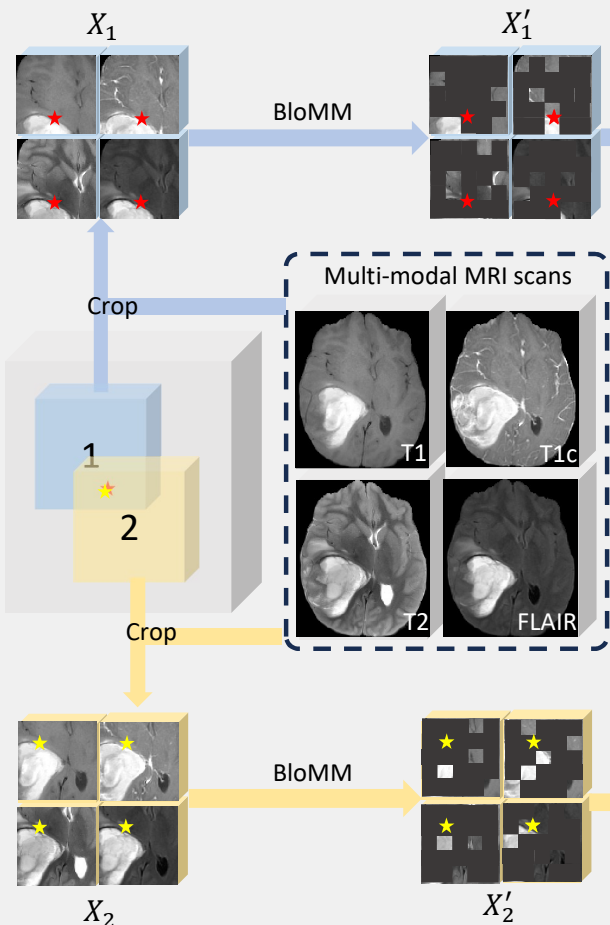
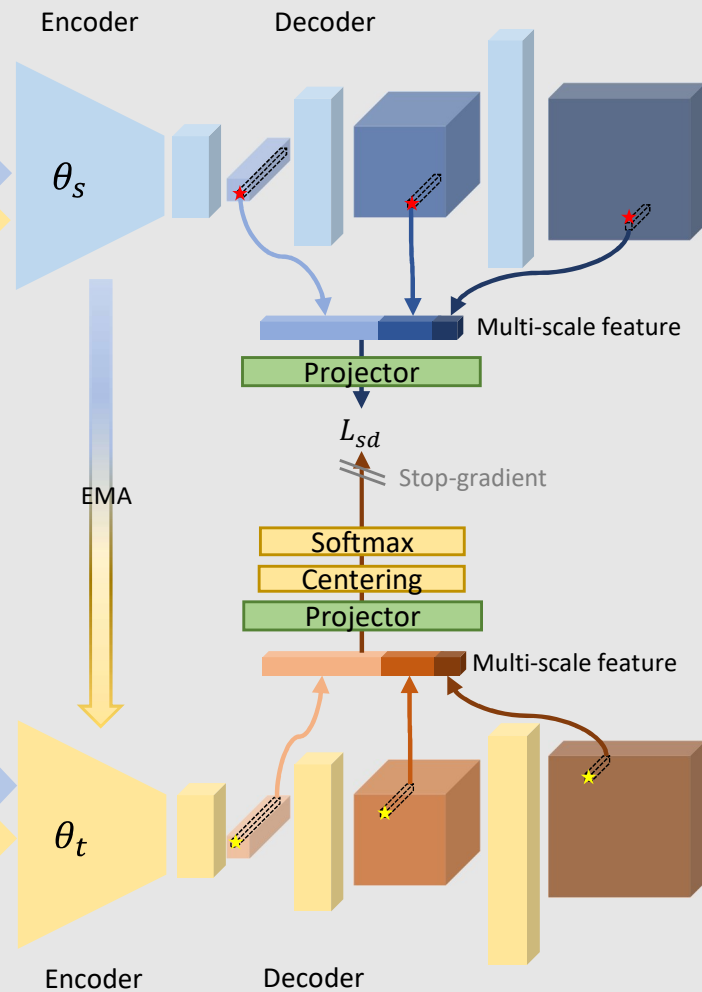


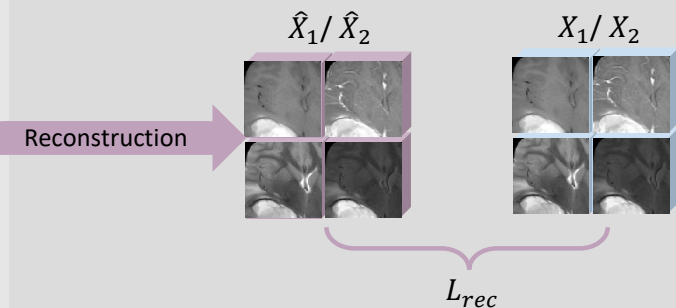
(a) Siamese Block-wise Modality Masking



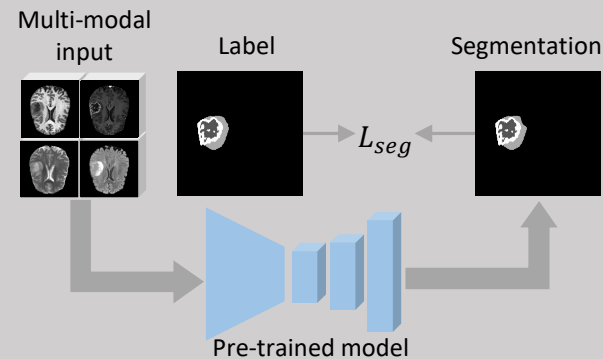
(b) Voxel-wise Multi-scale Feature Self-distillation



(c) Modality and Patch Reconstruction



(d) Downstream Fine-tuning



- Teacher's data flow
- Student's data flow
- Reconstruction's data flow
- Fine-tuning's data flow
- ★ ★ Voxel pair in different modalities and patches
- L_{sd} Loss for voxel-level multi-scale self-distillation
- L_{rec} Loss for modality and patch reconstruction
- L_{seg} Loss for downstream fine-tuning