Proceedings of CONF-SPML 2026 Symposium: The 2nd Neural Computing and Applications Workshop 2025
DOI: 10.54254/2755-2721/2026.TJ29487

The Performance Difference Between Generative Adversarial Network and Diffusion Models in the Garment Synthesis

Chenghuan Piao

FOSE, Univeristy of Nottingham, Ningbo, China
scycp4@nottingham.edu.cn

Abstract. As the development of the Artificial Intelligence (AI) technology, it is more important for the garment and fashion industry to apply AI to their own products, specially design aspect. The increasing number of fashion and garment companies attach significance to the different models because it can reduce lots of time and money during the manufacture. As a result, this essay will introduce the performance comparison between two popular and effective networks: the Generative Adversarial Networks (GAN) and the Diffusion models. This article focuses on the introduction of GAN and diffusion models for garment synthesis. In this case, this study will introduce the basic concept of GAN and diffusion models respectively and introduce features of these models. Eventually, it can compare their performance between GAN and diffusion models and get the conclusion that diffusion models commonly have better performance than GANs and diffusion models have more computing resources compared to GANs.

Keywords: GAN, diffusion, performance comparison

1. Introduction

In order to pursue more support and profits, garment and fashion industry gradually adapt artificial intelligence in the manufacture. In fact, innovations driven by artificial intelligence and machine learning have become a key factor for garment and fashion companies in the current century [1]. One of the most important aspects is garment synthesis. Traditionally, it is difficult for designers to generate an appropriate scheme for garment because of complex work processes, ranging from sketching and modeling [2]. Therefore, many researchers and designers tried to find and train models for simplifying these obstacles [1]. Among different models, the generative adversarial network and diffusion models are potentially effective methods for garment related applications. As a result, it is vital to contrast the difference between two networks.

Generative adversarial network leaves a significant impression for researchers because of high performance on the fashion synthesis [3]. Similarly, researchers attach importance on the diffusion networks because they possess powerful generative ability and finish different aspect tasks [4]. In summary, both two models show their potential applications for garment related problems.

In previous problem, the garment synthesis is extended different applications according to different situations. For example, users can receive different related fashion images according to their input texts and change clothes for the given photos. As a result, this demanding situation may

© 2025 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.TJ29487

lead to ineffectiveness of common GAN models. Many researchers realized drawbacks of GANs and started to modify and improve the architecture of GAN, ranging from conditional generative adversarial networks to progressive generative adversarial networks. Similarly, diffusion models often consider other derivations, such as stable diffusion and blend diffusion. Moreover, in order to solve garment part leakage and attribute confusion caused by stable diffusion, researchers developed new process called DiffCloth to solve this problem [5]. It is also convenient to operate on the garment manufactures.

Moreover, in order to solve factual problems, researchers also consider other factors to adapt the factual applications, such as body skeletons and new data set labeled different fashion.

This paper will introduce and contrast the performance between generative adversarial and diffusion networks in different garment synthesis. It will list related different models and their most impressive features or constructions. After that, this essay will compare different models by their generative images and try to list their pros and cons. Finally, this passages will try to analyze their factual applications and future challenges and prospects.

2. Method

2.1. GAN

2.1.1. Preliminaries of GAN

Generative Adversarial Networks (GAN) are a type of deep learning model which are composed by two neural networks: the generator and the discriminator [6]. The generator takes responsibility for creating data, while the discriminator distinguishes between generated and real data. Through this adversarial training, the generator improves its ability to produce realistic outputs by fooling the discriminator. Through the zero-sum game between generator and discriminator, generative adversarial networks can be applied in many aspects. Therefore, generative adversarial networks possess more advantages than other deep learning networks [7].

For the garment synthesis, researchers often use different derivations. One of the most effective models is the conditional Generative Adversarial Networks (cGAN) and the reason is that the common GAN model cannot control the generative data modes. In the cGAN process, generative process is set based on feature vectors defined by users [1].

2.1.2. Poly-GAN

The Poly-GAN model is an innovative Generative Adversarial Network (GAN) designed for fashion image synthesis. One of the most significant features of the ploy-GAN is flexibility so that it can accept different conditions as inputs [3].

For the architecture of this model, it can be divided into two parts, generator and discriminator. For the generator, there are three main compositions in the encoder called Conv Module, ResNet Module, and Conv-Norm Module. There are f 3 convolution layers, each followed by a ReLU activation function in the Conv module. In order to strengthen the feature influence, this model set conditional inputs in every layer. The reason is that writers found that features would reduce as they propagate deeper in the network. In the Conv-Norm Module, it can connect the outputs from Resnet Modules and Conv Modules to exploit the vital difference between two outputs. While decoder adopts the traditional design which is constructed by a series of ResNet modules. Meanwhile, it allows decoder to collect samples in each block [3].

DOI: 10.54254/2755-2721/2026.TJ29487

For the discriminator, ployGAN adopted the same discriminator with super resolution GAN to reduce blurry images [3]. The decoder consists of a series of ResNet modules followed by transposed convolution to up-sample the input in each block. Due to the skip connection, the number of layers receiving information from decoder reduces.

2.1.3. BC-GAN

BC-GAN can synthesize multiple collocated fashion items at one time. Meanwhile, BC-GAN can precisely learn the mapping among upper and lower clothes and overcome the spatial non-alignment of the semantics during translation.

For the construction of BC-GAN, its generator consists of three important sections: a mapping network, an encoding network, and a synthesis network. Generator will synthesize a set of images with random noise. Meanwhile, it is vital for the model to develop a style embedding discriminator to observe generative images.

In order to acquire more accurate labels and description of fashion garment, researchers try to construct data set called DiverseOutfits composed of different upper and lower clothing pairs. Moreover, researchers will use LPIPS to avoid high visual similarity between two garments. However, based on this rule, the richness of dataset also demonstrates that most clothes can match the appropriate partner [8].

2.1.4. Design-AttGAN

Design-AttGAN is a derivation of AttGAN for solving poor performance in the garment synthesis [1]. The reason for this situation is that original AttGAN model has a negative influence on classification tasks. While classification loss is trained as an individual target function int the design-AttGAN. This feature allows this model to generate more samples.

This model was tested by using the fashion dataset which concludes 22 binaries labeled and open-source code provided by TensorFlow [1].

Meanwhile, this model tried to connect the corresponding labels with clothes categories. To ensure the balance of different kinds of garments, researchers narrowed the data set [1].

2.2. Diffusion model

2.2.1. Preliminaries of diffusion model

A diffusion model is a generative model that gradually adds noise to data during training and then reverses this process to generate new samples. The work process can be divided into two parts the forward diffusion process and the reverse sampling process. It starts by defining a forward process to gradually increase the data with increasing noise in some time. The model is trained to learn the reverse process, denoising the data to track the original distribution. This approach allows the model to generate high-quality samples by iteratively improving noisy inputs. Moreover, it demonstrates an impressive performance in the garment synthesis.

2.2.2. FashionSD-X

FashionSD-X achieves a new stable diffusion low rank adaptation (LoRA) which is a parameter-efficient training method [4]. LoRA can reduce the weight if training parameters by freezing the original weights and learning rank decomposition matrices.

DOI: 10.54254/2755-2721/2026.TJ29487

For the selection of data set, researchers adopted the Fashion Try-On Datasets. The reason is that models training needs to generate fashion-centric images using multimodal conditions.

In order to prove the adherence between generative images and text, this program uses the CILP which is a transformer-based vision-language model. Based on these situations, researchers mention different standards to judge models, ranging from sketch similarity to surveys.

2.2.3. DiffDesign

DiffDesign based on the WÜRSTCHEN architecture, incorporates the Expert Garment Knowledge Enhancement Module. Meanwhile, it tries to finetune parameters to enhance semantic alignment and generative garment attributes [9].

Because traditional text-to-image models, ranging from Stable Diffusion to Stable Cascade, frequency falters in the specific domain. To solve this barrier, the program integrates ERNIE-ViLG to strength the fine-grained semantic perception of the model.

In order to ensure high quality garment images, this program extract data from polypore and corresponding garment categories. Additionally, researchers acquire labels by BLIP2 model a target detection model (Grounding DINO).

After comparison with other models, this program claims that these models can supreme most of existing methods in the image quality, design accuracy, and attribute matching.

3. Comparison, applications and discussion

3.1. The performance comparison between diffusion models and GAN

In this section, it will use the generative results and analyze their advantages and disadvantages to explain them.



Figure 1. Design-AttGAN for different tasks [1]

DOI: 10.54254/2755-2721/2026.TJ29487

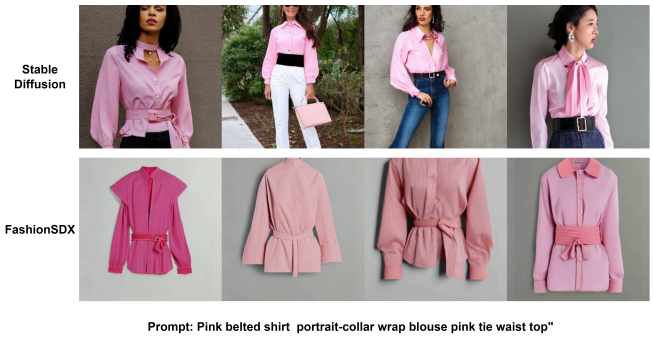

Figure 2. The comparison between FashionSDX and stable diffusion [4]

As above images shown in Figure 1 and Figure 2 demonstrated, there are still some fluctuations of image quality for the Design-attGAN. The reason may be that attGAN does not fit in garment synthesis. AttGAN cannot almost protect any pattern on the clothes due to negative impact on the reconstruction [1]. This feature will result in the performance of Design-attGAN which is origin from attGAN.

Meanwhile, its researchers also acknowledge that they need to pay more attention to the stability of models and high-resolution generative images [1].

While FashionSDX possesses more powerful ability to generate appropriate images. The reason may be that selected data set owns higher resolution and diffusion models possess more accurate and stable so that diffusion models may take more advantages than GAN in the garment synthesis [4].



Figure 3. The performance of DiffDesign among different models [9]

DOI: 10.54254/2755-2721/2026.TJ29487



Figure 4. The performance of BC-GAN [8]

Another comparison between GAN and diffusion models is between Diffdesign and BC-GAN shown in Figure 3 and Figure 4. They both generate the target image according to given input. BC-GAN can generate images by given clothes images at one time. It can identify the given input whether the upper clothes or not and provide corresponding multiple images. While BC-GAN will generate its result by input texts.

Meanwhile, it is apparent that GANs are currently faster in synthesizing images and use less resources compared with diffusion models during the training phase [9]. For the experimental environment, diffdesign use more number of gpus (a100), while BC-GAN use a1000. This illustrate that training of diffusion models need better resources.

Additionally, both generative images show that it is important to pay attention to the data set. The reason is that it is difficult to find an appropriate data set on the Internet to fit in research demands. As a result, it is necessary to process data set to pursue higher resolution and accurate results.

3.2. Applications

Typically, traditional garment creation needs lots of labor and it takes lots of time and cost to create new clothes. However, it is likely that Ai models can reduce cost and improve efficiency.

UCLA and Amazon cooperate to develop new model called garment dreamer which is a diffusion-based model. This model exploit 3D Gaussian Splatting(3DGS) guidance to construct 3d grids to ensure the consistency between garment deformation and texture synthesis.

Meanwhile, this model introduces a garment augmentation module and employs Neural Texture Fields (NeTF) composed with Variational Score Distillation (VSD) to generate diverse and more specific details [2].

3.3. Challenges and prospects

This paper introduces some architectures of GAN and diffusion models and tries to use four aspects to explain the differences and try to explain the pros and cons between GAN and diffusion models.

There are some challenges to contrast different models. First of all, different parameters may result in ineffectiveness of some rules. For example, if this study supposes diffusion models need more resources and time than GAN in the garment synthesis. The complex structure is likely to

DOI: 10.54254/2755-2721/2026.TJ29487

become the most significant barrier because researchers cannot know the corresponding relationship between parameters of two models.

As for prospects, this article will expect that increasing number of models can be applied by different software, especially shopping software [10]. In this case, users can design their own scheme by garment synthesis models to the factories. This may change people's custom in purchasing clothes.

4. Conclusion

In summary, this paper lists different GAN and diffusion networks and tries to compare their performance through their outputs, and it is implicit to observe that diffusion models needs more resources. After that, this article claims that diffusion commonly has a better performance in the garment related applications. Moreover, there is a top priority for the development of garment synthesis. It is important for other researchers to provide a useful dataset for training. The reason is that one of the most vital difficulties of garment synthesis is shortage of data sets. Meanwhile, this paper observes that some data sets lack the abundant categories of different clothes. This will result in the fact that the model cannot be applied in the factual life. As a result, it is urgent to seek a data set to reduce the training obstacles.

References

[1] Chenxi, T. and Mohsen, M. (2020). Garment Design with Generative Adversarial Networks [online] Available at: https: //arxiv.org/abs/2007.10947 [Accessed 29 September 2025]

[2] Li, B et al. (2025), GarmentDreamer: 3DGS Guided Garment Synthesis with Diverse Geometry and Texture Details, 2025 International Conference on 3D Vision (3DV), Singapore, Singapore, pp. 1416-1426, doi: 10.1109/3DV66043.2025.00134.

[3] Nilesh, P. and Andreas, S. (2020) Poly-GAN: Multi-conditioned GAN for fashion synthesis. Neurocomputing. [online] 414: pp.356-264 Available at: https: //www.sciencedirect.com/science/article/pii/S0925231220312625 [Accessed 29 September 2025]

[4] Abhishek, K. and Ioannis, P. (2024) FashionSD-X: Multimodal Fashion Garment Synthesis using Latent Diffusion Available at: https: //arxiv.org/abs/2404.18591 [Accessed 29 September 2025]

[5] Xujie, Z., Binbin, Y. and Michael, K. (2023) DiffCloth: Diffusion Based Garment Synthesis and Manipulation via Structural Cross-modal Semantic Alignment [online] Available at: https: //arxiv.org/pdf/2308.11206 [Accessed 29 September 2025]

[6] Ian, G., Jean, P. and Mehdi, M. (2014) Generative Adversarial Nets Available at: https: //arxiv.org/abs/1406.2661 [Accessed 29 September 2025]

[7] Foroumandi, E., Gavahi, K. and Moradkhani, H. (2024) 'Generative Adversarial Network for Real-Time Flash Drought Monitoring: A Deep Learning Study', Water resources research, 60(5). Available at: https: //doi.org/10.1029/2023WR035600 [Accessed 29 September 2025].

[8] Zhou., D et al. (2024), BC-GAN: A Generative Adversarial Network for Synthesizing a Batch of Collocated Clothing, in IEEE Transactions on Circuits and Systems for Video Technology, vol. 34, no. 5, pp. 3245-3259, doi: 10.1109/TCSVT.2023.3318216.

[9] Wu., S. (2025) DiffDesign: A diffusion model using garment Knowledge-Enhanced for Fashion Design Synthesis, ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Hyderabad, India, pp. 1-5, doi: 10.1109/ICASSP49660.2025.10889966.

[10] Lyu, Y., Wang, Z., Ye, Q., Sun, Y., & Chao, J. (2025). Body data-driven garment pattern construction in digital fashion innovations: a review. Textile Research Journal, 00405175251352800.