



(a) Video-Diagram Contrastive Loss

(b) Video-Manual Contrastive Loss

(c) Intra-Manual Contrastive Loss

