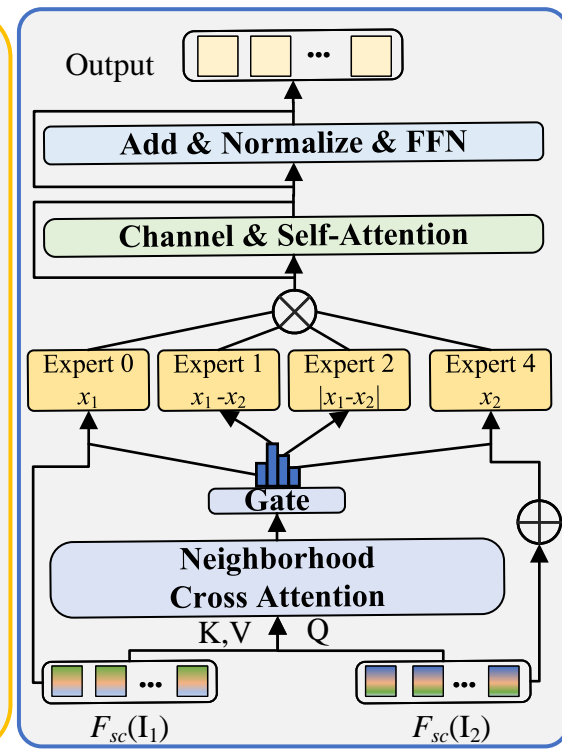
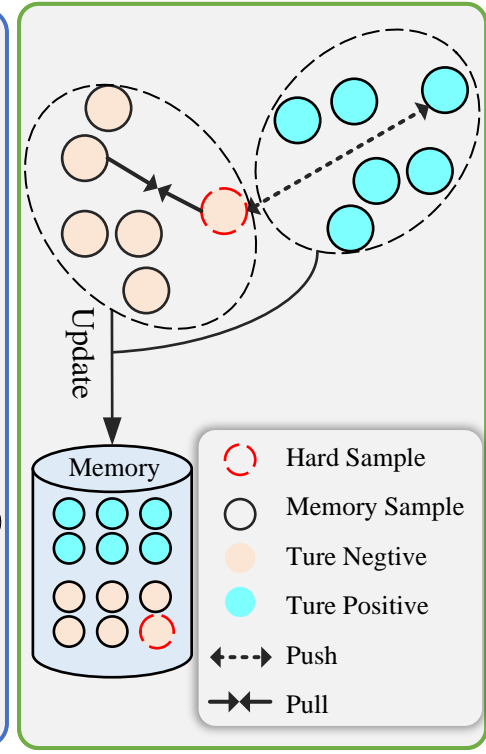


(a) SAM-CLIP Feature Encoder (SCFE)



(b) Cross-Perception Feature Module (CPFM)



(c) Hard Sample Contrast Module (HSCM)