentation accuracy ($\mathcal{J}\&\mathcal{F}$) and memory footprint on DAVIS-17 Val dataset.