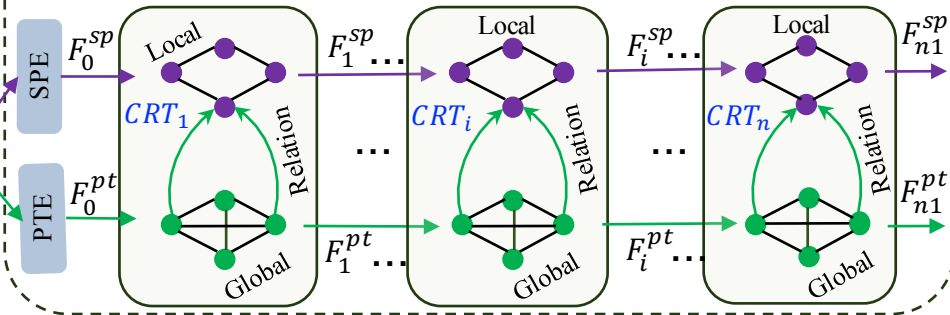




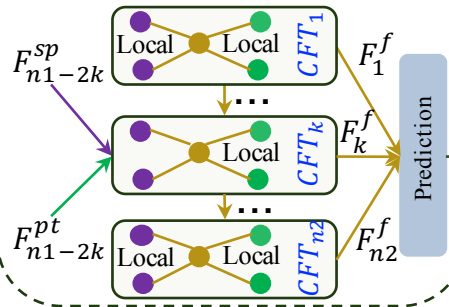
Image

Superpixel
branch
Patch
branch

Collaborative Relation Transformer

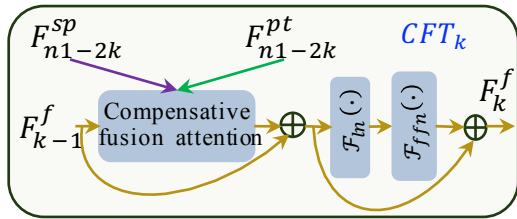
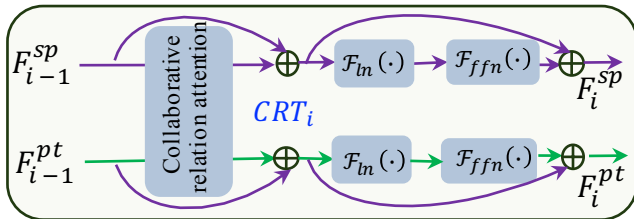


Compensative Fusion Transformer



y

F^{sp} Feature from
superpixel embedding
 F^{pt} Feature from patch
embedding
 F_k^f Feature fused with
structures and semantics



CRT_i i-th Transformer unit

CFT_k k-th Transformer unit

y Predicted salient object

SPE Superpixel embedding PTE Patch embedding \oplus Addition operation $F_{ffn}(\cdot)$ Feed-forward network $F_{ln}(\cdot)$ Layer normalization