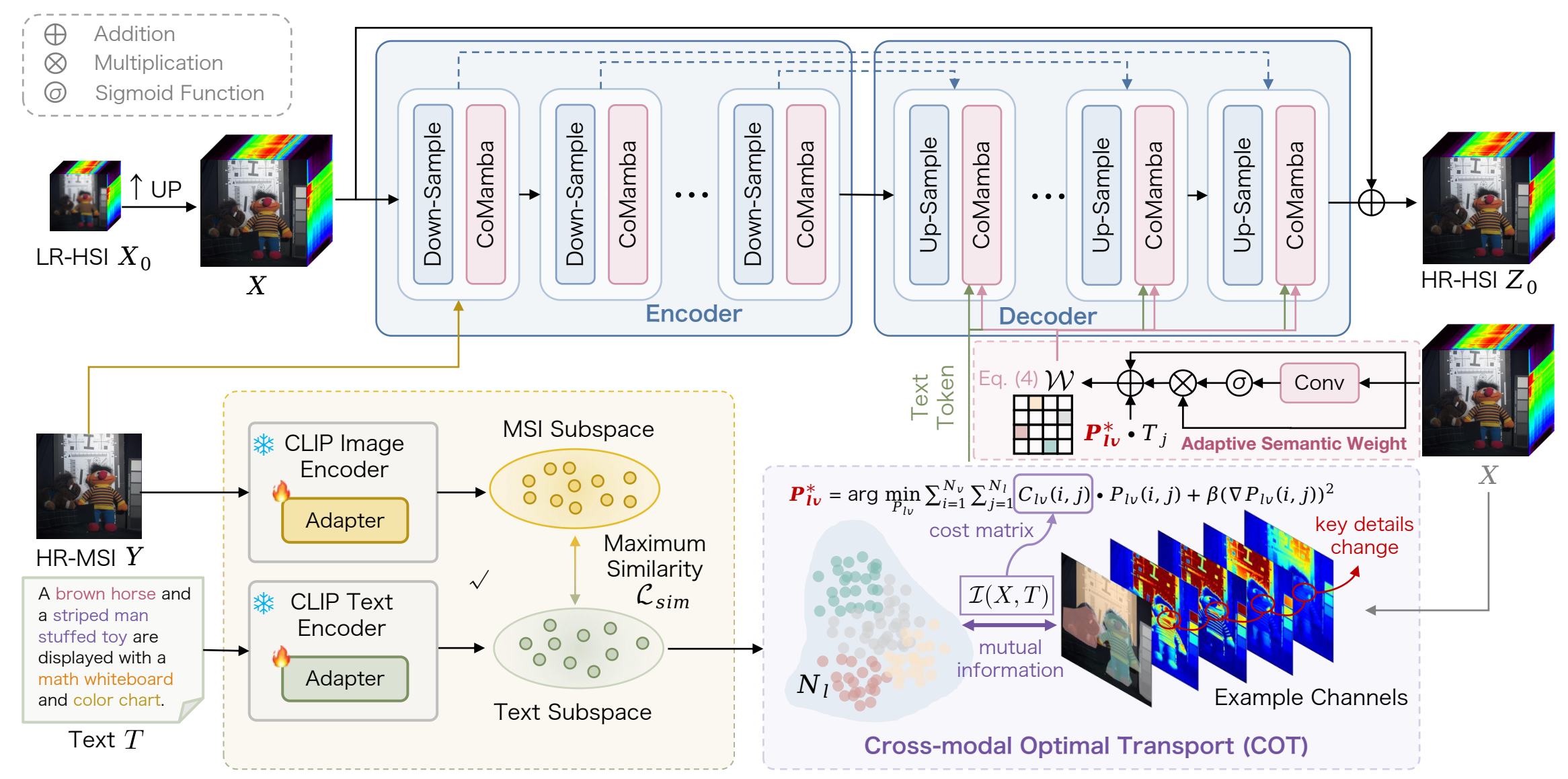
Addition
Multiplication
Sigmoid Function
↑ UP
LR-HSI X_0
X
Down-Sample
CoMamba
Encoder
Up-Sample
Decoder
HR-HSI Z_0
Text Token
Eq. (4) $\mathcal{W}$
Conv
$P^*_{lv} \cdot T_j$
Adaptive Semantic Weight
HR-MSI Y
CLIP Image Encoder
Adapter
MSI Subspace
Maximum Similarity $\mathcal{L}_{sim}$
CLIP Text Encoder
Text Subspace
A brown horse and a striped man stuffed toy are displayed with a math whiteboard and color chart.
Text T
$P^*_{lv} = \arg\min_{P_{lv}} \sum_{i=1}^{N_v} \sum_{j=1}^{N_l} C_{lv}(i,j) \cdot P_{lv}(i,j) + \beta(\nabla P_{lv}(i,j))^2$
cost matrix
$\mathcal{I}(X,T)$
mutual information
N_l
key details change
Example Channels
Cross-modal Optimal Transport (COT)