# Face Aging Using Conditional GAN

StackGAN is a two-stage network. Each stage has two generators and two discriminators. StackGAN is made up of many networks, which are as follows:

* **Stack-I GAN**: text encoder, Conditioning Augmentation network, generator network, discriminator network, embedding compressor network
* **Stack-II GAN**: text encoder, Conditioning Augmentation network, generator network, discriminator network, embedding compressor network

\*\*\* INSERT IMAGE OF STACK-GAN ARCHITECTURE \*\*\*

The text encoder network

The sole purpose of the text encoder network is to convert a text description (t) to a text embedding

(http://localhost:5001/assets/b202e2a1-36da-4851-8a99-1cf84585df1d.png). I won't train the text encoder network. I will be working with pre-trained text embeddings, but you can train yours by using the steps discussed in this paper: <https://arxiv.org/pdf/1605.05395.pdf>. The text encoder network encodes a sentence to a 1,024-dimensional text embedding. The text encoder network is common to both of the stages

The conditioning augmentation block

A **conditioning augmentation** (**CA**) network samples random latent variables http://localhost:5001/assets/d86d468c-b92c-41c9-83f1-75baf727847c.png from a distribution, which is represented as http://localhost:5001/assets/cb774d54-76ee-4c94-89a7-94e8b16c4493.png

Advantages:

* It adds randomness to the network.
* It makes the generator network robust by capturing various objects with various poses and appearances.
* It produces more image-text pairs. With a higher number of image-text pairs, we can train a robust network that can handle perturbations

STAGE-1

STAGE-II

TERMS TO READ UP ON

* diagonal covariance matrix
* upsampling