

# MomentDiff: Generative Video Moment Retrieval from Random to Real

Anonymous Author(s)

Affiliation

Address

email

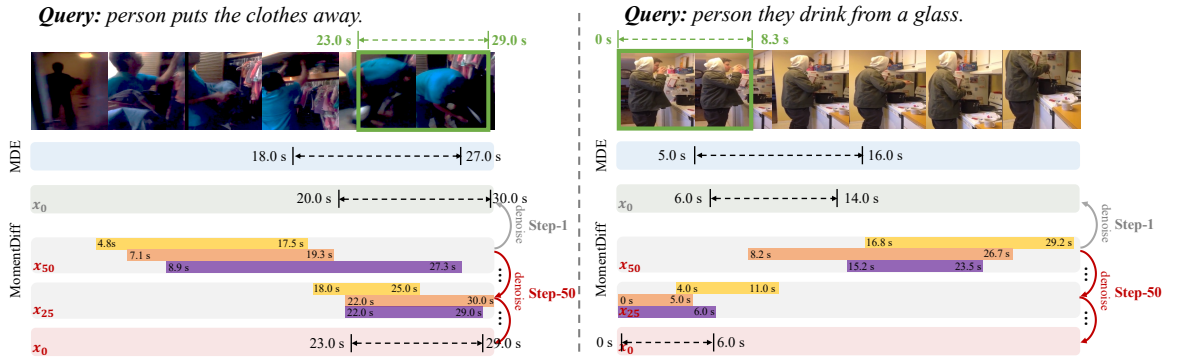


Figure 1: Visualization of the diffusion process on Charades-STA. We show the prediction results of two methods: MomentDETR (MDE) and our MomentDiff. In our approach, we show the results of single-step (step=1) and iterative generation (step=50). For clarity, we show 3 random spans ( $x_{50}$ ) with Gaussian initialization, and progressively get the top-1 result ( $x_0$ ) according to the confidence score. The results of MomentDETR and single-step generation are often not fine enough, and our MomentDiff can get high IOU generated span through 50-step diffusion denoising. Green box: right segment.