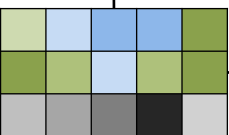




GT Y



Output \hat{Y}



Embedding Z

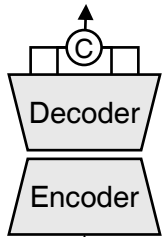
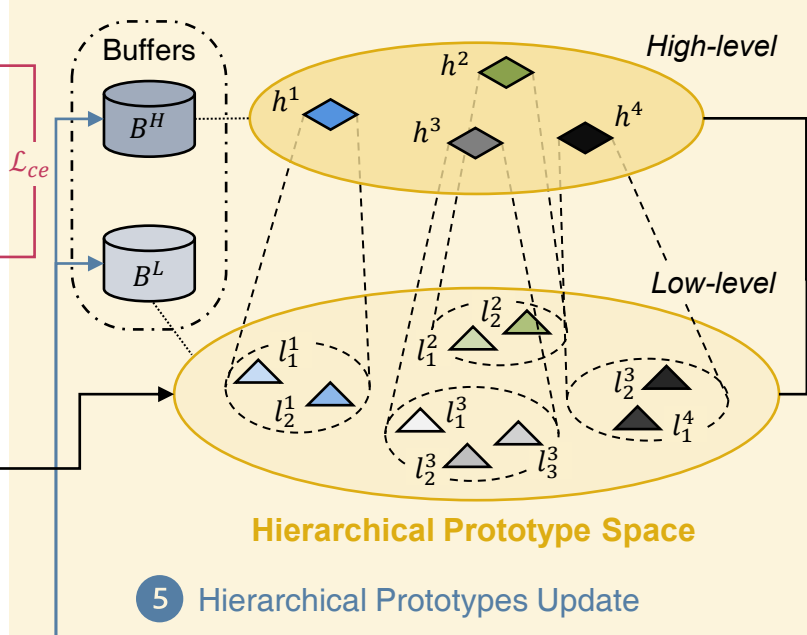
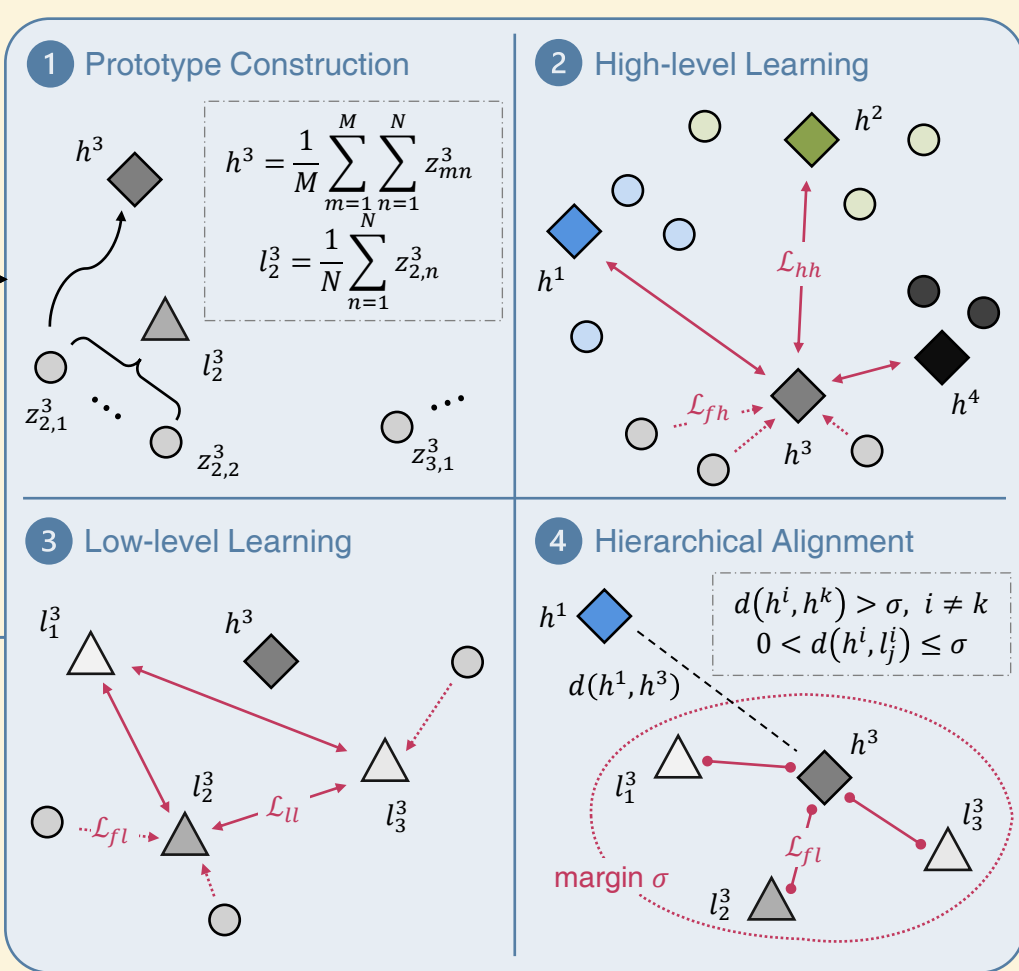


Image I



- Ⓢ Multi-scale Concatenation
- ◆ High-level Prototypes h^i
- ▲ Low-level Prototypes l_j^i
- Pixel Embeddings $z_{j,n}^i$
- ⋯⋯⋯ Attract
- ↔ Repel
- Align



Prototype-based Hierarchical Contrastive Learning