

Adversarial Attack



men in suits are
walking in a park

Prediction: entailment

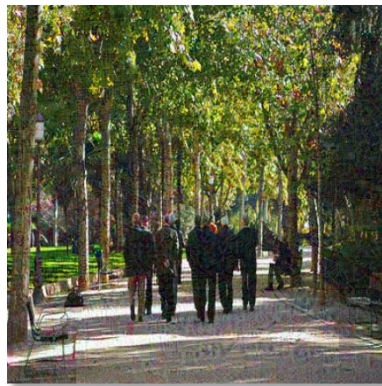
**Benign
Samples**

Image
Encoder

Text
Encoder

Transformer
Encoder/
Decoder

Victim Model



men in suits are
walking in a beautiful

Prediction: contradiction

**Adversarial
Samples**