Generative Fashion for Indian Clothing

Harshil Jain, Rohit Patil, Utsav Jethva, Ronak Kaoshik, Shaurya Agarawal, Ritik Dutta, Nipun Batra

Indian Institute of Technology Gandhinagar, Gujarat, India

{jain.harshil,rohit.patil,jethva.utsav,kaoshik.ronak,shaurya.agarawal,ritik.dutta,nipun.batra}@iitgn.ac.in

ABSTRACT

Deep learning-based innovations, particularly GANs, have recently shown great success in fashion modelling for various use cases such as pose and face generation. A famous work, FashionGAN[1], can generate images with modified clothing as per natural language description and uses the DeepFashion dataset, which primarily contains clothing styles of the Western countries. Currently, no dataset caters to Indian style and clothing. Hence, we present a dataset of 12k images and descriptions pertaining to the Indian culture as well as a baseline approach with this work. Deep learning-based innovations in the Indian Fashion context are a relatively new area of research, and we hope our work will be a starting point for other researchers. Code and Dataset: https://github.com/ronakkaoshik42/Generative_fashion

1 BACKGROUND

Generative Adversarial Networks have shown impressive results for various tasks like super-resolution of an image, neural style transfer, etc. Our work deals with the modification of the clothing of a person as per a textual description. We use a two-step GAN architecture along with the input segmentation maps to achieve the same. Deep learning-based innovations in the Indian fashion space will be a boon, considering the vast and diverse community of fashion designers. Our work establishes a baseline for the same.

2 DATASET CURATION

We collected our dataset from websites such as Amazon[1] and Myntra[2], which are leading fashion e-commerce websites in India having millions of clothing listed on their platform using Selenium library. We scraped a total of 12304 images along with their description from the two above mentioned sources. Our dataset consists of 6459 images of men and 5845 images of women in various outfits. We collected images with numerous clothing styles, including Indian ethnic wear and modern western style. Table 1 extensively describes our dataset along with the statistics.

---

[1] www.amazon.in
[2] www.myntra.com

Men		Women	
Category	Count	Category	Count
Sherwani	404	Kurta-Kurtis	826
Kurta	1766	Sarees	860
Formal	870	Lehenga	936
Casual Shirts	868	Dress	896
T-Shirts	902	Shirt	621
Ethnic Mix	1649	Ethnic Mix	1706
Total	**6459**	**Total**	**5845**

Table 1: Data distribution among various categories for both the genders

3 METHODOLOGY AND ARCHITECTURE

The input includes a picture of a person and a textual description of the desired output as in FashionGAN [1], we have to generate an image of the same person (with preservation of pose and body structure) wearing the outfit described in the text.

During training, the same image acts as the input and output. To preserve the person's overall body structure, we extracted out segmentation maps S_0 using Self Correction Human Parsing [2] of each image I_0 in our database using the existing trained model of Self-correction Human Parsing. These segmentation maps consist of 15 different labels. The textual description is represented as Word2vec word embeddings w for each image.
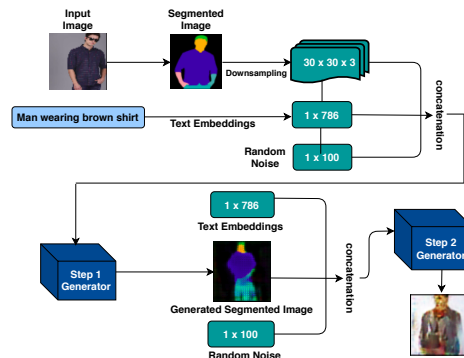

Figure 2: Proposed Architecture

Thus, given the description w and segmentation map S_0 of the image I_0, we have to generate image I_{out}. There are two steps for the framework, generating a segmentation image and, from that segmentation image, generate a coloured image of the person. So the two generators are as follows:

$$G_{seg}(z_1, S_{0_{downsampled}}, w_1) \rightarrow \tilde{S} \tag{1}$$

$$G_{img}(z_2, \tilde{S}_{downsampled}, w_2) \rightarrow I_{out} \tag{2}$$

G_{seg} is the generator in step-1 and G_{img} is the generator in step-2, z_1 and z_2 are noise vectors, w_1 and w_2 are text embeddings.

REFERENCES

[1] Yi Rui Cui, Qi Liu, Cheng Ying Gao, and Zhuo Su. 2018. FashionGAN: Display your fashion design using Conditional Generative Adversarial Nets. *Computer Graphics Forum* (2018). DOI:http://dx.doi.org/10.1111/cgf.13552
[2] Peike Li, Yunqiu Xu, Yunchao Wei, and Yi Yang. 2019. Self-Correction for Human Parsing. (2019). arXiv:cs.CV/1910.09777