

Garment Image



Image
Encoder



Projection



VAE
Encoder



...

Garment UNet



...

Dressing UNet



...

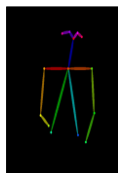
VAE
Decoder



Denoised Image



Optional Conditions



ControlNet



IP-Adapter

Noisy Image



VAE
Encoder



...

Caption

"A woman wearing red short sleeves and black pants against a white background."

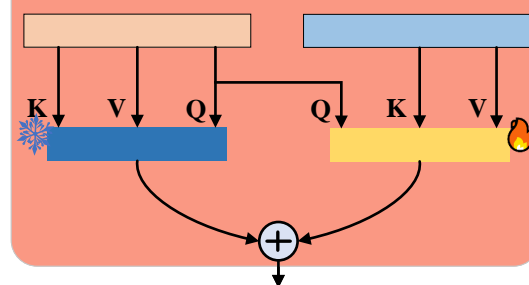
Text
Encoder



Hybrid Attention

Latent Noise

Latent Garment



Self Attention



Cross Attention



Hybrid Attention



Frozen Module



Trainable Module