DESCRIPTION.

In paper “AttnGAN: Fine-Grained Text to Image Generation with Attentional Generative Adversarial Networks” (text: http://openaccess.thecvf.com/content_cvpr_2018/papers/Xu_AttnGAN_Fine-Grained_Text_CVPR_2018_paper.pdf code: https://paperswithcode.com/ paper / attngan-fine-grained-text-to-image-generation) a state-of-the-art solution for generating images given text description was suggested. The paper suggests an attention-based approach with all the word in the input query being processed separately, which may be not that stable in a case of noisy text.

In paper “A Generative Adversarial Approach for Zero-Shot Learning from Noisy Texts” (text: http://openaccess.thecvf.com/content_cvpr_2018/papers/Zhu_A_Generative_Adversarial_CVPR_2018_paper.pdf code: https://github.com/EthanZhu90/ZSL_GAN, the authors suggest a GAN that can handle noisy inputs. However, they use a single module for encoding input strings.

The task is to create and implement a GAN that is based on AttnGAN and somehow uses the text encoding module from the second paper (so the images are now generated given input words and noise-reducing encoding of the input sentence presented in the second paper) .
1. You are expected to decide how to integrate the encoding module from the second paper to AttnGAN.
2. You are expected to describe the resulting architecture. It would be nice if you can motivate your choice.
3. You must implement it using PyTorch. You can use the source code of both papers and any other implementations.
4. It is not necessary to train the model and show results, however, it would be nice if you can show that it somehow works (results do not matter).

Please, share a link to your github project with a comprehensive description.