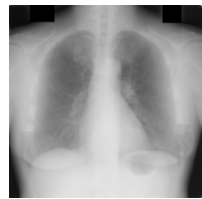
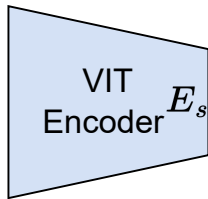


(Stage I)

Input Bone Suppressed CXR

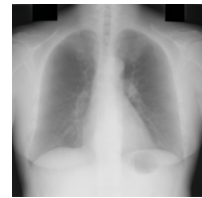


I_s



Multi-head
Codebook Attention
Learning

G_s VIT
Decoder



I_{rec}

CNN
Discriminator

real/fake

f	r	f	r
f	f	r	f
r	f	r	f
f	r	r	r

Output Bone Suppressed CXR

MSE

(Stage II)

Input Bone Shadow CXR



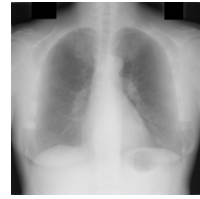
I_b

Cross-Covariance
Attention Blocks
Network

VIT
Encoder E_b

Multi-head
Codebook Attention
Learning

G_s VIT
Decoder



I_{res}

Output Bone Suppressed CXR