



INFORMATICS  
INSTITUTE OF  
TECHNOLOGY

INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with

UNIVERSITY OF WESTMINSTER

**Fashionable**

A Final Project Report by

Mr. Aarthif Nawaz

Supervised by

Mr. Guhanathan Poravi

Submitted in partial fulfillment of the requirements for the BEng/BSc in Software  
Engineering degree at the University of Westminster

**Date: May 2022**

## Abstract

Fashion designers encounter a lot of difficulties when crafting and designing in-vogue fashion designs. As a result, many fashion-design productions don't come large-scale to the market and are below standard. These problems affect the demand and supply chain of the fashion design industry.

These problems can be addressed with the usage of text-to-image synthesis. Text-to-image synthesis is the process of transforming text descriptions into high-quality two-dimensional images. After critically analyzing the existing systems, text descriptions with a large number of words that transform into images have not yet been identified in the text-to-image synthesis domain. So, the author has decided to bridge the gap by building a novel algorithm using Attn generative adversarial networks ensembled with a contrastive learning approach to synthesize fashion-design-based descriptions to high-quality fashion designs.

This system is developed using deep learning, following a multi-level architecture of GAN networks. An image-text encoder is simulated and trained to emphasize the words provided in a text description and make it semantically consistent with the images. Additionally, during training, contrastive loss of the image and text is computed to minimize the distance of textual descriptions related to the same image and maximize those related to different images. an Attn GAN network is employed to train the text description. After training a maximum of 800 epochs, the GAN model was able to generate images for a variety of text descriptions for classes (Shirts, Trousers, Blazers, Shorts & Tops and Dresses). Also, with the use of ESRGAN trained on the FashionGen dataset, the final image that was generated was of good resolution.

**Keywords:** Text-to-Image synthesis, Image Generative Models, Computer Vision, Generative Adversarial Networks, PyTorch

## Subject Descriptors

1. Computing Methodologies >> Machine Learning >> Neural Networks
2. Computing Methodologies >> Machine Learning >> Machine Learning Algorithms >> Ensemble Methods
3. Computing Methodologies >> Artificial Intelligence >> Computer Vision >> Computer Vision Representation >> Image Representation