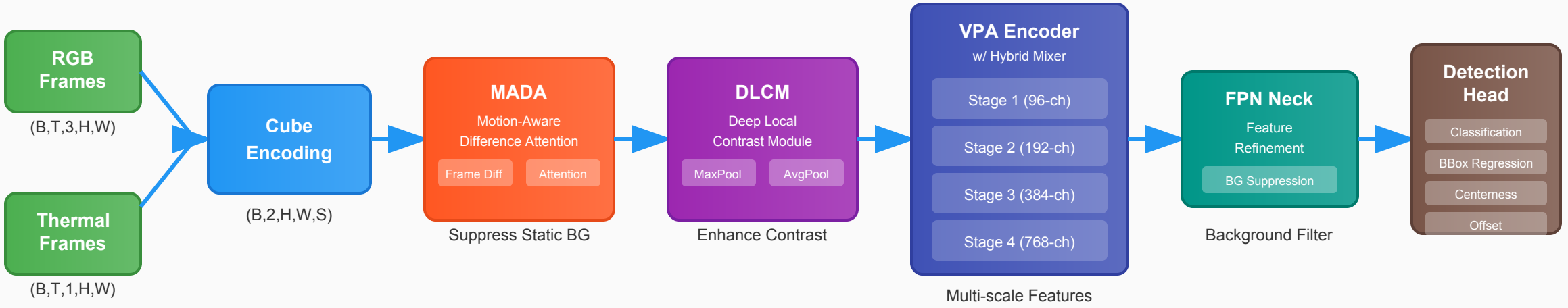
TLCFormer: Temporal-Local-Contrast Transformer Architecture



Core Innovations

MADA: Motion-Aware Difference Attention

1. Temporal Gradient:

   D_pre = |I_t - I_{t-1}|
   D_next = |I_{t+1} - I_t|

2. Motion Saliency Map:

   M_raw = D_pre * D_next

3. Feature Modulation:

   I'_t = I_t * (1 + a * sigmoid(F(M_raw)))

DLCM: Deep Local Contrast Module

1. Background Estimation:

   mu_bg = AvgPool_9x9(X)

2. Target Intensity:

   L_max = MaxPool_3x3(X)

3. Contrast Response:

   C = L_max^2 / (mu_bg + eps)
   X_out = X + beta * C

Hybrid Energy-Preserving Mixer

1. Dual-Path Pooling:

   P_max = MaxPool2d(X)
   P_avg = AvgPool2d(X)

2. Feature Fusion:

   P_hybrid = Concat(P_max, P_avg)

3. Channel Compression:

   X_out = X + GELU(Conv1x1(P_hybrid))