World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(REVIEW ARTICLE)

Check for updates

Fashion image generation using generative adversarial neural network

P. Kamakshi Thai [1], Sai Jayanth Bandaru [2,*], Abhishek Sharma [2] and Akshay Devala [2]

[1] *Assistant Professor of Department of CSE(AI&ML) of ACE Engineering College.*
[2] *Students of Department CSE(AI&ML) of ACE Engineering College.*

World Journal of Advanced Research and Reviews, 2025, 25(01), 850-853

Publication history: Received on 03 December 2024; revised on 08 January 2025; accepted on 10 January 2025

Article DOI: https://doi.org/10.30574/wjarr.2025.25.1.0123

Abstract

Fashion image generation is a significant challenge at the intersection of artificial intelligence (AI) and creative industries, with applications in design, e-commerce, and virtual try-on systems. Conditional Generative Adversarial Networks (CGANs) extend the capabilities of standard GANs by allowing control over generated content based on specified conditions, such as clothing type, color, or texture. This Study investigates the use of CGANs for generating high-quality, attribute-specific fashion images. The study includes designing a CGAN architecture, training the model on the Deep Fashion dataset, and optimizing performance through rigorous experimentation

Keywords: Fashion image generation; Generative Adversarial; Conditional Generative Adversarial Networks (CGANs); CGAN Architecture

1. Introduction

Fashion is a dynamic industry that thrives on creativity, innovation, and consumer trends. With the rapid digitization of the sector, there is growing interest in leveraging AI to enhance various processes, such as automated design, personalized recommendations, and virtual try-on systems. Image generation, in particular, offers unique opportunities to explore designs and styles in a virtual environment, reducing reliance on physical prototypes. Generative Adversarial Networks (GANs), introduced by Goodfellow et al., have revolutionized the field of image generation by producing visually realistic outputs. However, traditional GANs lack control over the specific features of generated images.

2. Literature survey

The increasing demand for innovative and personalized fashion design has driven rapid advancements in machine learning (ML) and Generative Adversarial Networks (GANs). Researchers have explored various techniques for fashion image generation, each addressing distinct challenges and offering unique contributions to the field.

Pernuš et al. (2023) introduced FICE, a text-conditioned fashion image editing model that enhances GAN inversion with semantic, pose-related, and image-level constraints. Their approach produced highly realistic edited images, outperforming existing methods. However, the study did not evaluate diverse fashion styles and overlooked biases in textual descriptions. Future work was suggested to expand datasets and address input bias for more robust editing.

Pandey and Savakis (2019) proposed Poly-GAN, a multi-conditioned GAN architecture designed for fashion synthesis tasks such as virtual try-on. Their model demonstrated superior results in generating garment images aligned with target poses. However, challenges included limited generalizability to unconventional poses and scalability issues with

* Corresponding author: Sai Jayanth Bandaru.

large datasets. Future research was recommended to optimize pose adaptability and improve scalability for broader applications.

Li et al. (2022) developed M6-Fashion, a two-stage framework integrating style knowledge and multi-modal control for high-fidelity fashion image generation and editing. While M6-Fashion excelled in generating diverse and detailed designs, it faced difficulties in non-autoregressive generation and real-time deployment. Future directions include improving generation speed and exploring applications in real-time design systems.

Raffiee and Sollami (2020) presented GarmentGAN, a model for realistic garment transfer by separating body features from clothing. Their method achieved impressive results in virtual try-on, but limitations included occasional artifacts and challenges in handling complex poses. Future studies were suggested to enhance garment alignment and reduce artifacts in dynamic scenarios.

Shen et al. (2020) introduced GD-StarGAN, an improved version of StarGAN tailored for garment design. The model enabled multi-domain image-to-image translation, generating garment images with enhanced texture quality. However, the study lacked evaluations on dynamic garment transformations. Future research was recommended to improve versatility in texture adaptation and domain transfer.

Other notable studies explored diverse approaches and techniques in fashion image generation. For instance, deep learning (DL) models like StyleGAN, CycleGAN, and BigGAN were applied to generate high-resolution and realistic fashion images. These methods achieved impressive results but required extensive labeled data and high computational resources. Virtual try-on systems paired with pose estimation models and UNet architectures enabled realistic garment transfer, though challenges with body pose variability and self-occlusions persisted. Hybrid models combining GANs and Variational Autoencoders (VAEs) were explored for controlled image generation, achieving high accuracy but at the cost of increased computational complexity. Ensemble techniques such as stacking GANs and multi-stage training enhanced performance but required significant data preprocessing and fine-tuning.

Overall, the literature showcases a variety of methods, each contributing to specific aspects of fashion image generation. Future research emphasizes the need for diverse, high-quality datasets, lightweight models for resource-constrained environments, and robust techniques for improving the generalization and interpretability of GAN-based systems.

2.1. Architecture

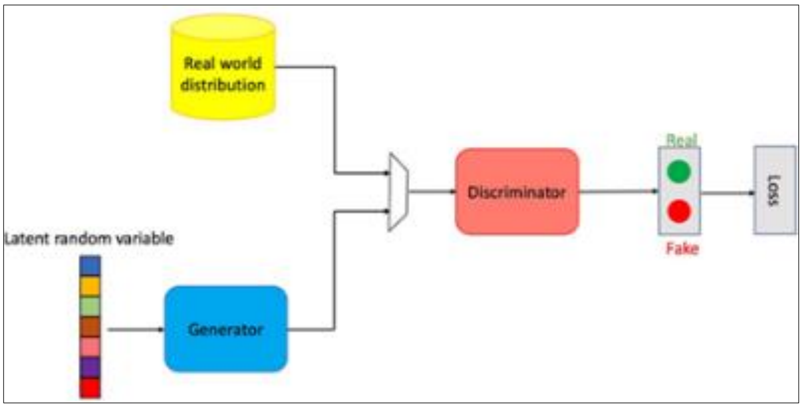

Figure 1 Generative Adversarial Network (GAN) Architecture

The Architecture depicts a Generative Adversarial Network (GAN), a powerful deep learning architecture designed to generate synthetic data that closely resembles real-world data. GANs consist of two key components: a generator and a discriminator.

The generator acts as the creative force within the GAN. It takes random noise as input and attempts to produce data samples that mimic the characteristics and distribution of real-world data. The generator's objective is to deceive the discriminator by generating data that appears authentic.

In contrast, the discriminator acts as a classifier, tasked with distinguishing between genuine data samples and those generated by the generator. It evaluates each data sample and outputs a probability score indicating whether the sample is real or fake. The discriminator's goal is to accurately identify the origin of each sample, thereby preventing the generator from fooling it.

The training process involves an adversarial game between the generator and discriminator. The generator produces synthetic data, which is then evaluated by the discriminator. Based on the discriminator's assessment, both the generator and discriminator update their internal parameters using backpropagation to improve their performance. This iterative process continues until the generator can produce data that is indistinguishable from real data by the discriminator.

3. Proposed Methodology

Conditional GANs (CGAN) involves leveraging the power of GANs while conditioning the image generation process on specific attributes to create more targeted and realistic fashion images. The first step is to collect and preprocess a suitable fashion dataset, such as Deep-Fashion or Fashion-MNIST, which includes images of various clothing items along with associated labels or attributes. The images are normalized to a fixed size and pixel range, while the conditional data, such as garment type, color, or any other relevant feature, is encoded as categorical labels or embeddings. Conditional Generative Adversarial Networks (CGANs) are an extension of traditional Generative Adversarial Networks (GANs) that incorporate additional information such as labels, features, or textual descriptions into the data generation process.

3.1. CGAN Architecture

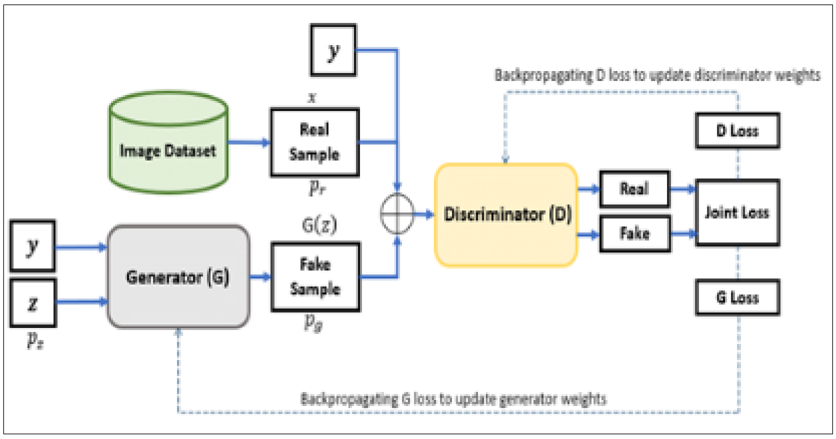

Figure 2 Conditional Generative Adversarial Network (CGAN) Architecture

The Architecture illustrates a Conditional Generative Adversarial Network (CGAN), a powerful deep learning architecture. CGANs comprise two key components: a generator and a discriminator. The generator acts as the creative force, taking random noise as input and attempting to generate synthetic images that closely resemble the real-world data. In our case, the generator aims to produce realistic fashion images, such as clothing, shoes, or accessories.

The discriminator, on the other hand, plays the role of a judge. It evaluates both real images from the dataset and synthetic images generated by the generator. The discriminator's objective is to accurately distinguish between these two sources, determining which images are authentic and which are generated.

3.2. Training Process and Loss Functions

The generator loss generally exhibits a decreasing trend over the 30 epochs, suggesting that the generator is improving its ability to generate synthetic data that can effectively deceive the discriminator. However, there are noticeable fluctuations and even slight increases in the loss at certain points, which is a typical characteristic of GAN training. This dynamic behaviour reflects the adversarial nature of the training process, where the generator and discriminator are constantly adapting to each other's strategies.

Similarly, the discriminator loss also displays fluctuations but tends to stabilize at a relatively low level. This indicates that the discriminator is becoming increasingly adept at distinguishing between real and generated data. The fluctuations in the discriminator loss are also a natural consequence of the adversarial training process, as the discriminator strives to maintain its accuracy in the face of the generator's evolving capabilities.

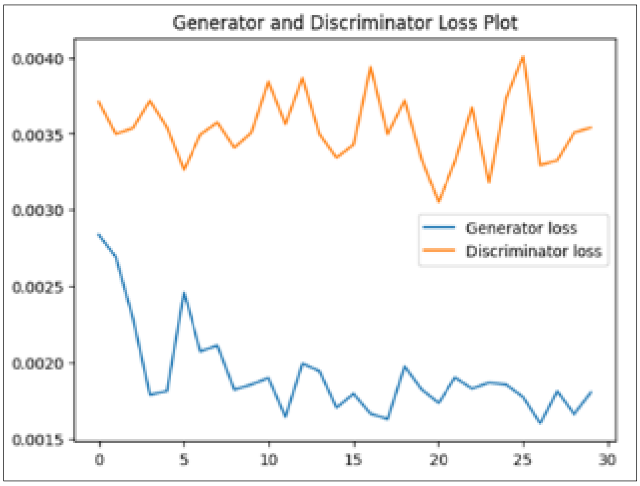

Figure 3 Loss Plot Diagram of CGAN Training

4. Conclusion

In this study, we explored the use of Conditional Generative Adversarial Networks (CGANs) for fashion image generation using the Fashion MNIST dataset. CGANs demonstrated the ability to generate recognizable fashion items conditioned on specific clothing categories, such as T-shirts, trousers, and dresses. However, the quality and diversity of the generated images were limited by the dataset's low resolution and grayscale format. The lack of colour information and the limited variety of clothing styles constrained the model's ability to create highly detailed and realistic fashion images. Despite these challenges, CGANs proved to be a promising tool for generating basic fashion items, suggesting potential for improvement through the use of advanced architectures and higher-resolution datasets in future work.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] FICE: Text-Conditioned Fashion Image Editing with Guided GAN Inversion (Pernuš et al., 2023)

[2] Poly-GAN: Multi-Conditioned GAN for Fashion Synthesis (Pandey and Savakis, 2019)

[3] M6-Fashion: High-Fidelity Multi-modal Image Generation and Editing (Li et al., 2022)

[4] GarmentGAN:Photo-realistic Adversarial Fashion Transfer (Raffiee and Sollami, 2020)

[5] GD-StarGAN: Multi-domain Image-to-Image Translation in Garment Design (Shen et al., 2020)