

Frozen \bigcirc Update $\mathcal{L}_1/\mathcal{L}_2$: Focal Loss for ViT/CNN $\mathcal{L}_3/\mathcal{L}_4$: Discrepancy Loss \cdots Strong augmentation — Weak augmentation $\mathcal{L}_{V \to C}$: CNN loss guided by ViT-generated pseudo-labels $\mathcal{L}_{C \to V}$: ViT loss guided by CNN-generated pseudo-labels θ_1/θ_2 : Threshold in ViT/CNN