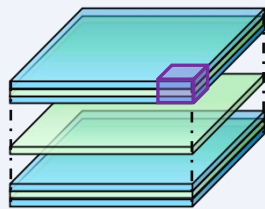
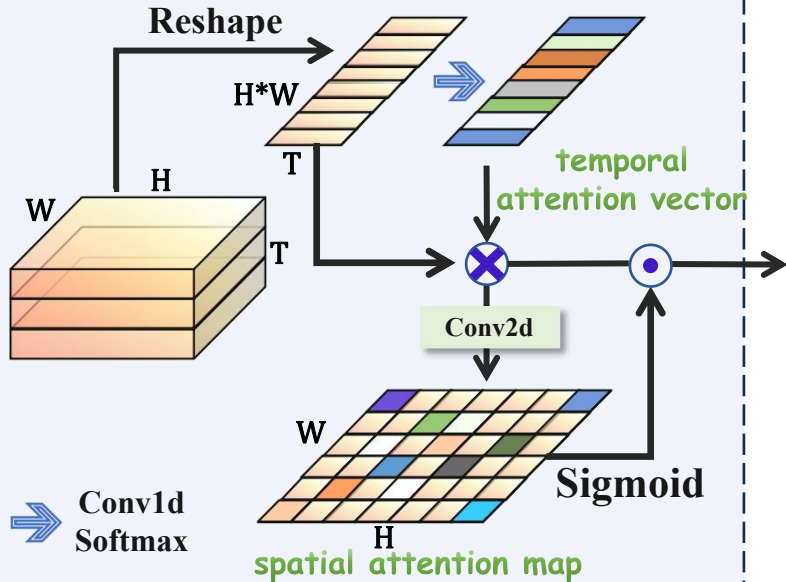
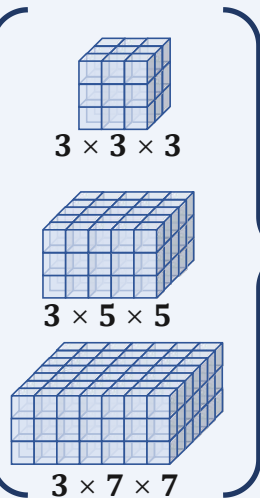


Cross-stitch



temporal-like
features

Multi-scale 3D difference feature
extraction



Spatial-temporal attention guidance