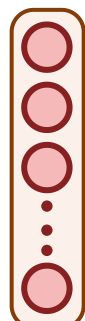


Input

Fully Connected Layers

Output

$\mathbf{X}_{\text{ST}}^{\text{A}}$



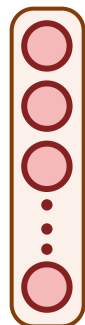
...



...



Shared
Parameter



...



...



Supervision

\hat{O}

d



Proximity Loss

Backpropagation

$$L_P(\hat{O}, d) = L_{\text{MSE}}(\hat{O}, p(d))$$