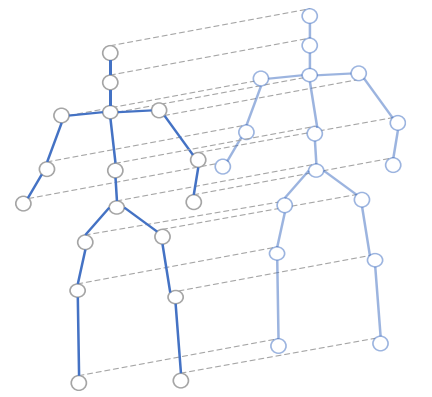
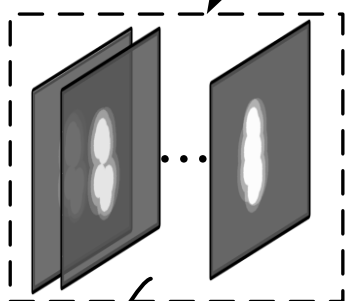


Stage-1

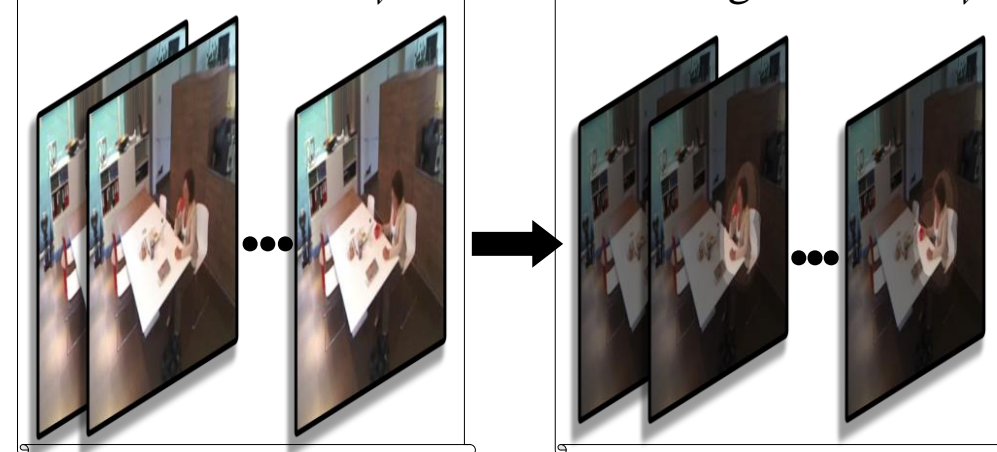
Skeleton



Joints of interest



Soft Shading

Total Frames F_v Shading Frames F_v^S 

Stage-2

Skeleton backbone

 f_s

Activation

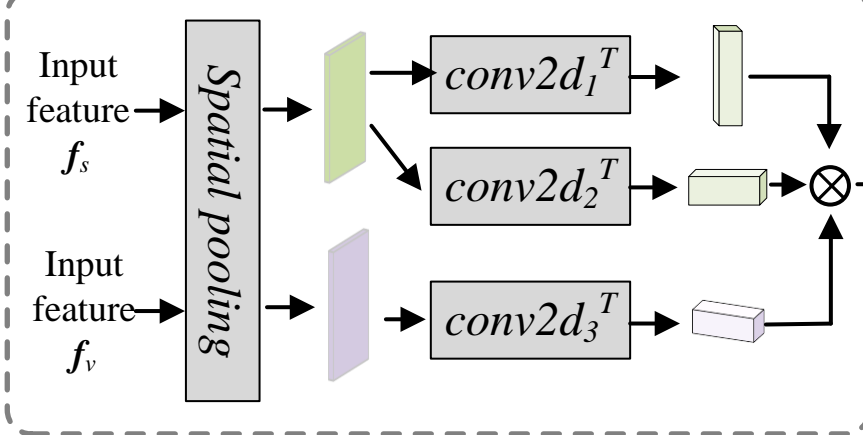
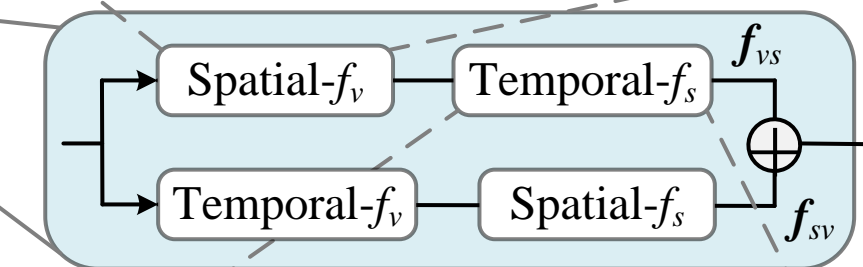
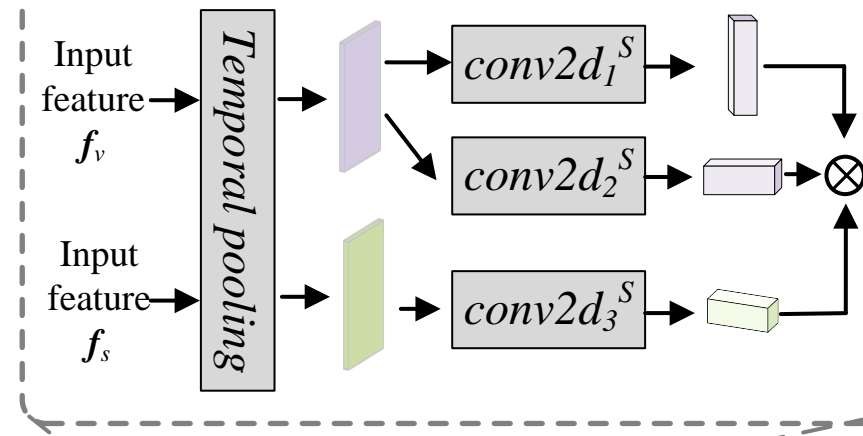
 \tilde{f}_s

Spatial-Temporal Guidance

 f_v

Interest frame extraction

Video backbone



Stage-3

Classifier

Shared

Classifier

Knowledge Distillation

Adaptive Fusion

❄️ Frozen parameters

⊗ Element-wise multiplication

⊕ Element-wise add

Learnable weights

Learnable weights