



(a) Top-1 Decoding (b) Anchor Decoding (c) Spatio-Temporal Aware Decoding

