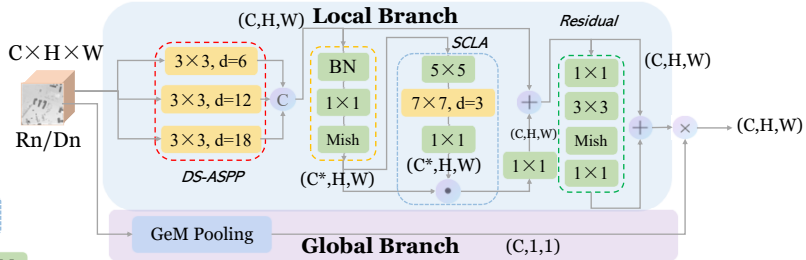
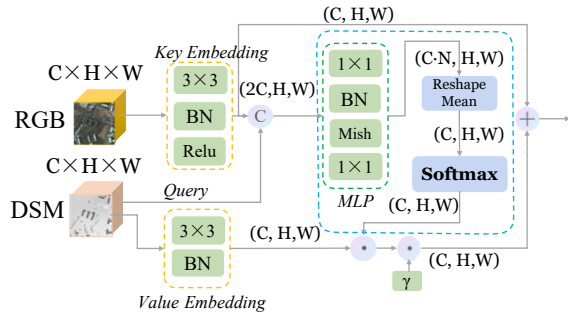


(a) Overview of the proposed framework



(b) Contextual Representation Enhancement Module (CREM)



(c) Cross-Modal Fusion Module (CMFM)