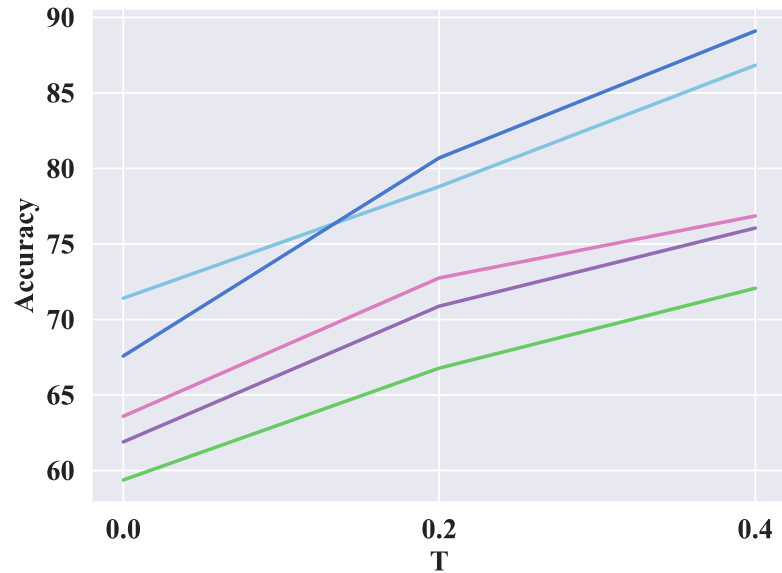


(a) Testing on audio-dominated subset



(b) Testing on visual-dominated subset

— Baseline — OGM-GE — Grad-Blending — Greedy — Audio unimodal — Visual unimodal