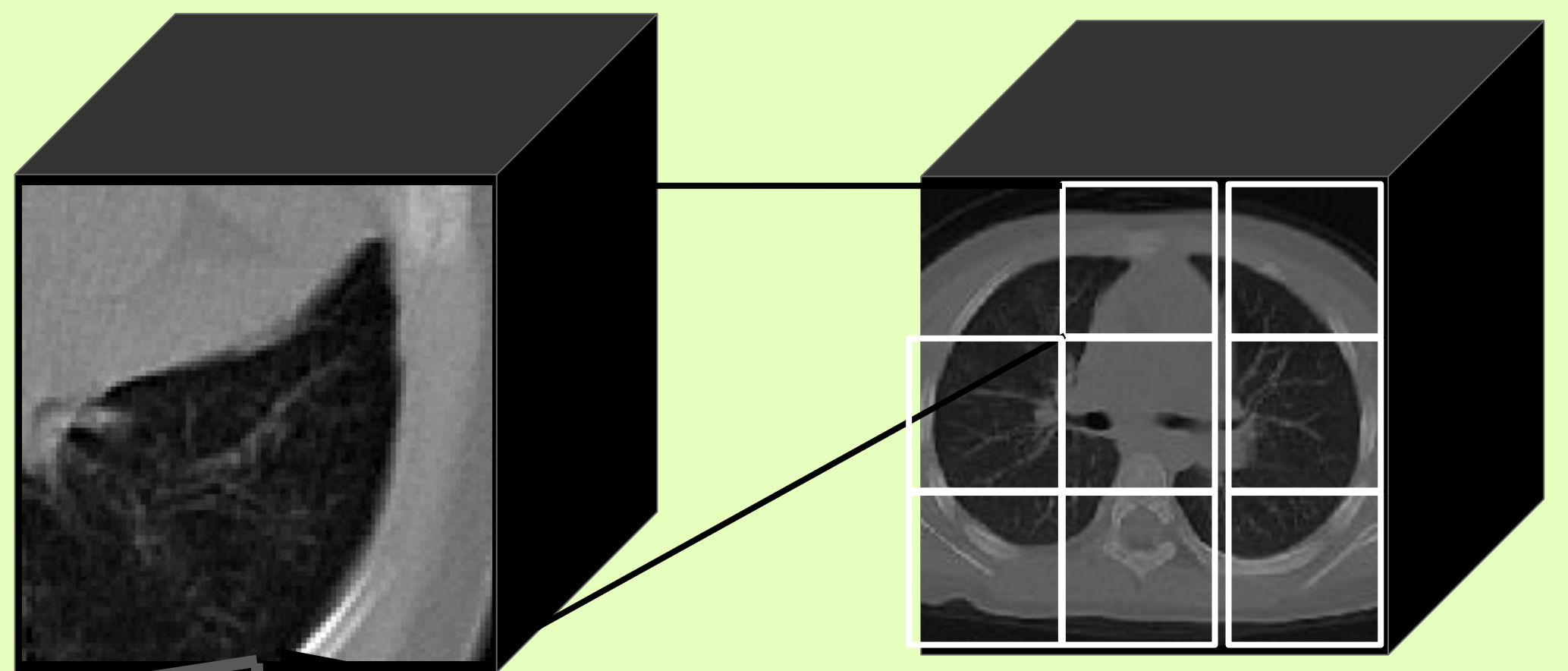
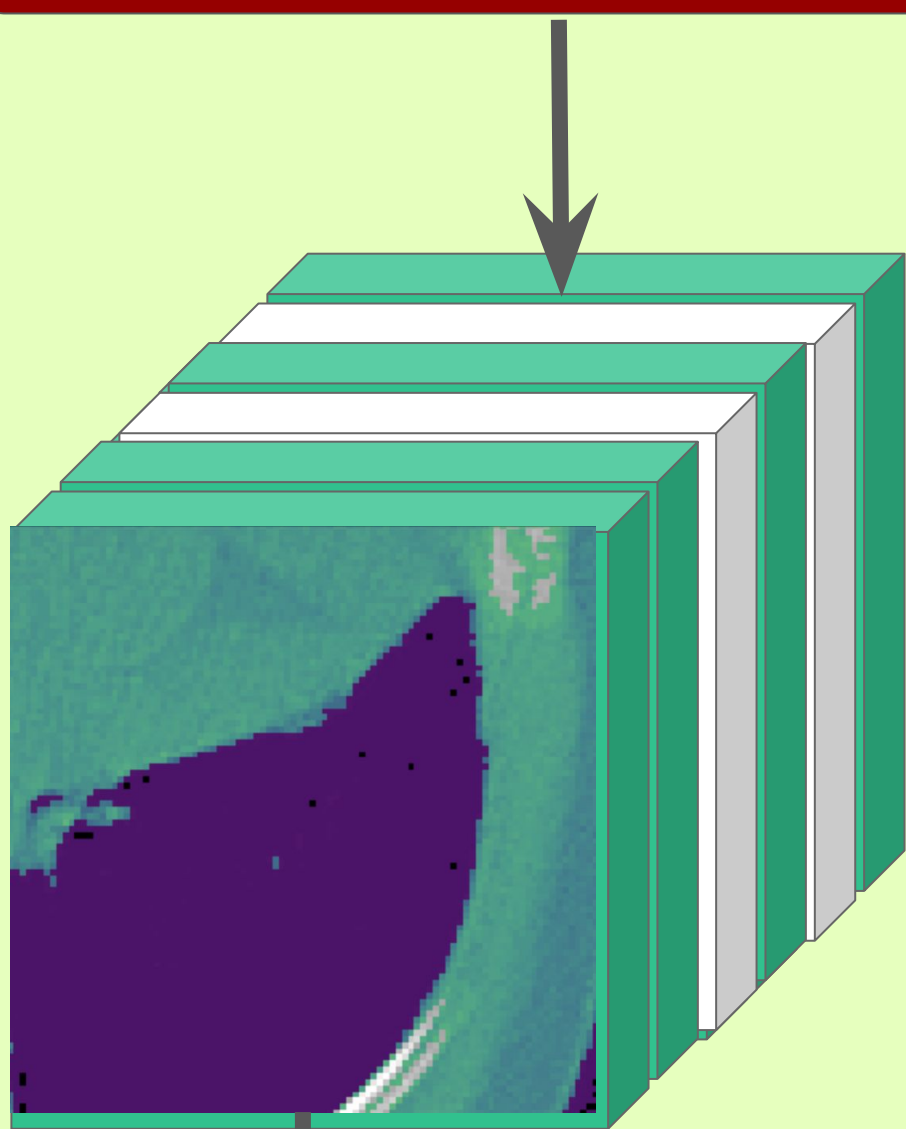


3D Medical Volume

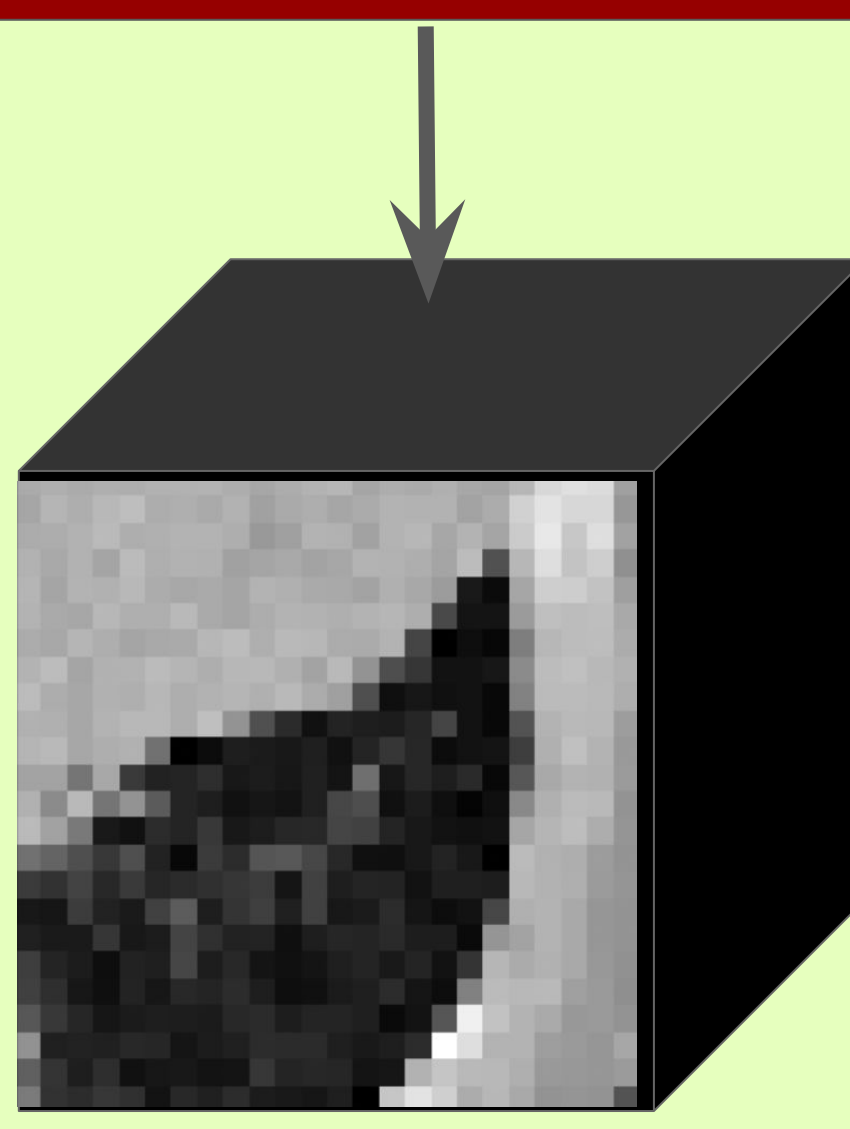


3D Medical Volume

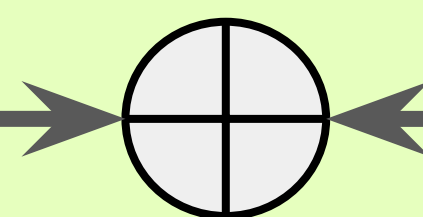
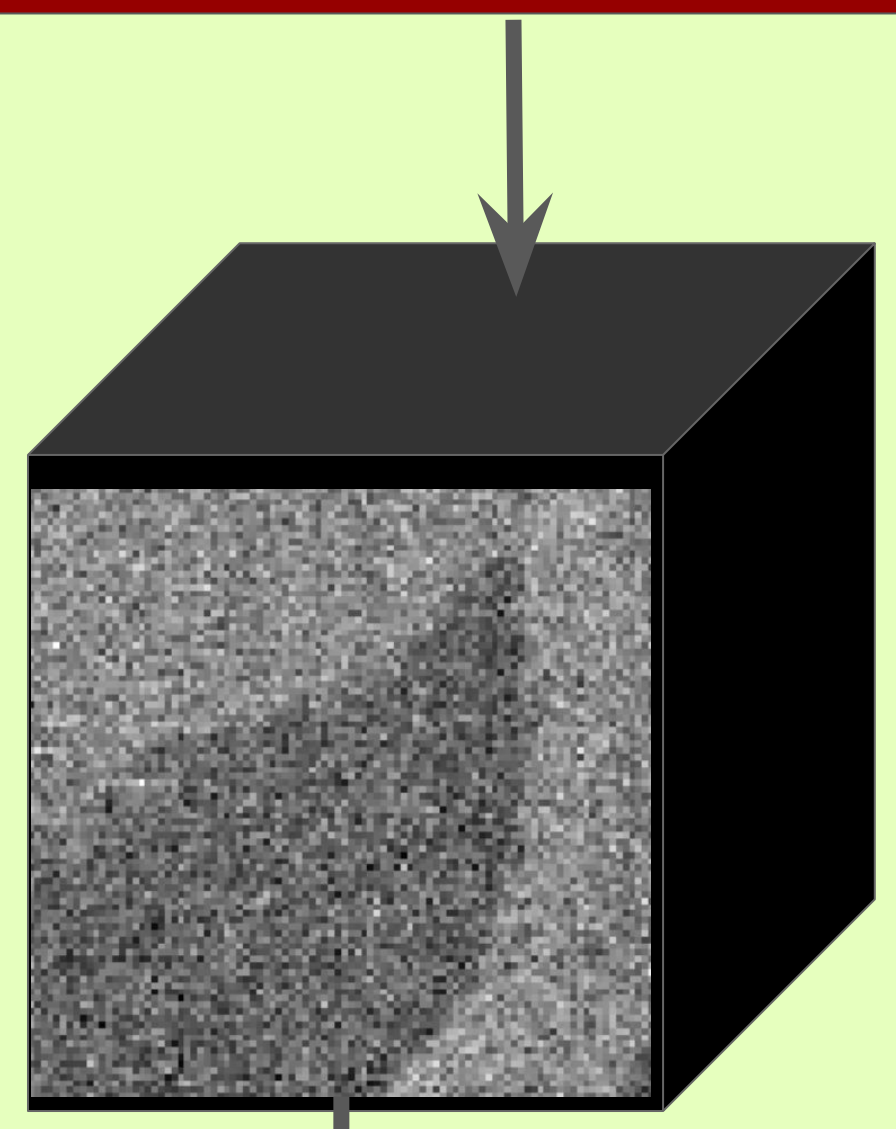
Local Masking



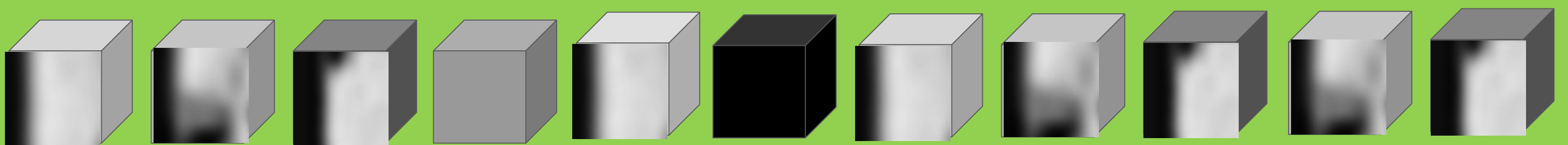
Downsampling



Add Noise



Disrupted Tokens

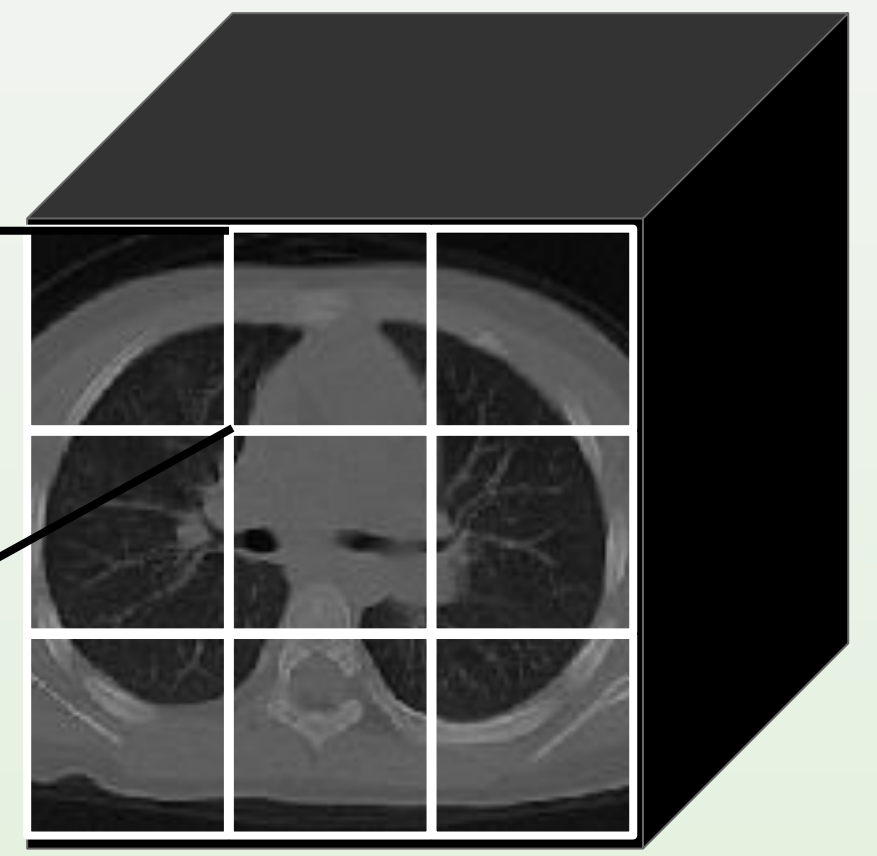
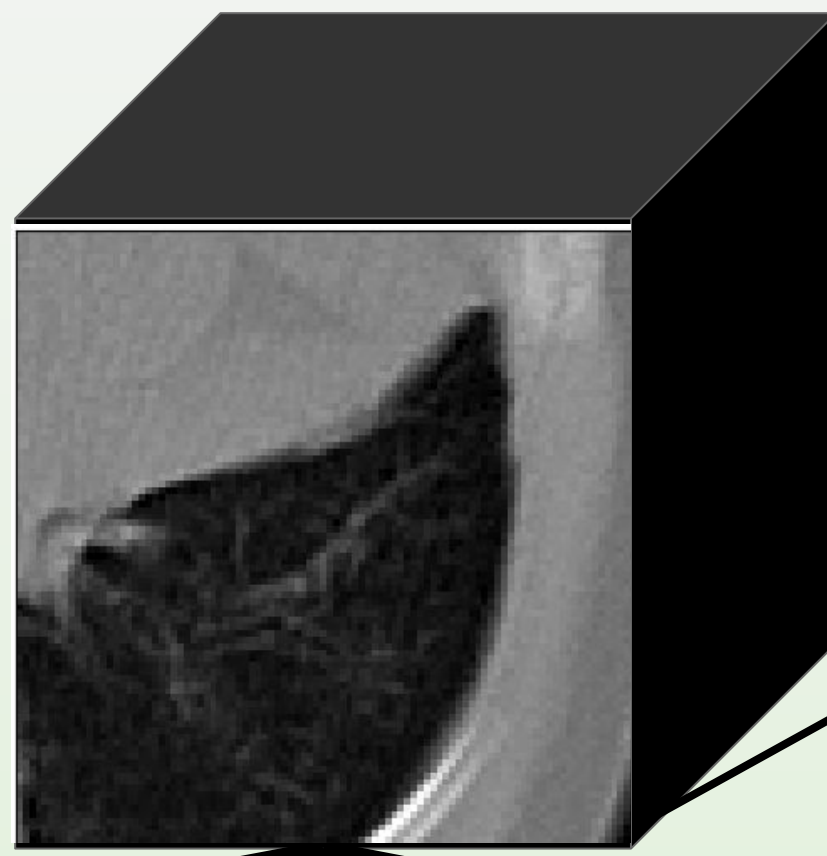
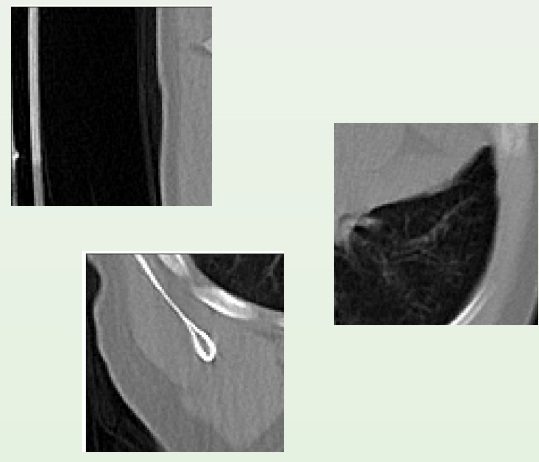


Transformer Encoder

Reconstruction

Cross Modal
Contrastive Learning

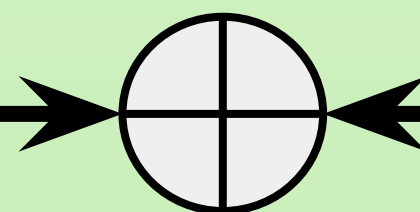
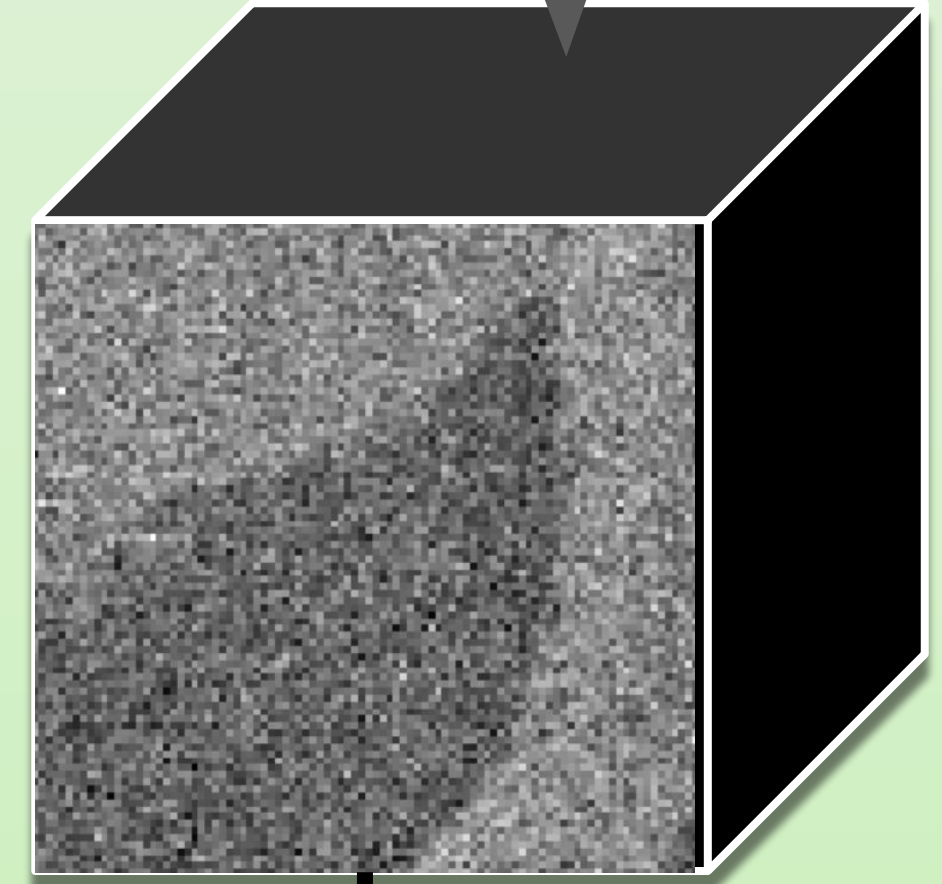
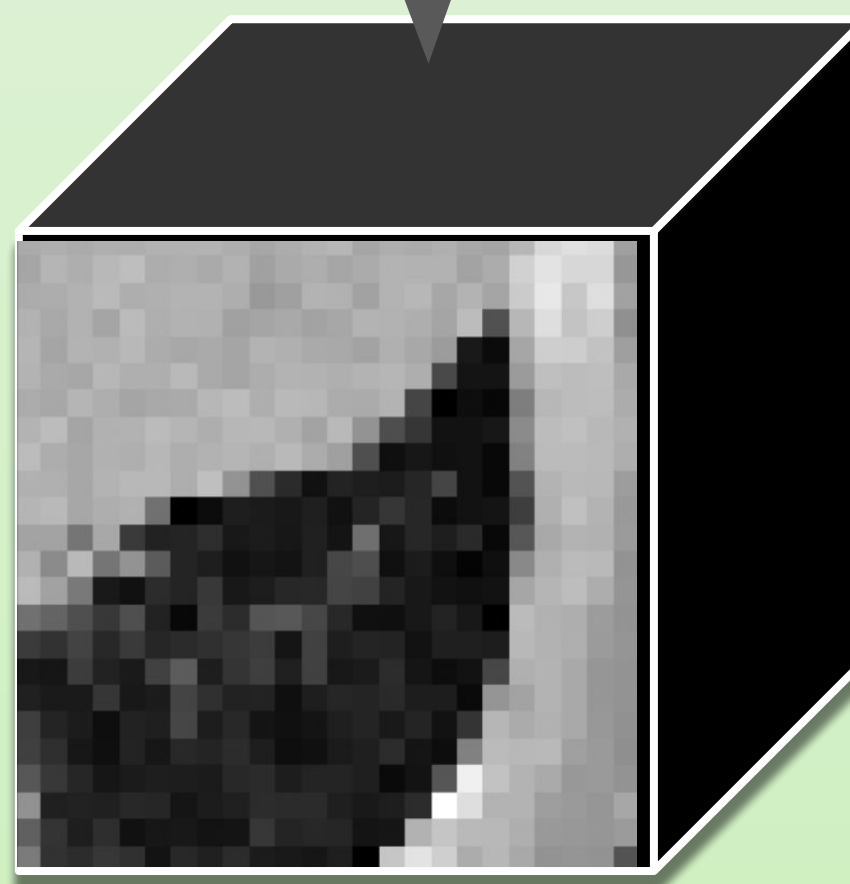
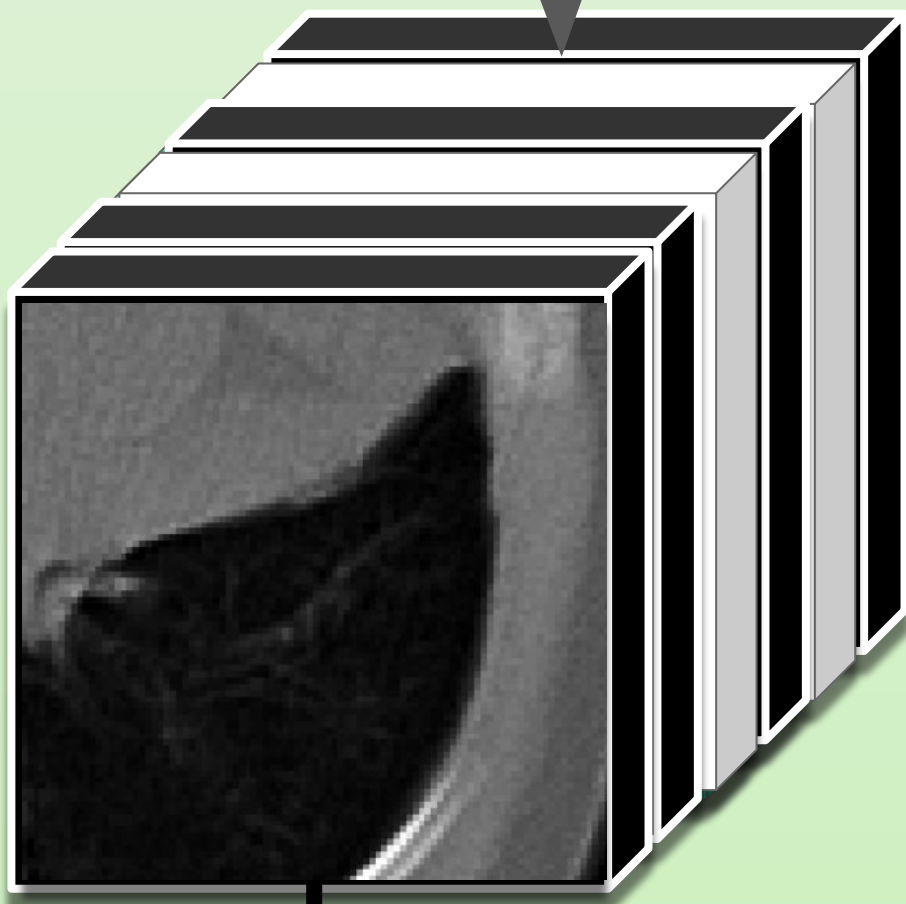
3D Medical Volume



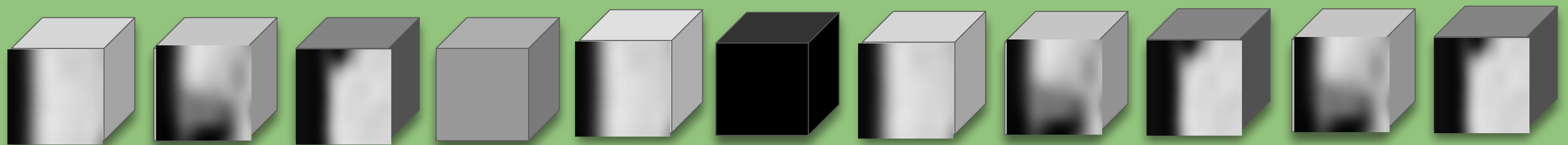
Local Masking

Downsampling

Add Noise



Disrupted Tokens

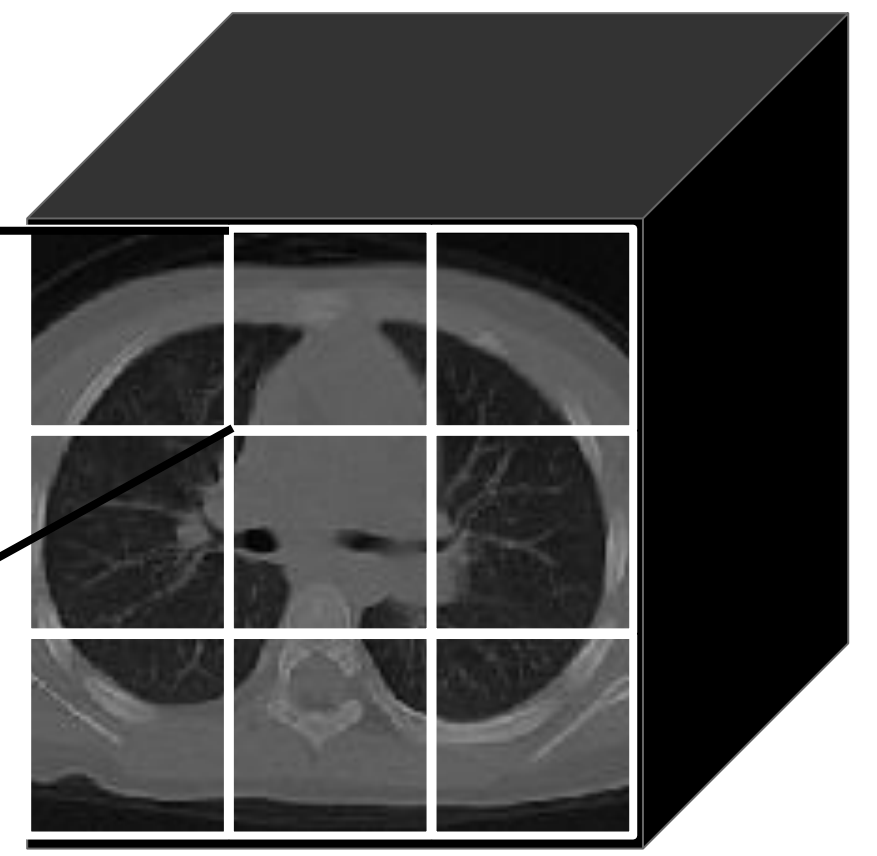
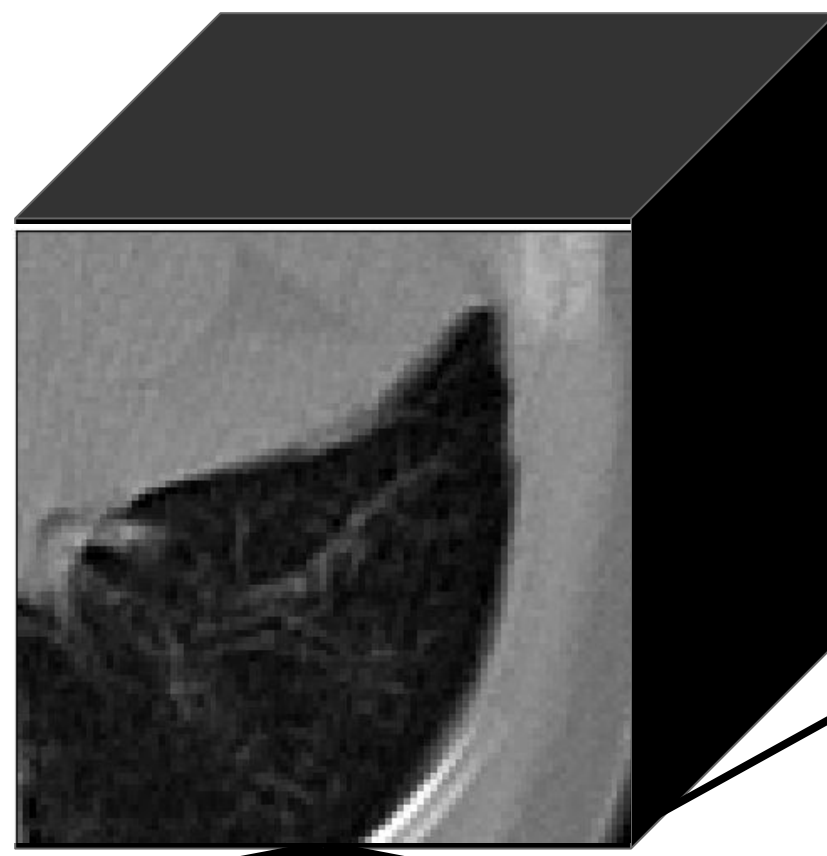


Transformer Encoder

Reconstruction

Cross Modal Contrastive Learning

3D Image Patch

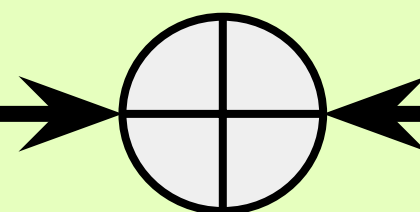
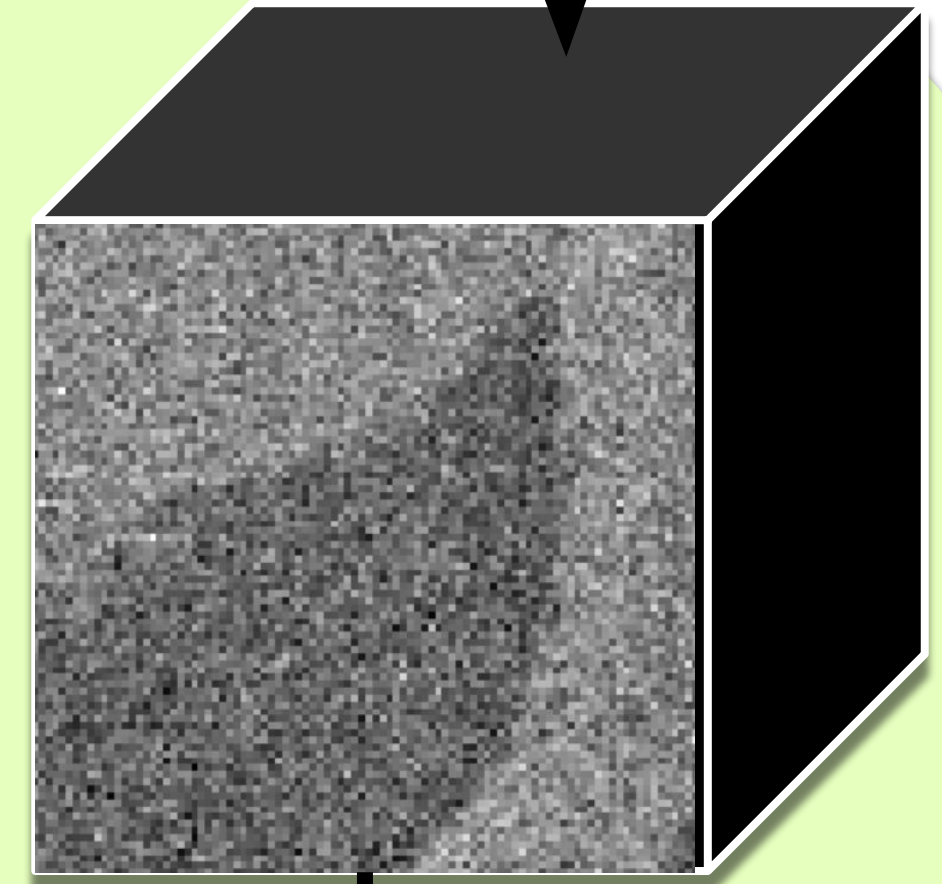
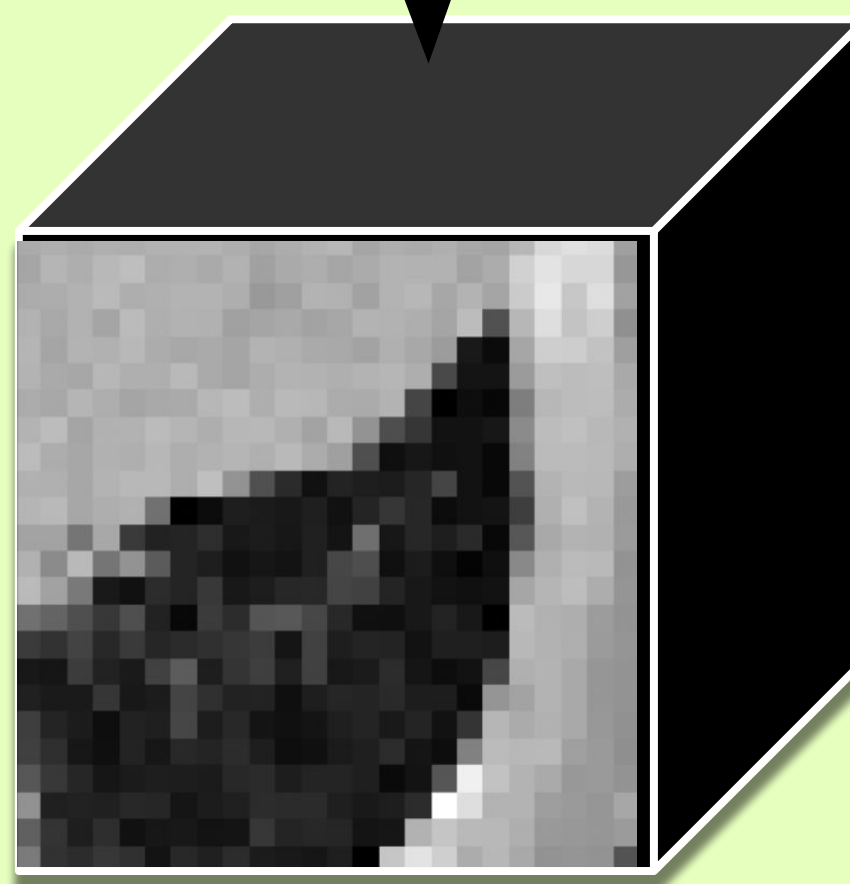
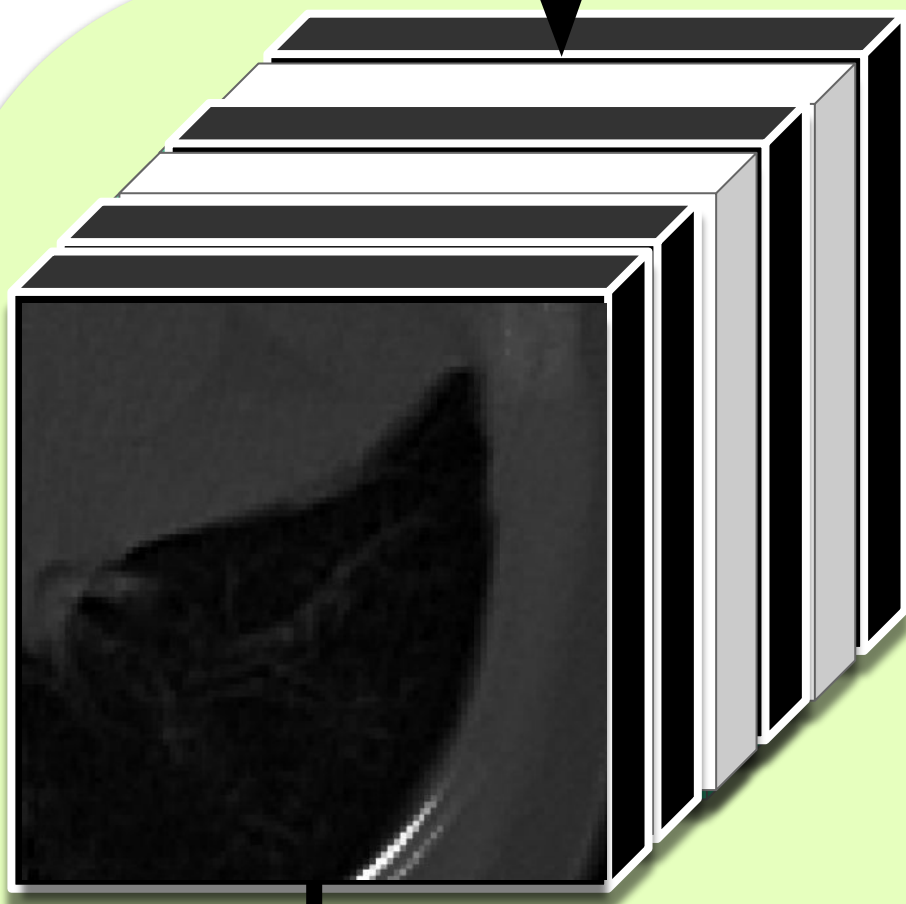


3D Medical Volume

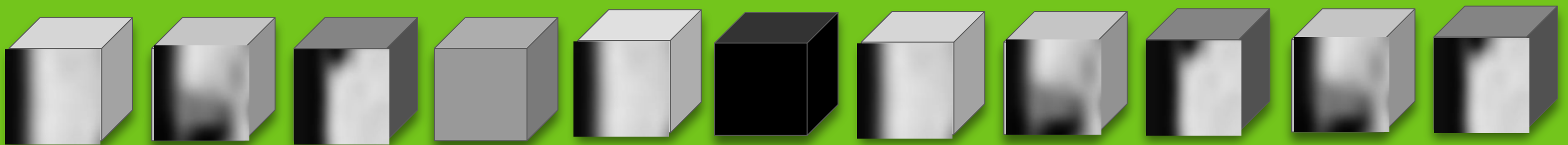
Local Masking

Downsampling

Add Noise



Disrupted Tokens



Transformer Encoder

Reconstruction

Cross Modal Contrastive Learning