In this paper, we study image captioning as a conditional GAN training, proposing both a context-aware LSTM captioner and co-attentive discriminator, which en