A serious man

Text Encoder

Text Features

Decoupled Cross-Attention

U-Net

Cross Attention

Cross Attention

Image Encoder

Self-Attention

Lin

LN

Images Features

Image Encoder

Image Encoder

Image Encoder

Frozen Modules

Trainable Modules