YAŞAR UNIVERSITY

GRADUATE SCHOOL

PHD THESIS

ARTIFICIALLY INTELLIGENT

FASHION DESIGNER AND EVALUATOR

WITH GENERATIVE ADVERSARIAL NETWORKS

CANER KIVANÇ HEKİMOĞLU

THESIS ADVISOR: PROF. (PHD) MEHMET SÜLEYMAN ÜNLÜTÜRK

PHD IN COMPUTER ENGINEERING

BORNOVA / İZMİR
08 2024

JURY APPROVAL PAGE

We certify that, as the jury, we have read this thesis and that in our opinion, it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy.

Jury Members:	**Signature:**
Prof. (PhD) Mehmet Süleyman ÜNLÜTÜRK Yaşar University	.....
Assoc. Prof. (PhD) Korhan KARABULUT Yaşar University	.....
Assist. Prof. (PhD) Ayşe Övgü KINAY Dokuz Eylül University	.....
Assoc. Prof. (PhD) Mete EMİNAĞAOĞLU Yaşar University	.....
Prof. (PhD) Aybars UĞUR Ege University	.....

Prof. (PhD) Yucel Ozturkoglu

Director of the Graduate School

ABSTRACT

ARTIFICIALLY INTELLIGENT FASHION DESIGNER AND EVALUATOR WITH GAN

Hekimoğlu, Caner Kıvanç

PhD, Computer Engineering

Advisor: Prof. (PhD) Mehmet Süleyman Ünlütürk

August 2024

Recently, the fashion industry has been incorporating advanced technologies more and more in order to satisfy the needs of a varied and competitive market. This doctoral thesis investigates the application of Generative Adversarial Networks (GANs) as a fashion design and assessment method. The objective of the study is to create an advanced artificial intelligence system that can produce fashion images of high-quality and realism. This system will enhance the design process and elevate the virtual shopping experience.

The study focusses on various approaches of GAN, including CycleGAN, Neural Transfer, and StyleGAN, to improve different aspects of fashion design, such as the transformation of clothes and patterns. The effectiveness of these models is evaluated through detailed computational experiments, demonstrating their ability to revolutionize the creative process in fashion design by providing innovative and efficient solutions.

This thesis showcases notable progress in automating the processing of fashion images, providing designers with powerful tools to explore novel ideas and visualize concepts without the necessity of physical prototypes. Integrating GANs not only speeds up the design process but also decreases expenses related to material waste and extended prototyping. By using extensive datasets of past designs and fashion trends, the AI-driven system creates original and cutting-edge fashion ideas, encouraging innovation and helping designers stay ahead of the competition.

Furthermore, this study aims to meet the increasing need for sustainability in the fashion sector by reducing material waste through the use of digital sample production. The utilization of AI technology is in line with current environmental objectives,

establishing GANs as a significant contributor in promoting sustainable fashion practices. The effective utilization of GANs in this particular situation highlights their ability to not only improve artistic procedures but also contribute to more ecologically conscious fashion design practices.

Keywords: Generative Adversarial Networks, AI in Fashion, Neural Networks, Fashion Design Automation, Image Synthesis, Convolutional Neural Networks

ÖZ

ÜRETKEN ÇEKİŞMELİ AĞLAR İLE YAPAY ZEKA DESTEKLİ MODA TASARIMCISI VE DEĞERLENDİRİCİ

Hekimoğlu, Caner Kıvanç

Doktora, Bilgisayar Mühendisliği

Danışman: Prof. Dr. Mehmet Süleyman Ünlütürk

Ağustos 2024

Son zamanlarda moda endüstrisi, çeşitli ve rekabetçi bir pazarın gereksinimlerini karşılamak için ileri teknolojileri giderek daha fazla kullanmaktadır. Bu doktora tezi, Üretken Çekişmeli Ağlar (ÜÇA) kullanımını moda tasarımı ve değerlendirme aracı olarak araştırmaktadır. Çalışmanın amacı, olağanüstü kalitede ve gerçekçilikte moda görüntüleri üretebilen gelişmiş bir yapay zeka sistemi yaratmaktır. Bu sistem, tasarım sürecini iyileştirecek ve sanal alışveriş deneyimini yükseltecektir.

Çalışma, giysileri ve desenleri dönüştürmek gibi moda tasarımının farklı yönlerini mümkün kılmak için CycleGAN, Neural Transfer ve StyleGAN gibi birden fazla ÜÇA tabanlı metodolojiye odaklanmaktadır. Bu modellerin etkinliği, kapsamlı hesaplamalı deneylerle değerlendirilmekte olup, yaratıcı süreci dönüştürme kapasitelerini yenilikçi ve verimli çözümler sunarak göstermektedir.

Bu tez, moda görüntü işleme otomasyonunda önemli ilerlemeleri sergilemekte, tasarımcılara fiziksel prototiplere ihtiyaç duymadan yeni kavramları keşfetmek ve fikirleri görselleştirmek için güvenilir araçlar sağlamaktadır. ÜÇA'ların entegrasyonu, sadece tasarım sürecini hızlandırmakla kalmaz, aynı zamanda malzeme israfı ve uzun prototipleme ile ilgili maliyetleri de azaltır. Geçmiş tasarımlar ve moda trendlerinden oluşan kapsamlı veri setlerini kullanarak, yapay zeka destekli sistem, özgün ve yenilikçi moda fikirleri üretir, yeniliği teşvik eder ve tasarımcıların piyasada önde kalmalarına yardımcı olur.

Ayrıca, bu çalışma, dijital örnek üretimi yoluyla malzeme israfını azaltarak moda sektöründe sürdürülebilirlik ihtiyacını karşılamayı amaçlamaktadır. Yapay zeka teknolojisinin kullanımı, mevcut çevresel hedefleri destekler ve ÜÇA'ları

sürdürülebilir moda uygulamalarını teşvik etmede önemli bir katkı sağlayıcı olarak konumlandırır. ÜÇA'ların bu özel durumdaki etkili kullanımı, sadece sanatsal süreçleri geliştirme kapasitelerini değil, aynı zamanda daha çevre dostu moda tasarım yöntemlerine katkıda bulunma potansiyellerini de vurgulamaktadır.

Anahtar Kelimeler: Üretken Çekişmeli Ağlar, Modada Yapay Zeka, Sinir Ağları, Moda Tasarım Otomasyonu, Görüntü Sentezi, Konvolüsyonel Sinir Ağları

ACKNOWLEDGEMENTS

I would like to extend my deepest gratitude to my thesis advisor, Mehmet Süleyman Ünlütürk, for his exceptional patience and meticulous guidance throughout the entire duration of my thesis study.

I wish to convey my sincere appreciation to the members of my thesis committee, namely Mete Eminağaoğlu and Korhan Karabulut. Whose invaluable insights and mentorship significantly contributed to the enhancement of my work's quality.

I would like to express my enduring love to my family, who are always supportive, loving, motivated, and caring to me in every possible way in my life.

Caner Kıvanç HEKİMOĞLU

Izmir, 2024

TEXT OF OATH

I declare and honestly confirm that my study, titled "ARTIFICIALLY INTELLIGENT FASHION DESIGNER AND EVALUATOR WITH GAN" and presented as a PhD Thesis, has been written without applying any assistance inconsistent with scientific ethics and traditions. I declare, to the best of my knowledge and belief, that all content and ideas drawn directly or indirectly from external sources are indicated in the text and listed in the list of references.

Caner Kıvanç HEKİMOĞLU

…….

12.08.2024

TABLE OF CONTENTS

JURY APPROVAL PAGE iii

ABSTRACT v

ÖZ ix

ACKNOWLEDGEMENTS xiii

TEXT OF OATH xv

TABLE OF CONTENTS xvii

LIST OF FIGURES xxi

LIST OF TABLES xxv

SYMBOLS AND ABBREVIATIONS xxvii

1. CHAPTER: INTRODUCTION 1

2. CHAPTER: PROBLEM DEFINITION 5

3. CHAPTER: LITERATURE REVIEW 9

3.1.Background 9

3.1.1 Convolutional Neural Networks 9

3.1.2 Generative Adversarial Networks: 10

3.1.3 Image Segmentation 19

3.2 AI in Fashion 24

4. CHAPTER: DATASET SELECTION 29

4.1 Brand-Based T-Shirt Dataset 30

4.2 Segmentation applied Dataset 32

4.3 Pattern Dataset 33

5. CHAPTER: COMPUTATIONAL EXPERIMENTS & RESULTS 35

5.1 CycleGAN approach: (with background) 39

5.2 CycleGAN approach: (Maskedbackground) 43
5.3 Neural Transfer approach (Garment to Garment): 47
5.4 Neural Transfer Approach (Pattern to Pattern): 50
5.5 Neural Transfer Approach (Pattern to Garment): 53
5.6 StyleGAN approach: (With background): 55
5.7 StyleGAN approach: (Maskedbackground) 59
6. CHAPTER: CONCLUSION AND FUTURE WORK 65
7. REFERENCES 69
APPENDIX 73
APPENDIX 1.RANDOM OUTPUTS 73
APPENDIX 2. CURATED CLOSE UPS 75

LIST OF FIGURES

Figure 3.1 Deep CNN Architectures 9

Figure 3.2 Simple GAN Architecture 10

Figure 3.3 CycleGAN Style Transfer 15

Figure 3.4 Sketch-to-Shoe project 16

Figure 3.5 Generated People Faces 17

Figure 3.6 CNN Architecture 19

Figure 3.7 U2 -Net architecture 21

Figure 3.8 Sample Result iMaterialist Dataset: Left: Original, Middle: Mask Right: Masked Image 23

Figure 3.9 Sample Result CuTeks Dataset: Left: Original, Middle: Mask Right: Masked Image 23

Figure 3.10 Fashion Landmark Detection 25

Figure 4.1 Garment samples 30

Figure 4.2 Dataset From Cuteks Mango 32

Figure 4.3 Masked Results 33

Figure 4.4 Pattern Samples 34

Figure 5.1 Proposed GAN Architecture 38

Figure 5.2 CycleGAN First Epoch 40

Figure 5.3 CycleGAN Epoch: 15000 41

Figure 5.4 CycleGAN Epoch: 35000 42

Figure 5.5 CycleGAN Masked Epoch: 0 44

Figure 5.6 CycleGAN Masked Epoch: 20000 45

Figure 5.7 CycleGAN Masked Epoch: 40000 46

Figure 5.8 Transferring Style (Left) to Content (Middle) to create a new garment.. 48

Figure 5.9 Samples of successful images 49

Figure 5.10 Transferring Style (Left) to Content (Top) End product (bottom right cells) to create a new pattern 51

Figure 5.11 Samples of pattern-to-pattern transferred images 52

Figure 5.12 Samples of Pattern to Garment images; Transferring Style (Left) to Garment (Middle) End product (Right) to create a new style Garment 54

Figure 5.13 StyleGAN Unmasked Epoch: 0 56

Figure 5.14 StyleGAN Unmasked Epoch: 15000 56

Figure 5.15 StyleGAN Unmasked Epoch: 30000 57

Figure 5.16 StyleGAN G and D Loss Mean 57

Figure 5.17 StyleGAN G and D Loss Standart Deviation 58

Figure 5.18 StyleGAN masked Epoch: 0 60

Figure 5.19 StyleGAN masked Epoch: 15000 60

Figure 5.20 StyleGAN masked Epoch: 30000 61

Figure 5.21 StyleGAN masked G and D Loss Mean 61

Figure 5.22 StyleGAN masked G and D Loss STD 62

Figure 8.1 Random seed between 10000-10060 73

Figure 8.2 Random seed between 20000-20060 74

Figure 8.3 Masked StyleGAN30000 Seed1409 75

Figure 8.4 Masked StyleGAN30000 Seed6257 75

Figure 8.5 Masked StyleGAN30000 Seed6452 76

Figure 8.6 Masked StyleGAN30000 Seed1832 76

LIST OF TABLES

Table 3.1 GAN Algorithms 11

Table 4.1 Garment Samples by Brand 29

Table 4.2 Shirt Samples by Brand 31

Table 5.1 CycleGAN Parameters 39

Table 5.2 CycleGAN parameters masked 43

Table 5.3 Neural Transfer approach Garment x Garment parameters 47

Table 5.4 Neural Transfer approach pattern x pattern parameters 50

Table 5.5 Neural Transfer Approach Pattern x Garment Parameters 53

Table 5.6 StyleGAN parameters 55

Table 5.7 StyleGAN masked parameters 59

Table 5.8 FID50k Comparison Table for GAN Results 63

SYMBOLS AND ABBREVIATIONS

ABBREVIATIONS:

AI Artificial Information

GAN Generative Adversarial Networks

LPIPS Learned Perceptual Image Patch Similarity

FID Fréchet Inception Distance

GAN Generative Adversarial Network

CNN Convolutional Neural Network

RGB Red, Green, Blue

GPU Graphics Processing Unit

CPU Central Processing Unit

MSE Mean Squared Error

kimg Thousand of Images (In training)

1. CHAPTER: INTRODUCTION

The fashion industry today grapples with the dual imperative of catering to an increasingly diverse customer base while simultaneously navigating the demands of a highly competitive and fast-paced market (Yuan & Moghaddam, 2020). As online shopping revolutionizes how consumers engage with fashion, the need for high-quality, informative product visuals has become paramount. Virtual try-on and image synthesis techniques present a promising avenue to address these challenges, empowering fashion designers and consultants to efficiently generate and manipulate garment images to serve their clients better.

Generative Adversarial Networks (GANs) have emerged as a potent deep-learning approach for the production of photo-realistic images. Fashion designers can accelerate the ideation process and overcome cognitive biases by utilizing these models, while online retailers can improve the virtual shopping experience for their customers. The practical benefits and key challenges of diversifying fashion generation and enabling dynamic virtual try-ons are the focus of this paper, which examines the application of GANs in the context of fashion design and online retail. (Sun et al., 2019) In the realm of image generation, the field of deep learning has experienced remarkable advancements in recent years, with GANs at the forefront of this revolution. The capacity of these models to generate high-quality photo-realistic images has been demonstrated, thereby introducing new opportunities for a diverse array of applications, such as virtual fashion. A proliferation of technological innovations, including augmented reality, virtual reality, and connected fitting rooms, has fostered the rapid growth of the online fashion industry. These innovations have allowed retailers to provide a more immersive and personalized shopping experience. (Raffiee & Sollami, 2020)

Generative AI in fashion offers a notable benefit by having the ability to decrease expenses. The conventional fashion design process is highly resource-intensive, necessitating substantial investments in materials, labor, and time. Designers generate

preliminary concepts, fabricate prototypes, and engage in numerous iterations prior to completing a design. This procedure can span several months and entails significant material wastage in the event that prototypes fail to match the specified criteria. GANs expedite this procedure by rapidly producing several design iterations, enabling designers to explore various aesthetics and ideas without the necessity of tangible prototypes. This not only expedites the design cycle but also diminishes the expenses linked to manufacturing many samples. Consequently, fashion firms are able to distribute resources in a more effective manner, prioritizing the use of superior materials and sophisticated manufacturing methods instead of prolonged prototyping (Goodfellow et al., 2014).

In the realm of fashion design, creativity serves as the essential driving force. However, even the most gifted designers may encounter periods of creative stagnation. Generative AI solves this problem by offering a constant stream of novel and inventive design concepts. GANs utilize large databases of historical designs and fashion trends to develop novel and non-traditional looks that challenge conventional fashion constraints. This AI-powered methodology helps the creative process by providing novel viewpoints and amalgamations that designers may not have contemplated. It enables designers to investigate a wider spectrum of options, promoting innovation and uniqueness. The capacity to produce a wide range of designs also enables fashion firms to maintain a competitive edge by adapting to changing consumer tastes and staying up-to-date with current trends, thus assuring the continued relevance and attractiveness of their collections.

The integration of GANs in fashion design does not supplant human designers but instead alters their position. Designers may now prioritize curating and improving AI-generated outputs, saving them from spending excessive time on manual design activities. This shift enables designers to utilize their experience and creativity in choosing and improving the most outstanding concepts generated by AI, guaranteeing that the end products adhere to the brand's aesthetic and quality criteria. The integration of AI and human designers in a collaborative manner results in streamlined workflows and superior design outcomes. Designers can optimize their time and expertise by focusing on tasks that demand human judgment and creativity, such as generating new concepts for collections, formulating marketing strategies, and engaging with

consumers. The integration of AI's generative powers with human creativity yields a creative process that is more dynamic and productive.

Generative artificial intelligence (AI) is also essential in improving the overall user experience. AI-powered virtual try-on systems enable consumers to visually assess the appearance and fit of clothing on their bodies without needing a real try-on. This technique decreases the probability of returns caused by ill-fitting or unsatisfactory product appearance, enhancing overall consumer happiness and loyalty (Kim & Forsythe, 2008). Moreover, AI-powered personalization empowers fashion firms to provide customized suggestions according to individual interests and behaviors. Through the analysis of consumer data and feedback, artificial intelligence (AI) has the capability to recommend designs that are highly likely to resonate with individual customers, thereby generating a customized shopping experience. Customization at this level not only enhances sales but also fortifies the bond between the company and its clients.

The fashion sector is increasingly being examined for its environmental impact and is facing mounting pressure to implement sustainable practices. Generative AI enhances sustainability by optimizing resource utilization and minimizing waste. The capacity to produce digital samples obviates the necessity for several tangible prototypes, hence reducing material wastage. In addition, AI has the capability to enhance production processes by minimizing overproduction and ensuring that only the required numbers of clothing are manufactured. Generative AI supports the industry's move towards sustainability by encouraging the efficient utilization of materials and minimizing the environmental impact of the design process. This not only has a positive impact on the environment but also improves the brand's standing among environmentally aware consumers.

Multiple fashion firms have already started utilizing the capabilities of generative AI, showcasing its ability to revolutionize the sector. The cooperation between Tommy Hilfiger, IBM, and the Fashion Institute of Technology (FIT) resulted in the development of an AI-powered design assistant called "Reimagine Retail." This collaboration demonstrated how AI can improve creativity by offering data-driven insights and inspiration for the creation of new collections. Stitch Fix is an internet-based service for personalized fashion advice that uses artificial intelligence (AI) to

offer tailored fashion suggestions. This involves a combination of AI-generated designs and human expertise, resulting in highly individualized shopping experiences.

The ongoing development of generative AI is expected to lead to greater expansion of its applications in the fashion industry. Possible future advancements include improved customization, collaborative design platforms, and advanced virtual try-on systems. AI has the potential to utilize biometric data to produce clothing that is precisely customized, resulting in an unprecedented level of personalization. AI-driven collaborative platforms have the potential to enable real-time cooperation among designers, consumers, and AI systems, allowing for the joint development of fashion items.

To summarise, generative AI has the capacity to profoundly revolutionize the fashion sector by tackling significant obstacles and fostering creativity. AI is positioned to alter the fashion design process by decreasing expenses, fostering innovation, and advocating for sustainability. The fusion of artificial intelligence and human creativity holds the potential for a dynamic and efficient future in the fashion industry. This future enables firms to swiftly adapt to trends and consumer preferences, all while upholding rigorous standards of quality and sustainability.

The primary objective of this investigation is to optimize the creative process of ready-to-wear garment production in order to accelerate the process and reduce the time and personnel required to develop new designs. The remaining portion of the investigation is structured as follows: The problem definition is provided in Chapter 2, while the Literature and Background are explained in Chapter 3. Chapter 4 provides an introduction to our methodology and datasets, while Chapter 5 contains experiments and results from various phases, plus their respective discussions. Chapter 6 concludes the subject matter and provides a list of potential future research topics.

2. CHAPTER: PROBLEM DEFINITION

Advanced technology is increasingly necessary for the modern fashion business to optimize operations and foster innovation. An important obstacle fashion enterprises encounter is the lack of effectiveness of conventional approaches in creating and manufacturing clothing prototypes. The existing procedure, which entails human designers generating visual representations on computers and subsequently fabricating tangible prototypes utilizing diverse materials, is both time-consuming and resource-intensive. The conventional approach not only hampers the pace of work but also includes numerous checkpoints and production stages that can impede decision-making and escalate expenses.

The process of making decisions to predict demand adds extra complexity to this situation. An accurate demand prediction technique is essential for maximizing the order percentages of developed clothing samples. Nevertheless, this system necessitates a continuous influx of novel designs in order to operate efficiently. The system's performance can be hindered by the variety and inconsistency introduced by the ongoing reliance on human designers to provide these samples. In addition, decision-makers from global design firms must assess these samples to verify they satisfy their criteria, introducing further intricacy and possible postponement.

Implementing GANs provides a promising method to tackle these challenges. GANs comprise two primary elements: the generating network and the discriminator network. The generative network can undergo training using historical design data in order to acquire knowledge on the process of creating garment prototypes by designers. As a result, it is capable of producing novel and distinct designs by leveraging the acquired patterns and styles. Contrary to conventional models like Feed-forward Neural Networks or Support Vector Machines, which depend on preprocessed data and lack the ability to immediately generate images, GANs have the capability to produce authentic clothing samples from unprocessed photographs. Convolutional Neural Networks (CNNs) improve this capability by effectively processing picture

input without the need for considerable preprocessing. CNNs have the ability to extract features from photos using convolutional layers. This allows for the creation of garment patterns that are both highly realistic and intricate in detail.

Conversely, the discriminator network has a crucial role in assessing the quality of the produced samples. Its purpose is to distinguish between authentic photos and those generated by the generative network. The discriminator network contributes to enhancing the precision and authenticity of the created designs by continuously rectifying its mistakes and assimilating the feedback. The adversarial training approach guarantees that both the generator and discriminator networks mutually improve their performance, resulting in a resilient AI system that can generate and evaluate high-quality clothing samples.

The use of GANs as an AI-powered fashion designer and evaluator offers numerous benefits. First, it greatly decreases the time and resources needed to create clothing prototypes. Second, GANs, through the automation of the design process, can produce a large number of designs within a brief timeframe, facilitating quick iteration and experimentation. Rapid speed and optimal efficiency are vital in an industry characterized by rapidly changing trends, where timely responsiveness to market needs is imperative.

Furthermore, the utilization of artificial intelligence in generative design cultivates ingenuity and advancement. GANs have the ability to investigate a wider spectrum of design possibilities compared to human designers alone, resulting in the creation of distinctive and unorthodox styles that have the potential to establish novel trends. The enhanced ability to think creatively can assist fashion companies in distinguishing themselves in a highly competitive market.

Also, GANs improve the precision and reliability of demand forecasting systems. AI can enhance the effectiveness of predictive models and enable more precise forecasting by consistently generating high-quality designs that provide dependable data. This can result in improved inventory control, decreased inefficiency, and enhanced financial gains.

Next, the AI-powered method aligns with the increasing focus on sustainability in the fashion sector. The capacity to produce digital samples diminishes the necessity for tangible prototypes, minimizing material wastage and mitigating the environmental

consequences of the design process. This sustainability advantage aligns with the growing customer desire for environmentally conscious behaviors and products.

To summarise, using GANs in the fashion design process effectively tackles numerous significant obstacles linked to conventional approaches. GANs improve productivity, creativity, and sustainability by automating the production and assessment of clothing prototypes. This sophisticated AI-powered method not only simplifies processes and lowers expenses but also encourages creativity and fulfills the changing requirements of the fashion industry. The apparel industry's adoption of GANs signifies a notable advancement towards a more efficient and innovative future.

3. CHAPTER: LITERATURE REVIEW

There are various ways to use AI techniques in fashion topics. It ranges from detecting the garment sample position to creating thematic style samples and figuring out a person's likeness of buying a product to combining different garment elements in order to get a more compatible combination of the garment. We can divide this category into two parts. The first one understands the fundamental concepts.

3.1. Background

3.1.1 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are a particular type of Neural Network that have shown state-of-the-art results on various competitive benchmarks. CNN first occurred in the work of (Lecun et al., 1989) for the processing of grid-like topological data (images and time series data). After that, lots of new CNN architectures and methods are developed to process images and time series data. In Figure 3.1, Deep CNN architectures are shown. CNNs are considered one of the best techniques for understanding image content and have shown state-of-the-art results in image recognition, segmentation, detection, and retrieval-related tasks.

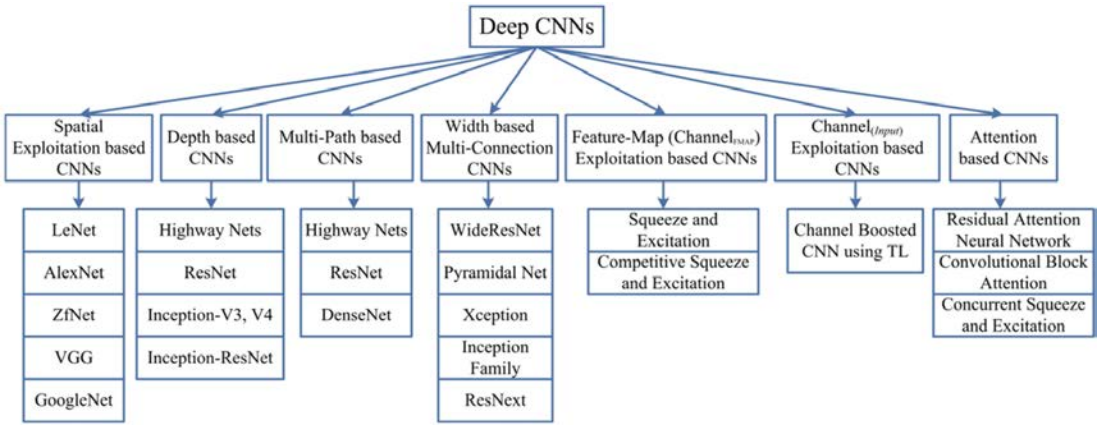

Figure 3.1 Deep CNN Architectures

Source: (Lao & Jagadeesh, 2014a)

Deep CNN networks can be used for fashion Classification. In (Lao & Jagadeesh, 2014b) work, they can retrieve the clothing style classification and clothing attribute classification from a given image that contains a clothing object in it, i.e., a model posing with cloth samples. They trained Alex Net (Krizhevsky et al., 2017), which is a convolutional network, to get promising results.

CNN also works with other models to reduce complexity. Because I used Gabor Filters for preprocessing images for my MSc thesis(Hekimoğlu, 2017). (Luan et al., 2018) the work became interesting to me. It is called the Gabor Convolutional Network. They are aiming to enhance the deep feature representations with steerable orientation and scale capacities by using Gabor filters. In that way, they reduce the complexity of incoming data to the network and thus make it easier to train.

3.1.2 Generative Adversarial Networks:

The Generative Adversarial Network (GAN), proposed in 2014(Goodfellow et al., 2014), has since become very popular in the field of generative modeling. GAN is composed of two networks, a generator G and a discriminator D (Figure 3.2). Both G and D are deep neural networks with convolutional and dense layers as appropriate. The generator G takes a random noise vector z as input and generates data $G_{(z)}$, which is intended to resemble real data. The discriminator D then receives both the generated data $G_{(z)}$ and real data x and attempts to distinguish between the two, labeling them as "fake" or "true". The generator and discriminator are trained simultaneously, with the generator trying to fool the discriminator. Once training is done, a new x can be generated by sampling from $G_{(z)}$ And using G.

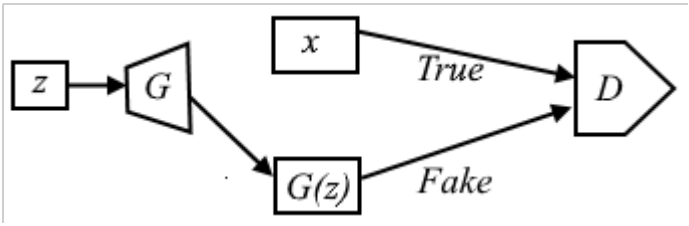

Figure 3.2 Simple GAN Architecture

Source: Author

According to (Gui et al., 2020) comprehensive review in Table 3.1, this representation categorizes various Generative Adversarial Networks (GANs) based on their

representative variants, objective functions, and training techniques. It organizes GANs into different groups, such as those focused on skills like improved GANs or structural innovations like multi-generator learning, highlighting the advancements in the field. Additionally, it emphasizes task-driven GANs, including those designed for semi-supervised learning, transfer learning, and reinforcement learning, illustrating the diverse applications of GAN technology.

Table 3.1 GAN Algorithms

GANs' Representative variants		InfoGAN [14], cGANs [15], CycleGAN [16], f-GAN [17], WGAN [18], WGAN-GP [19], LS-GAN [20]
GANs training	**Objective function**	LSGANs [21], [22], hinge loss based GAN [23]–[25], MDGAN [26], unrolled GAN [27], SN-GANs [23], RGANs [28]
	Skills	ImprovedGANs [29], AC-GAN [30]
	Structure	LAPGAN [31], DCGANs [32], PGGAN [33], StackedGAN [34], SAGAN [35], BigGANs [36], StyleGAN [37], hybrids of autoencoders and GANs (EBGAN [38], BEGAN [39], BiGAN [40]/ALI [41], AGE [42]), multi-discriminator learning (D2GAN [43], GMAN [44]), multi-generator learning (MGAN [45], MAD-GAN [46]), multi-GAN learning (CoGAN [47])
Task-driven GANs	**Semi-supervised learning**	CatGANs [48], feature matching GANs [29], VAT [49], Δ-GAN [50], Triple-GAN [51]
	Transfer learning	DANN [52], CycleGAN [53], DiscoGAN [54], DualGAN [55], StarGAN [56], CyCADA [57], ADDA [58], [59], FCAN [60], unsupervised pixel-level domain adaptation (PixelDA) [61]
	Reinforcement learning	GAIL [62]

Source: (Gui et al., 2020)

In Literature, GAN Algorithms are effectively used for generating an image from a given noise vector, which is a key point in (Goodfellow et al., 2014) work. It consists of a generator and discriminator network that try to beat each other. However, due to

the randomness of this noise vector, it can give very diverse results. That's why conditional GAN models appeared (Mirza & Osindero, 2014). They bring another perspective to making generator and classifiers conditioned on classes. This allows GANs to better represent multi-modal data.

GANs are also used in fashion design topics. In (L. Liu et al., 2019) work, unlike extant fashion mining literature, in which style is usually classified as similarity, they prefer to try to find clothing match rules based on semantic attributes according to the GAN model. Another example can be (Yildirim et al., 2019) work, which creates high-resolution images of fashion models wearing a custom outfit under an input body pose. That way, when purchasing new clothes, they can be seen on the input body and decide whether he/she wants them or not. (Kato et al., 2018) The "DeepWear" case study examines the application of DCGANs in fashion design, presenting a new feedback loop to enhance designers' efficiency and creativity. It delves into the workflow, impact, and comparison with traditional approaches, emphasizing the advantages of collaborative design involving human expertise and artificial intelligence within the fashion sector. Also, (Kato et al., 2019) study involved conducting interviews with participants who had experience in pattern-making with the brand of the learning source. The findings indicated that individuals found it easier to sketch patterns from the images generated if they achieved an average score of 2.8 or higher.

In my thesis, we need to bring out non-existent garment samples with the knowledge of historical garment sample data. It creates a new, unreal sample, which is one of the most powerful ways of GANs. The study of (Karras et al., 2018), which was called StyleGAN in June 2019, and an improved version of SyleGANv2 (Karras et al., 2019) can create non-existent faces with ground truth of other real human face images. There are lots of GAN architectures that work on image data and can get meaningful results for this thesis. Several GAN architectures will be implemented and measured for this project. That's why some metrics must be used to distinguish those architectures and show which one is superior to the other. There are lots of quantitative and qualitative metrics to measure GANs. In (Borji, 2019) work, more than 24 quantitative and five qualitative measures are discussed, and every one of them has pros and cons. At least two metrics will be used below.

Precision, Recall, and F1 Score: Those are proven and widely adopted techniques for quantitatively evaluating the quality of discriminative models. (Lucic et al., 2017)Precision measures the fraction of relevant retrieved instances among the retrieved instances, while recall measures the fraction of the retrieved instances among relevant instances. The F1 score is the harmonic average of precision and recall.

Inception Score: Inception Score (IS) is an objective metric for evaluating the quality of generated images, specifically synthetic images output by GAN models. IS is the first such criterion to have become widely adopted in the GAN literature (Salimans et al., 2016). The idea is to use the Google Inception network(Szegedy et al., 2015), which is pre-trained on the large-scale ImageNet data set to compute the class label distribution of the fake examples.

$$IS = exp\left(\mathbb{E}_{x\sim P_g} D_{KL}(p(y|x)||p(y))\right) \quad (1)$$

In Equation 1, where $x\sim P_g$ indicates images sampled from the generator's output ($p(y|x)$ is the conditional probability of class labels given an image as predicted by a pre-trained Inception network, and $p(y)$ is the marginal probability of class labels across all generated images. The D_{KL} divergence measures how different the predicted label distribution for an image is from the overall label distribution, with higher divergence indicating more distinctive images. The expected value $\mathbb{E}$ averages this divergence over all generated images, and the exponential function scales the result to a positive range, producing the IS which balances image quality and diversity.

The Frechet Inception Distance, or FID for short, is a metric for evaluating the quality of generated images and was specifically developed to evaluate the performance of generative adversarial networks. The FID score was proposed and used by (Heusel et al., 2017) in their 2017 paper titled "GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium." The FID score is used to evaluate the quality of images generated by generative adversarial networks, and lower scores have been shown to correlate well with higher-quality images. (Parmar et al., 2021)'s paper, it highlights the importance of correct image preprocessing in FID calculations and the impact of resizing and compression on the metric. It provides recommendations and a reference implementation to improve the reliability of FID evaluations.

$$FID = \left\|\mu_r - \mu_g\right\|^2 + Tr\left(\sum_r + \sum_g - 2\left(\sum_r \sum_g\right)^{1/2}\right) \quad (2)$$

In Equation 2, where μ_r and $\sum_r$ are the mean and covariance of the features for real images and μ_g and $\sum_g$ are the same for generated images. The term $\left\|\mu_r - \mu_g\right\|^2$ represents the squared distance between the means, and the rest of the equation measures the similarity of the covariances. A lower FID indicates that the generated images are more similar to the real images in terms of both content and diversity.

The FID50k metric is an adaptation of the Fréchet Inception Distance (FID) used for evaluating generative models, especially in the context of image generation. It measures the discrepancy between the distributions of real and generated images by comparing their high-level feature representations obtained from a pre-trained Inception network. The FID50k refers to the FID score computed using 50,000 samples from both real and generated datasets. This metric is widely used because it captures both the quality and diversity of generated images, making it a comprehensive evaluation tool for generative models.

$$FID50k = \left\|\mu_r{}^{50k} - \mu_g{}^{50k}\right\|^2 + Tr\left(\sum_r{}^{50k} + \sum_g{}^{50k} - 2\left(\sum_r{}^{50k} \sum_g{}^{50k}\right)^{1/2}\right) \quad (3)$$

In Equation 3, all operands are the same for generated images, all derived from 50,000 samples as FID metric. This metric captures both the quality and diversity of generated images, making it a comprehensive tool for evaluating the performance of generative models.

Learned Perceptual Image Patch Similarity, or LPIPS metric, is a perceptual similarity metric that measures the similarity between two images based on the activations of two image patches for some pre-defined network. LPIPS (Zhang et al., 2018) can be used to train GANs to create new fashion garment sample images by evaluating the quality of the generated images and comparing them to real images.

$$LPIPS = \sum_l \frac{1}{H_l W_l} \sum_{h=1}^{H_l} \sum_{w=1}^{W} \left\| w_l \odot \left(f_{h,w}^l(x_1) - f_{h,w}^l(x_2)\right) \right\|_2^2 \quad (4)$$

In Equation 4, The deep features, denoted as f, are obtained from layer l at spatial location h, w. The height and width of the feature maps at each layer are H_l and W_l, respectively, and these dimensions normalize the sum of the feature differences. The differences between the corresponding features of the two images are weighted by

learned weights w and summed across all layers and spatial locations. The L2 norm $\| \|_2^2$ is used to measure these differences, and the element-wise multiplication $\odot$ applies the learned weights. The resulting LPIPS score indicates perceptual similarity, with lower scores reflecting higher similarity.

There are a lot of GAN architectures for multiple problems in the Image synthesis area, such as image super-resolution, image completion/repair, image matting, and image-to-image translation. In our problem, we are mainly focusing on image-to-image translation. Some of the inspected image-to-image translation architectures are explained in the next part.

3.1.2.1 CycleGAN

To achieve image translation, the authors of CycleGAN(Zhu et al., 2017) used a very simple yet effective procedure. They made use of two GANs, and the generator of each GAN performed the image translation from one domain to another. The CycleGAN approach is presented with many impressive applications. In the Figure3.3 from Zhu et al. (2017) demonstrates the application of CycleGAN for style transfer, where input images are transformed into the styles of famous artists such as Monet, Van Gogh, Cezanne, and the traditional Japanese Ukiyo-e style.

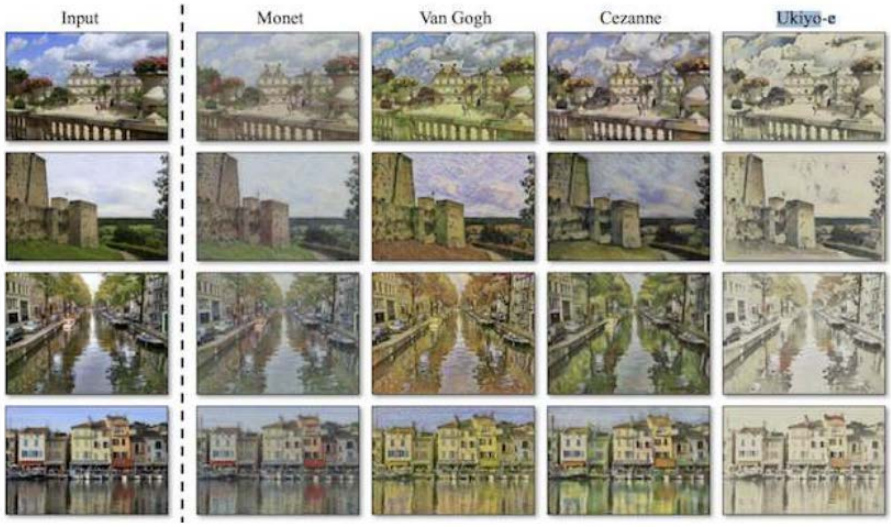

Figure 3.3 CycleGAN Style Transfer

Source: (Zhu et al., 2017)

3.1.2.2 Pix2Pix

Pix2Pix(Isola et al., 2018) focuses on a general-purpose solution to image-to-image translation problems using GAN. These networks not only acquire the ability to map input images to output images, but they also acquire the ability to learn a loss function that is used to train this mapping. This enables the application of a uniform methodology to problems that would typically necessitate distinct loss formulations. The Pix2Pix GAN has been showcased in various image-to-image translation tasks, including the conversion of maps to satellite photographs, black and white photographs to color, and sketches of products to product photographs. CycleGANs employ a similar methodology to Pix2Pix GAN for Image-to-Image Translation. However, CycleGANs differ in that they utilize unpaired images for training and incorporate an additional criterion, known as the cycle consistency loss, into their objective function. Figure 3.4 showcases the application of a model that converts line drawings (input) into realistic images of shoes (output) that closely resemble the ground truth photos. This demonstrates the model's ability to accurately generate detailed and visually convincing images from simple sketches.

Figure 3.4 Sketch-to-Shoe project

Source: (Isola et al., 2018)

3.1.2.3 StyleGan

The Style Generative Adversarial Network, or StyleGAN (Karras et al., 2018) for short, introduces a novel architecture for generative adversarial networks that allows for the unsupervised separation of high-level characteristics and random variation in generated images. This design enables precise control over image synthesis at different scales, resulting in improved image quality, interpolation capabilities, and disentanglement of underlying factors.

This proposed architecture is particularly applicable to fashion-related tasks as it offers the ability to manipulate the generation of high-level attributes such as posture and identity, along with stochastic variations, including clothing textures, patterns, and colors. With this, a robust solution for fashion-focused works by allowing refined control over image synthesis while enabling the production of diverse fashion garments of high quality. Figure 3.5 presents highly realistic images of faces that are entirely generated by a GAN. These synthetic images demonstrate the model's advanced ability to create human-like facial features that are indistinguishable from real photographs.

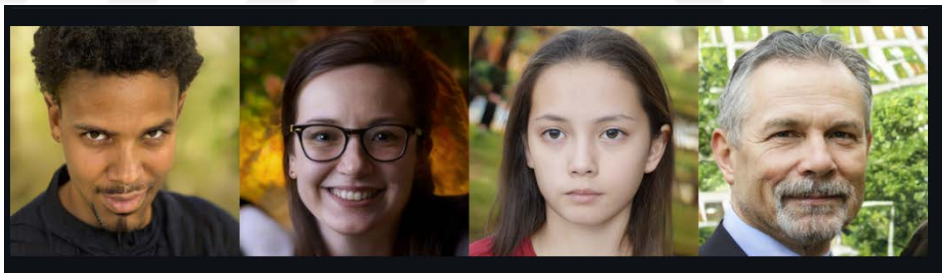

Figure 3.5 Generated People Faces

Source: (Karras et al., 2018)

3.1.2.4 A neural algorithm of artistic style

(Gatys et al., 2016) proposed a work that captures the style of artistic images and transfers it to other images using Convolutional Neural Networks (CNN). This work formulated the problem as finding an image matching both the content and style

statistics based on the neural activations of each layer in CNN. The fundamental element of style representation in their work is the Gram matrix. In simple words, a Gram matrix is a matrix created by multiplying a matrix with its transpose. (Gatys et al., 2016) do not explain gram matrices in detail, but in (Li et al., 2017) work, it is explained more clearly. Those Gram matrices can be created as feature maps with changing kernel size values. Basically, these matrices contain low-level features of an image, like lines, edges, dots, and curves, and if we perform deep convolution, we get some higher-level features like shaped objects. Different feature maps represent different features of an image, but these features are independent of each other, so they can capture objects and shapes but can't capture the style of an image style, whose occurrences are dependent on each other. A high correlation between feature maps like color, texture, and shape in the form of feature matrices and, if measured, the degree of correlation between these matrices can represent the style of an image. When feature 1 occurs with feature two, too, the correlation between feature maps can be calculated using dot products with respect to each other. Those matrices will be used to optimize CNN with these loss values. In Figure 3.6, the process of neural style transfer is illustrated, where a convolutional neural network separates and recombines the content and style representations of images. It shows how the content of one image (e.g., a town scene) is restructured with the style of another image (e.g., Van Gogh's "Starry Night") to produce a final output that blends the content and artistic style effectively.

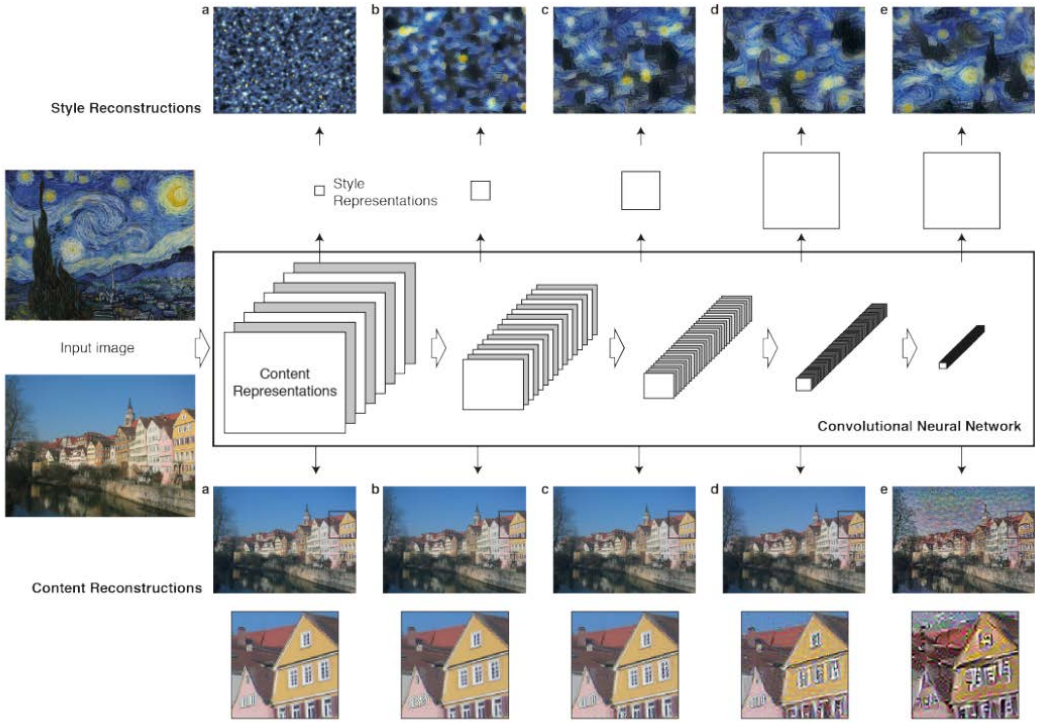

Figure 3.6 CNN Architecture

Source: (Gatys et al., 2016)

3.1.3 Image Segmentation

The fashion industry has seen a significant increase in the use of computer vision techniques to improve the online shopping experience and provide considerable economic value due to the abundance of visual information accessible. One of the most important tasks in computer vision is image segmentation, which involves dividing an image into regions or pixels that belong to different objects or categories. Image segmentation has many applications in fields such as medical imaging, autonomous driving, face recognition, and video editing. In the context of fashion, image segmentation can be used to identify and extract fashion garments from images, which can be used for various purposes, such as recommendation systems, trend analysis, and virtual try-on systems.

Semantic image segmentation is a type of image segmentation that involves grouping each pixel under a particular label. For example, any pixel belonging to a garment would be assigned under the same "garment" category. Semantic segmentation can provide a comprehensive understanding of the scene and its components, which is

useful for tasks like navigation, mapping, and localization. However, semantic segmentation is still full of challenges due to limited or sparse annotations, class imbalance, overfitting, long training time, and gradient vanishing. (Yu et al., 2023)

Convolutional neural networks (CNNs) have shown remarkable performance in semantic image segmentation. CNNs are deep neural networks that have the capability to classify and segment images. CNNs can be trained using supervised or unsupervised learning techniques, and they have been used in cutting-edge applications such as self-driving cars and medical imaging.

One of the challenges in semantic image segmentation of fashion garments is the presence of the background in the images. The background can interfere with the segmentation of the garment and reduce the accuracy of the segmentation. Therefore, cleaning the background from garment images can help extract more information from those images and improve the accuracy of the segmentation. This can be achieved using various techniques such as image cropping, background removal, and image retouching.

In this part, the use of deep neural networks, especially CNNs, for semantic image segmentation on fashion garments will be discussed. Cleaning the background from garment images can improve the accuracy of the segmentation.

3.1.3.1 U2Net architecture for segmentation

U2-Net (Qin et al., 2020)is a deep learning model that was originally designed for salient object detection (SOD). However, it has also been used for other tasks, such as image matting, background removal, and image segmentation. The nested U-structure, inspired by the original (Ronneberger et al., 2015) for biomedical image segmentation paper, showcases the adaptability and effectiveness of this architecture in different computer vision tasks. While the original U-Net model employs a single U-shaped structure for feature extraction and reconstruction, the U2-Net approach proposes a nested variant within the encoder component, showcasing the evolution and specialization of this fundamental architectural principle to achieve enhanced performance in salient object detection (SOD). Figure 3.7 below depicts a hierarchical

deep-learning architecture designed for medical image segmentation. The model consists of multiple encoder-decoder blocks En_1 to En_6 and De_1 to De_5 that process input images at different scales, progressively extracting features and refining the segmentation output through side outputs Sup1 to Sup6. The final segmentation result, Sup0, is obtained by fusing these middle outputs, ensuring that both global context and fine details are captured for accurate segmentation.

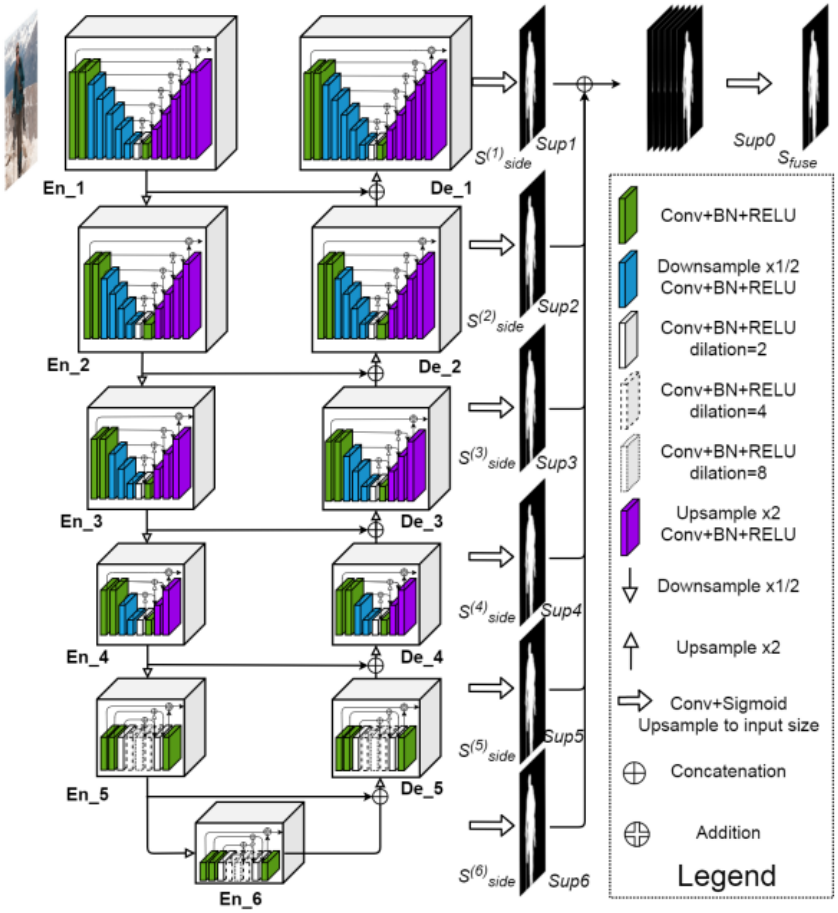

Figure 3.7 U2 -Net architecture

Source: (Qin et al., 2020)

3.1.3.2 Dataset for semantic segmentation

For this thesis, to use U2-Net for fashion garment segmentation, this model needs to be trained the model on a dataset of images that includes garment pictures. The iMaterialist (Fashion) 2019 dataset, provided as part of the FGVC6 challenge on Kaggle, is a comprehensive large-scale collection aimed at advancing image segmentation in the fashion industry. It consists of over 50,000 images, each annotated for one of three tasks: image-level annotations, object detection bounding boxes, and semantic segmentation polygons. This Dataset includes 46 types of fashion categories, such as 'dress,' 'shirt,' and 'pants,' providing strong support for developing detection and segmentation models that can identify various fashionable items.

All images in the Dataset are annotated with various attributes like color, texture, or fabric pattern, which makes the segmentation task more challenging but reflects real-world scenarios. These attributes are crucial for building models that detect not only the category of fashion items but also different features within them, making it a valuable resource for detailed fashion recognition. The pixel-wise segmentation masks enable precise location and shape boundary extraction for detailed analysis and accurate model training.

Organizers of the competition encourage participants to utilize this highly annotated Dataset to expand current fashion image analysis work. By offering a diverse and large dataset, the challenge aims to foster innovation in machine learning models for complex segmentation and classification tasks in the fashion domain. The ultimate goal is to improve the quality of automated fashion item recognition and segmentation, forming the basis for several applications, including e-commerce, digital wardrobe management solutions, and personalization systems in the fashion domain.

Since the original 46 categories in the dataset labels add to the complexity, categories are grouped into two categories (upper body and full-body) for simplification. The images are resized into 512 x 512 px squares for training.

To show the outputs of this process, Figure 3.8 presents a sample result from the iMaterialist dataset, showcasing the process of image segmentation. On the left, the original image of a garment is displayed. The middle image represents the mask generated by the segmentation model, which identifies the boundaries of the garment.

On the right, the masked image shows the garment isolated from its background, highlighting the effectiveness of the segmentation in accurately extracting the item from the original photo. The same process was employed for the CuTeks Dataset in Figure 3.9, and similar results were obtained.

Figure 3.8 Sample Result iMaterialist Dataset: Left: Original, Middle: Mask Right: Masked Image

Source: Author

Figure 3.9 Sample Result CuTeks Dataset: Left: Original, Middle: Mask Right: Masked Image

Source: Author

3.2 AI in Fashion

There are lots of uses of AI in the fashion sector. In this section, many examples will be given by researchers, fashion brands, and technology companies that have successfully integrated AI into their design processes, demonstrating the practical benefits and transformative potential of this technology.

A neural network named "Style M" has been designed to behave like a personal fashion stylist (Hsiao & Grauman, 2017), analyzed a person's previous fashion choices, and built a neural network that suggests new combinations for that individual with a higher buyer acceptance rate. This shows that machine learning and artificial intelligence techniques can help us decide what an individual wants to wear in his/her daily life, considering his/her fashion sense.

Another best fashion outfit composition project utilizes the deep learning system (L. Liu et al., 2019). It evaluates and scores the garment's aesthetic features and finds the possible combination of other items based on this score. Garment's aesthetic features are used as categories, colors, coherence, patterns, creative pairing choices, as well as styles for different ages and individuals. Those features are based on (Kobayashi, 1991) work about clothing semantic attributes. (L. Liu et al., 2019) found that even though color values on the image have the highest individual impact, the best outcome can be obtained by combining these features of the garment's aesthetic features.

Another emotion-based search would be (Yang et al., 2021). This research introduces a method to design fashion products that match consumers' emotional needs using modified CNN and GAN models. The method includes a product image recognition model and an intelligent design generation model, effectively demonstrated through shoe design.

Gathering information on fashion products is crucial. This can be done by detecting garments on the human body or gathering information about the fabric. Most object detection methods represent clothing regions by either bounding boxes or human joints. (Z. Liu, Yan, et al., 2016) presents fashion landmark detection or fashion alignment, which predicts the positions of functional key points defined on the fashion items, such as the corners of the neckline, hemline, and cuff. This Dataset has 123,016 images, each of which is labeled with eight landmarks. Fashion alignment is studied by cascading multiple convolutional neural networks. In Figure 3.10, it can be seen

that garment sample edge points are correctly found. This technique can be used to get more meaningful and rich information on garments.

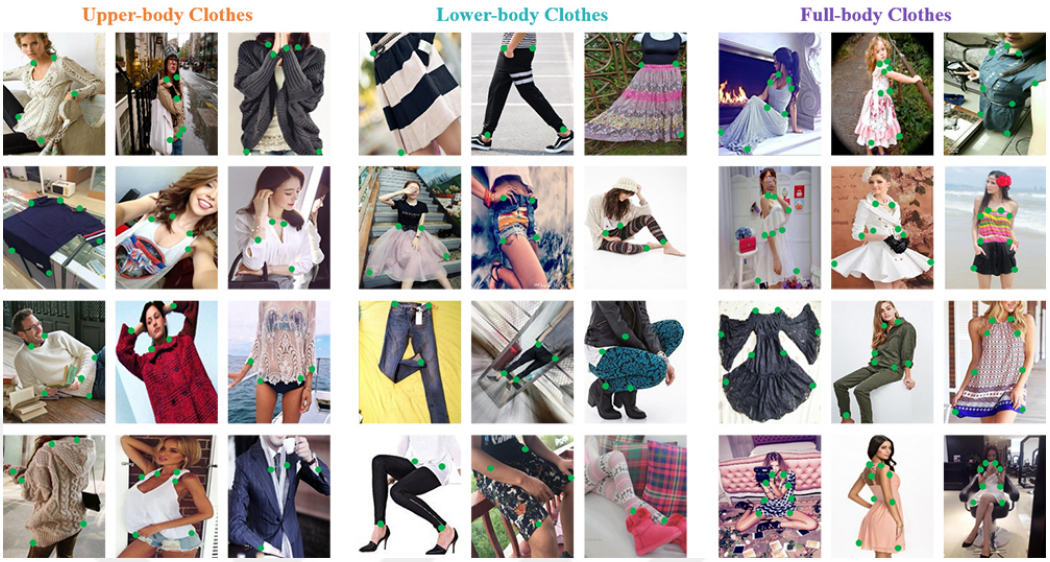

Figure 3.10 Fashion Landmark Detection

Source: (Z. Liu, Yan, et al., 2016)

Detecting fashion products by itself is not enough. There are lots of garment attributes and features. (Z. Liu, Luo, et al., 2016) created an extensive fashion set for fashion garments to classify over 800,000 images. In their work, they introduce DeepFashion, a large-scale clothes dataset with comprehensive annotations, which are richly annotated with massive attributes, clothing landmarks, and correspondence of images taken under different scenarios, including store, street snapshot, and consumer. It is optimized in an iterative manner. In (Hekimoğlu, 2017), I had features for shapes and touch sensitivities, but those extensive features can also be used.

In addition to online shopping, virtual try-on systems can enhance the in-store shopping experience. Retailers can implement these systems in physical stores to offer consumers a quick and hygienic way to try on multiple items without the need to change clothes. This is particularly beneficial in the current context of heightened hygiene awareness due to the COVID-19 pandemic. The integration of virtual try-on systems in stores can also create a more engaging and interactive shopping experience, attracting tech-savvy. (Hashmi et al., 2020) proposes a novel virtual try-on system that enhances the customer experience by providing a photo-realistic visualization of

custom outfits on users. The architecture integrates Generative Adversarial Networks (GANs) for generating realistic images and Neural Body Fit (NBF) for accurately mapping 3D poses. The system employs human segmentation to isolate the body from the background, ensuring precise garment fitting. Human pose estimation is achieved using Spatial Transformer Network (STN) and Parallel Single Person Pose Estimator (SPPE), capturing accurate body measurements. A similarity query reasoning mechanism provides visual recommendations of similar garments based on user selections.

There is much work involved in creating new garment samples. The author's (Bhattacharya et al. 2022) work FAR-GAN represents a significant advancement in the field of fashion apparel regeneration, focusing on the precise control of color in the synthesized images. This Generative Adversarial Network (GAN)--based tool integrates style information from fashion apparel with target color embeddings, allowing for the explicit control of the color of regenerated apparel items. The key innovation of FAR-GAN lies in its two-step encoding process, which augments features from both the fashion apparel and its edge map to extract detailed style information. This extracted information is then controlled with the desired color embedding during the decoding phase, facilitating the production of apparel images in specified colors.

(Shastri et al. 2022) Introduce Vastr-GAN, an innovative text-to-image synthesis model designed to create intricate Indian apparel designs from textual descriptions. This model strategically combines multiple trained Generative Adversarial Networks (GANs) to generate high-quality images that reflect the detailed specifications provided in the text inputs. To address the lack of datasets tailored to traditional Indian fashion, the authors developed an Indian Fashion dataset comprising approximately 16,000 images of apparel, such as Kurtis and kurtas, along with their corresponding textual descriptions. This Dataset enables Vastr-GAN to generate visually accurate and culturally relevant designs, filling a significant gap in the existing resources that predominantly cater to Western fashion.

(Yan et al. 2023) introduce a novel AI-based framework for fashion design that leverages the capabilities of Generative Adversarial Networks (GANs) to enhance designers' efficiency. This framework addresses the dual need for randomness and

controllability in the design process. The system comprises two main modules: a sketch-generation module and a rendering-generation module. The sketch-generation module, based on latent space, enables the creation of various sketches, while the rendering-generation module learns the mapping between textures and sketches, completing the design process. The results show that the proposed model significantly improves the design process by allowing designers to utilize the controllability of different conditions (e.g., sketch and texture) and the randomness of latent space, ultimately facilitating the creation of innovative and efficient fashion designs.

4. CHAPTER: DATASET SELECTION

In order to feed GAN and CNN models, we need to train them with relevant fashion garment data to get meaningful results. For this project, data set from my previous MSc Thesis with Mehmet Ünlütürk and TÜBİTAK project with CuTeks Company data will be used. From 2017 to 2018, image datasets consist of 4833 garment samples. A subset of garment images from brands like Bershka, Elcorte, Hema, Mango, Zara (Inditex), etc., were selected. Those garment samples have passed through several GAN architectures, mainly coded in Python. Most GAN architectures' base files exist on the GitHub repository. They need some adjustments to get data and some configurations. The distribution of garment samples can be seen in Table 4.1 below.

Table 4.1 Garment Samples by Brand

Brand Name	2017 Stats	2018 Stats
BERSHKA	698	274
HEMA	402	91
ELCORTE	238	26
MANGO	419	253
MORRISONS	116	27
REWE	110	16
SIPLEC	139	70
ZARA	1594	360
Total	3716	1117

Source: Author

Figure 4.1 Garment samples

Source: Author

It is shown that most of the garment sample has the same resolution and background in Figure 4.1. But there, some of the garment samples have faults, such as not being centered correctly, background differences, etc. Those samples must be corrected, centered, or else eliminated from the Dataset. Moreover, all resolutions must be the same in order to get reasonable results from the models. This problem was mainly solved by segmenting images with U2-Net. Images are standardized to a 512x512 pixel squared format, with background filling to ensure uniformity across the Dataset. This preprocessing step ensures consistency and improves the accuracy of the AI model during training.

4.1 Brand-Based T-Shirt Dataset

From the original Dataset, we will create a new dataset focusing exclusively on t-shirts, which are a significant portion of the original data. By isolating t-shirts, we ensure that the subset is relevant for creating a comprehensive dataset specifically for shirt design and analysis. The following Table 4.2 represents the estimated number of t-shirts for each brand:

Table 4.2 Shirt Samples by Brand

Brand Name	**2017 Stats**	**2018 Stats**
BERSHKA	558	219
HEMA	322	73
ELCORTE	190	21
MANGO	335	202
MORRISONS	93	22
REWE	88	13
SIPLEC	111	56
ZARA	1275	288
Total	2972	894

Source: Author

By focusing on this t-shirt subset, we can leverage the strengths of GANs to generate innovative and market-relevant T-shirt designs, ultimately enhancing the creative process and improving efficiency in the fashion design industry. Some of the T-shirt garment samples from MANGO company are given in Figure 4.2.

Figure 4.2 Dataset From Cuteks Mango

Source: Author

4.2 Segmentation applied Dataset

In addition to the Shirt data set, we have another dataset consisting of masked images of t-shirts. This Dataset is critical for training models that require image data to learn design patterns and styles. Masked images allow the model to focus on the garment itself, removing distractions from the background and other elements. All of the Dataset from the "Brand-based t-shirt dataset" is segmented using U2Net architecture, and some examples of results are shown in Figure 4.3.

Figure 4.3 Masked Results

Source: Author

4.3 Pattern Dataset

A comprehensive study was conducted as part of the TÜBİTAK project (5160117) to utilize artificial intelligence in predicting consumer preferences for garment designs. The objective of this project was to create an artificial intelligence model in order to predict fashion trends. It was a joint effort between Yaşar University and CuTeks Textile Industry and Trade Inc. The project examined 530 distinct garment designs supplied by CuTeks Textile, extracting 216 unique attributes for each design. The attributes encompassed color patterns, emotional responses (assessed through the Kobayashi method), and fabric touch values.

The database derived from this analysis functions as the foundation for predicting consumer preferences regarding particular designs. The garment patterns analyzed in

this study are diverse and represent distinct fashion aesthetics. The database encompasses a diverse array of designs, spanning from informal and contemporary to sophisticated and timeless. Every design is assigned with specific emotional vectors and fabric touch values. The AI model underwent training using these patterns, which were scrutinized using sophisticated image analysis techniques. The diagram presented below illustrates various instances of these patterns, highlighting the wide array of designs and the thorough analysis carried out. The project's findings demonstrate the effectiveness of employing artificial intelligence to precisely predict fashion preferences, offering valuable insights for the fashion industry.

Some pattern examples from brands like Mango, Bershka, Zara, etc, are in Figure 4.4 below.

Figure 4.4 Pattern Samples

Source: Author

5. CHAPTER: COMPUTATIONAL EXPERIMENTS & RESULTS

In this part, we detail a set of computational experiments addressing the performance evaluation methods related to fashion image analysis frameworks relying on state-of-the-art deep learning techniques. Given the heterogeneous nature of our Dataset, these experiments aim to test how well different models handle a variety of challenges in processing fashion images across varying conditions. In particular, we study CycleGAN with and without background, Neural Transfer for Garment-to-Garment and Pattern-to-Pattern transformations, and StyleGAN with a limited number of training data points in both conditions. Each experiment is thoughtfully proposed to demonstrate whether and where those models work well for which problems in fashion image processing.

First, we exploit the CycleGAN (Cycle-Consistent Generative Adversarial Network) in two scenarios with and without background. The CycleGAN with background also attempts to change the fashion of image sets while preserving their contextual environment, which includes settings and background elements. This strategy is valuable in applications such as virtual try-ons or fashion catalog generation, where it is important to preserve the layout of the image. On the other hand, our method explicitly handles fashion items in isolation without surrounding elements (CycleGAN without background), resulting in more accurate item alteration or replacement for tasks where the user wants to change particular parts of an outfit. We compared these two methods to investigate how contextual information influences CycleGAN performance in fashion image transformation tasks.

The next two uses of the Neural Transfer method are for Garment-to-Garment transfer and Pattern-to-Pattern transfer. Garment-to-Garment takes a garment and transfers the style and elements onto another, helping new fashion designs by mixing innovative features from different clothing. By using this technique, fashion design and customization can take on new life: garments that seemed undesirable at first can be made into something unique and personal. The Pattern-to-Pattern transfer, in contrast,

changes the patterns of clothes so that fabric designs can change while maintaining their original garment structure. The most common use of this system is for textile design and the customization of fashion pieces, allowing future options to expand for fashion designers, such as variant patterns and textures.

We evaluate the proposed pipeline with and without a background using the StyleGAN (Style-Based Generator Architecture for GAN) approach. Fashion images can predict the fashion items themselves and their contextual environment, which is significant for the systematic generation of entire fashion catalogs and virtual displays. On the other hand, for applications that might need full-visible fashion items (like e-commerce platforms or digital fashion showcases), a single isolated item is crucial. By removing the background from those target domain images, all vectors effectively focus on isolating an individual high-performance generation of each constituent in isolation. During the training of GANs, the number of thousands of images (kimgs) processed is a crucial parameter. Kimgs is a shorthand notation that denotes the total number of images (in thousands) that have been fed into the network during training. Monitoring the kimgs parameter helps evaluate the progress of the training and understand the stability of the GAN model.

Pytorch (Paszke et al., 2017) has been used to implement the models. Neural Style Transfer Experiments were conducted with Nvidia Geforce 3080 GPU, and Both CycleGAN and StyleGAN experiments were conducted with Nvidia Tesla A100 GPUs. These experiments yield a detailed review of different deep learning-based fashion image analysis methods and their successes and failures across datasets. The results of such experiments will provide useful knowledge for the deployment of these models for fashion applications and stimulate future research in automated fashion image processing.

The proposed GAN architecture is shown in Figure 5.1. This architecture integrates the U2Net and StyleGAN frameworks to enhance and stylize images of fashion products. The left side of the diagram represents the U2Net architecture, which is designed for high-precision image segmentation. In this specific example, the input garment image is segmented into various regions based on color or other visual features. The segmentation process involves hierarchical downsampling to extract

details at different scales, followed by upsampling to reconstruct a segmented image that retains the critical features identified at each scale.

The right side of the diagram shows the StyleGAN architecture, which is used to generate images based on the segmented output from U2Net. StyleGAN operates by taking a normalized random noise vector and transforming it through a mapping network into an intermediate latent space, which provides fine control over the generated image's features. The synthesis network then begins with a constant input tensor and iteratively applies convolutional layers and adaptive instance normalization (AdaIN) to progressively increase the resolution and detail of the image. Gaussian noise is introduced at various stages to ensure variability and realism in the generated outputs. The final product is a high-resolution, stylized image of the T-shirt, where the segmentation information from U2Net guides StyleGAN in emphasizing or modifying specific aspects of the garment design.

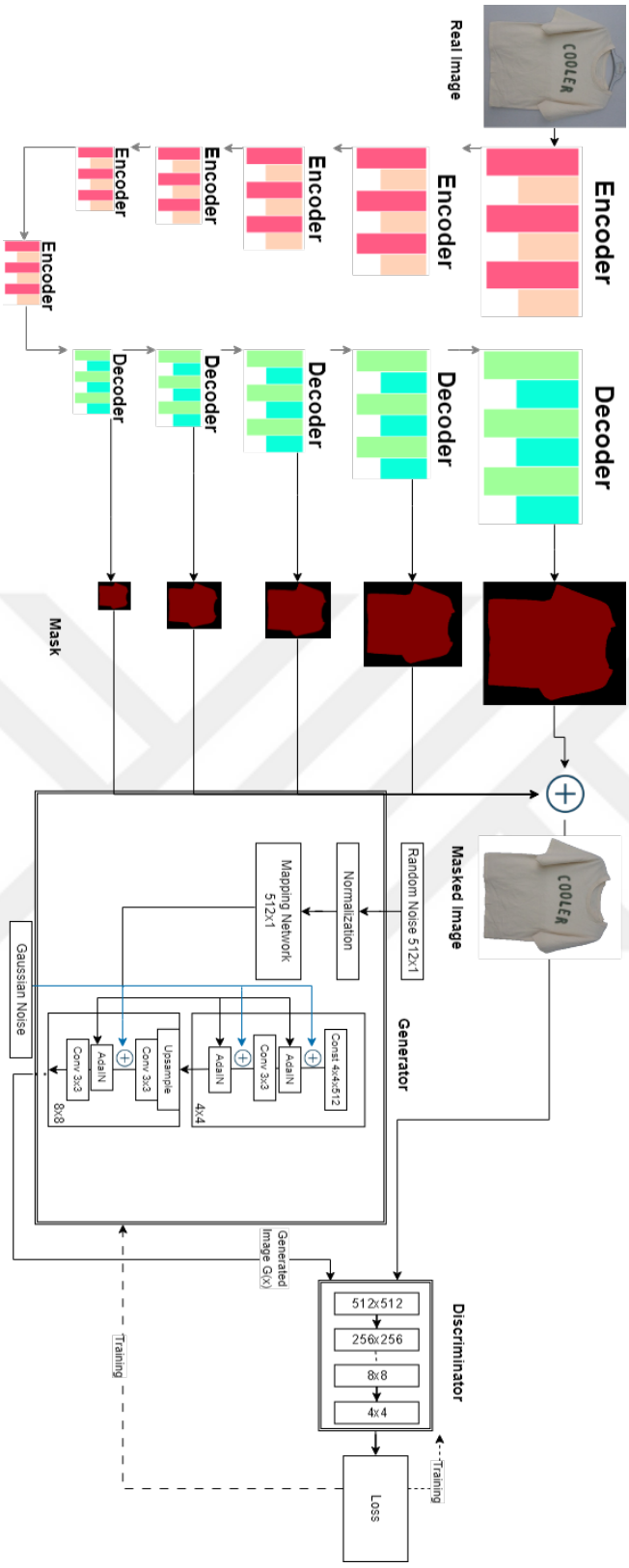

Figure 5.1 Proposed GAN Architecture

Source: Author

5.1 CycleGAN approach: (with background)

Table 5.1 CycleGAN Parameters

Number of epochs of training	35000
Batch size	8
Learning rate	0.02
Decay of first-order momentum of gradient	0.5
Decay of second-order momentum of gradient	0.999
Number of CPU threads to use during batch generation	6
Size of image height	512
Size of image width	348
Number of image channels	3
Interval between saving generator outputs	1000
Cycle loss weight	10
Identity loss weight	3
GAN criterion loss method	MSE loss
Cycle loss method	L1loss

Source: Author

The experiment hyperparameters for the CycleGAN approach with the background that is given in Table 5.1 are as follows. The model is trained for 35,000 epochs with a batch size of 8. The learning rate is set at 0.02, with the first-order momentum decay at 0.5 and the second-order momentum decay at 0.999. During batch generation, 6 CPU threads are used. The input images have dimensions of 512x348 pixels and consist of 3 channels (RGB). Generator outputs are saved at intervals of every 1,000 epochs. The cycle loss weight is set to 10, and the identity loss weight is 3. The GAN criterion is evaluated using the Mean Squared Error (MSE) loss, while the cycle consistency is measured using L1 loss.

Figure 5.2 CycleGAN First Epoch

Source: Author

Figure 5.3 CycleGAN Epoch: 15000

Source: Author

Figure 5.4 CycleGAN Epoch: 35000

Source: Author

For this experiment, the image format is organized in a grid format, where the odd-numbered rows display the real, unaltered images, and the even-numbered rows display the corresponding generated images produced by the CycleGAN. The results can be observed through Figures 5.2, 5.3, and 5.4. The generated results vary in quality, with some images retaining key design elements while others appear distorted or blurred. This indicates that the model struggles to maintain high fidelity, particularly due to the lack of effective cycling between domains, which is crucial for preserving content across transformations.

5.2 CycleGAN approach: (Maskedbackground)

Table 5.2 CycleGAN parameters masked

Number of epochs of training	40000
Batch size	8
Learning rate	0.05
Decay of first-order momentum of gradient	0.5
Decay of second-order momentum of gradient	0.999
Number of CPU threads to use during batch generation	8
Size of image height	512
Size of image width	512
Number of image channels	4(RGBA)
Interval between saving generator outputs	1000
Cycle loss weight	10
Identity loss weight	3
GAN criterion loss method	MSE loss
Cycle loss method	L1loss

Source: Author

The experiment hyperparameters for the CycleGAN approach that is given in Table 5.2, which uses a masked background technique, are as follows. The model undergoes training for 40,000 epochs with a batch size of 8. The learning rate is fixed at 0.05, while the first-order momentum decay is 0.5 and the second-order momentum decay is 0.999. During batch generation, 8 CPU threads are utilized. The input images have a resolution of 512x512 pixels and include four channels (RGBA). Generator outputs are saved at intervals of every 1,000 epochs. The cycle loss weight is set to 10, and the identity loss weight is 3. The Mean Squared Error (MSE) loss function is used to evaluate the GAN criterion, and L1 loss is applied for cycle consistency evaluation.

Figure 5.5 CycleGAN Masked Epoch: 0

Source: Author

Figure 5.6 CycleGAN Masked Epoch: 20000

Source: Author

Figure 5.7 CycleGAN Masked Epoch: 40000

Source: Author

For this experiment, the image format is organized as in the previous CycleGAN experiment. The results can be observed through Figures 5.5, 5.6, and 5.7. After Epoch 40,000, the results shown in this image are significantly improved, producing garment images that closely resemble the originals. The better performance can be attributed to the effective segmentation of the garment from the background, which allows the model to focus more accurately on the relevant features of the clothing. This targeted approach enhances the quality of the generated images, leading to more consistent and visually similar outputs compared to earlier iterations.

5.3 Neural Transfer approach (Garment to Garment):

Table 5.3 Neural Transfer approach Garment x Garment parameters

Number of epochs of training	25000
Model	Pre-trained VGG19 model with ImageNet weights
Learning rate	12
Decay steps	100
Decay rate	0.6
Optimizer	Adam optimizer
Size of image height	512
Size of image width	348
Number of image channels	3
GAN criterion loss method	MSE loss

Source: Author

Table 5.3 presents a Neural Transfer approach tailored for garment-to-garment transformation. This method utilizes a pre-trained VGG19 model with ImageNet weights and undergoes training over 25,000 epochs. The initial learning rate is set at 12 and decays every 100 steps by a factor of 0.6. Optimization is performed using the Adam optimizer. The input images have dimensions of 512x348 pixels and contain three channels, adhering to the RGB format. The model's performance and accuracy are assessed using the Mean Squared Error (MSE) loss function, which ensures precise transformations between garments with minimal error.

Figure 5.8 Transferring Style (Left) to Content (Middle) to create a new garment

Source: Author

Figure 5.9 Samples of successful images

Source: Author

The neural transfer approach of Garment to Garment has successfully shown its potential in the transfer of style elements between different garments and, hence, in the achievement of unique, creative fashion designs. In Figure 5.8, it can be seen that the style of the garment image on the left is used to transfer the content of the garment image in the middle, and the result is shown on the image on the right. Through the employment of gram matrices for both feature representation and feature mixing, this technique proved that it had the power to merge different garment styles accurately. Its potential to allow experiments with new styles and concepts just underlines the

promise that it has in revolutionizing fashion design. Samples of successful images can be seen in Figure 5.9.

5.4 Neural Transfer Approach (Pattern to Pattern):

Table 5.4 Neural Transfer approach pattern x pattern parameters

Number of epochs of training	1000
Model	Pre-trained VGG19 model with ImageNet weights
Learning rate	12
Decay steps	100
Decay rate	0.6
Optimizer	Adam optimizer
Size of image height	24
Size of image width	24
Number of image channels	3
GAN criterion loss method	MSE loss

Source: Author

Table 5.4 outlines a Neural Transfer approach designed for pattern-to-pattern transformation. In this approach, a pre-trained VGG19 model with ImageNet weights is employed. The model undergoes training for 1,000 epochs, starting with a high learning rate of 12, which decays every 100 steps by a factor of 0.6. The optimization process is managed by the Adam optimizer. The input images used in this approach have a resolution of 24x24 pixels and consist of 3 channels, adhering to the RGB color model. The effectiveness of the neural transfer process is evaluated using Mean Squared Error (MSE) loss, which ensures that the transformed patterns closely match the target patterns with minimal error.

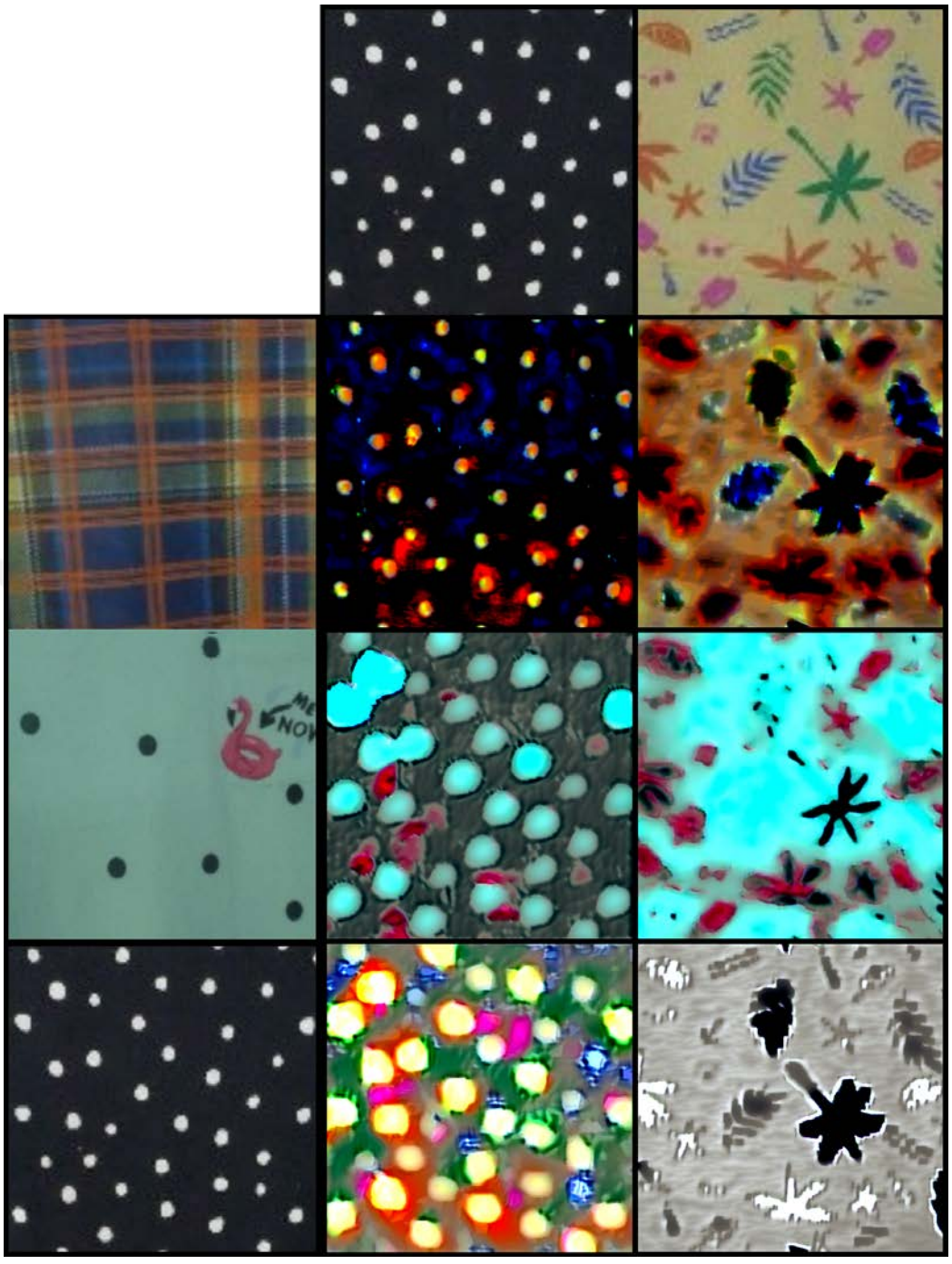

Figure 5.10 Transferring Style (Left) to Content (Top) End product (bottom right cells) to create a new pattern

Source: Author

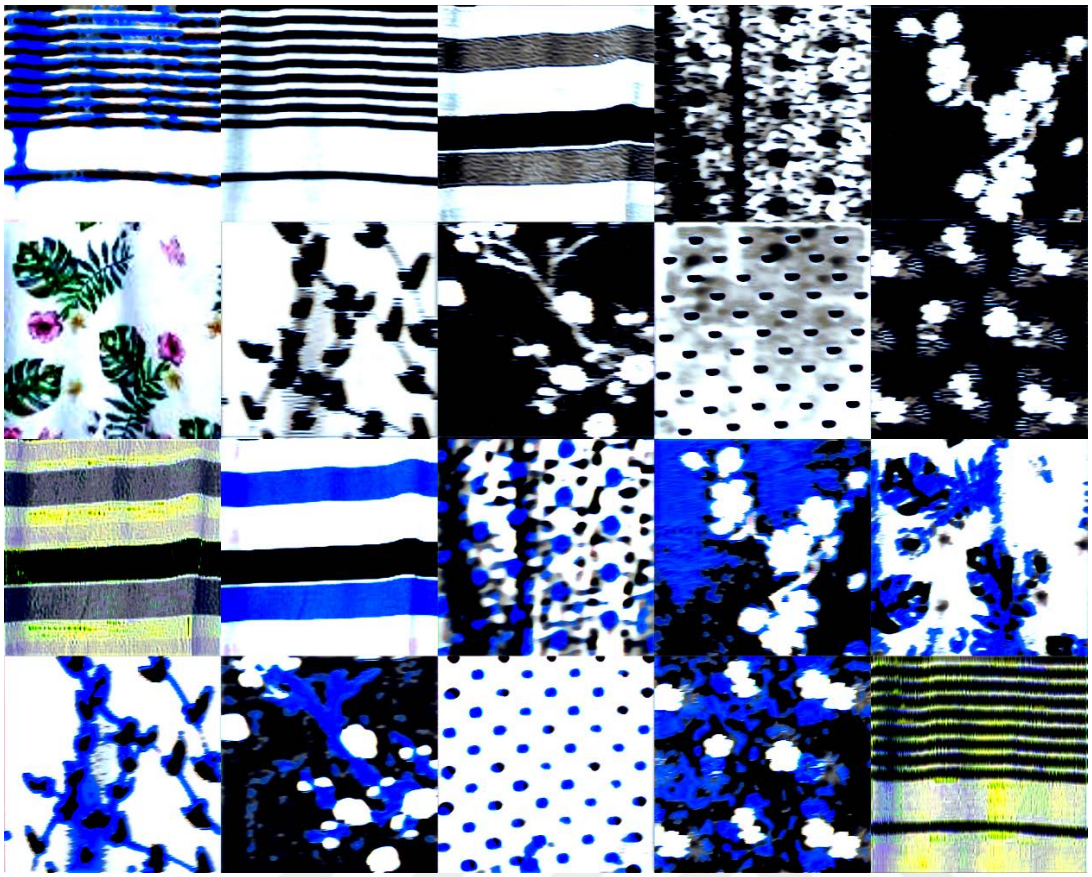

Figure 5.11 Samples of pattern-to-pattern transferred images

Source: Author

The pattern-to-pattern neural transfer technique pretty well demonstrated the technique's capability to print fabric patterns without losing the physical structure of the original garments. Figure 5.10 demonstrates the process of style transfer, where the visual style from images in the left column is applied to the content of images in the top row, resulting in the newly generated patterns displayed in the bottom right cells. Through the style transfer process, the stylistic elements of the left column images seem to be integrated with the structural and thematic content of the top row images. The application of Gram matrices in the use of style and texture within patterns gave a very varied and creative use in textile design. This methodology has offered great potential for expanding the creative horizons of fashion designers as well as bringing many innovations in fabric design. In Figure 5.11, each picture represents a unique combination of original and transferred patterns, demonstrating the diversity and complexity achievable through pattern transfer techniques.

5.5 Neural Transfer Approach (Pattern to Garment):

Table 5.5 Neural Transfer Approach Pattern x Garment Parameters

Number of epochs of training	1000
Model	Pre-trained VGG19 model with ImageNet weights
Learning rate	12
Decay steps	100
Decay rate	0.6
Optimizer	Adam optimizer
Size of image height	24
Size of image width	24
Number of image channels	3
GAN criterion loss method	MSE loss

Source: Author

Table 5.5 outlines a Neural Transfer methodology designed to apply patterns to clothing items. This method employs a pre-trained VGG19 model that has been trained using ImageNet weights. The model undergoes training for 1,000 epochs, commencing with an initial learning rate of 12. The learning rate is subsequently reduced by a factor of 0.6 every 100 steps. The Adam optimizer is employed to optimize the parameters of the model. The input images for this process have dimensions of 24x24 pixels and consist of 3 channels representing the colors red, green, and blue (RGB). The neural transfer's performance is assessed by measuring the Mean Squared Error (MSE) loss. This ensures that the generated images faithfully reproduce the patterns on the garments with minimal deviation.

Figure 5.12 Samples of Pattern to Garment images; Transferring Style (Left) to Garment (Middle) End product (Right) to create a new style Garment

Source: Author

The Pattern to Garment neural transfer technique demonstrated good results in terms of successfully implementing complicated patterns over garment structures. The approach used in the method, through gram matrices for style representation, was capable of representing patterns with fine detail. This technique, therefore, offers a promising stride for bespoke fashion design and automation of the manufacturing process since it suggests significant strides in manufacturing efficiency that are to be realized in the fashion industry. Figure 5.12 showcases how a distinct pattern is applied to the original garment design, resulting in a transformed garment with an updated style. The end products demonstrate the potential of this technique to generate innovative fashion designs by integrating different patterns into existing garment templates.

5.6 StyleGAN approach: (With background):

Table 5.6 StyleGAN parameters

Number of epochs of training	30000
Model	StyleGAN
Learning rate	0.25
Batch size	32
Total kimgs	2500
Optimizer	Adam optimizer
Size of image height	512
Size of image width	512
Number of image channels	3
GAN criterion loss method	FID-FID50k

Source: Author

Table 5.6 details a StyleGAN approach involving images with backgrounds. The model is trained over 30,000 epochs, maintaining a learning rate of 0.25 and a batch size of 32. Throughout the training process, a total of 2,500 thousand images (kimgs) are processed. The optimization is carried out using the Adam optimizer, ensuring efficient and effective learning. The generated images have dimensions of 512x512 pixels and consist of 3 channels corresponding to the standard RGB color model. The performance and quality of the GAN are assessed using the FID as the criterion loss method, which ensures that the generated images closely match the statistical distribution of real images. The generated networks' FID50k values are also calculated for 5k epochs. A lower FID50k score is better.

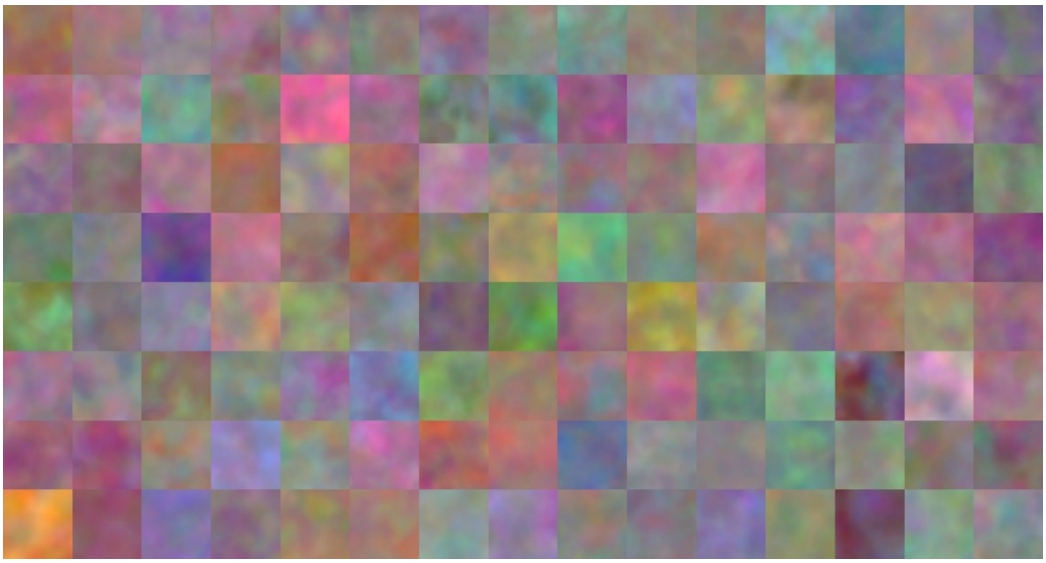

Figure 5.13 StyleGAN Unmasked Epoch: 0

Source: Author

The network has started with a Fid50k value of 420.5244, which is reasonable for the initial value.

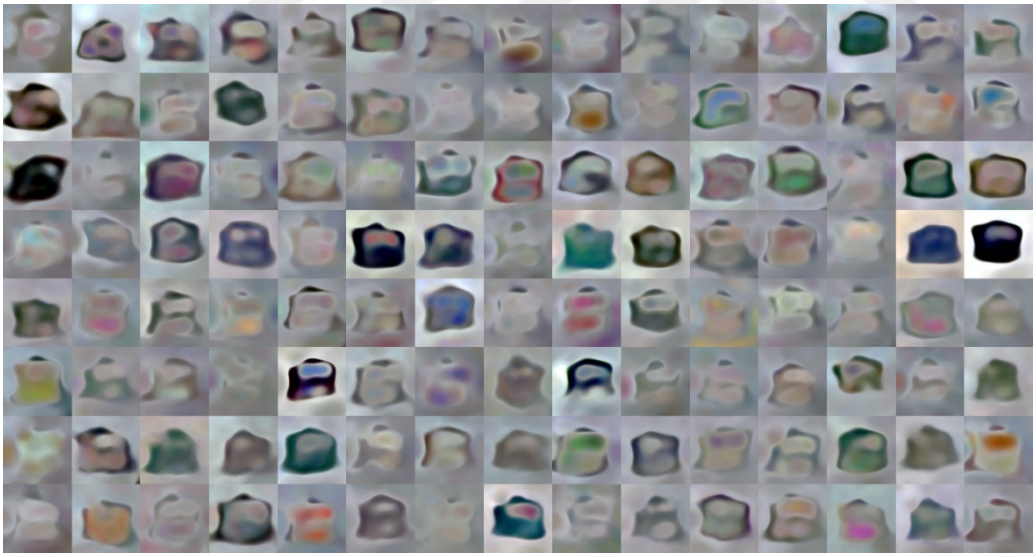

Figure 5.14 StyleGAN Unmasked Epoch: 15000

Source: Author

In Epoch 15000, the Fid50k value is 351.4369.

Figure 5.15 StyleGAN Unmasked Epoch: 30000

Source: Author

The final Fid50k value for Epoch 30000 is 321.4369.

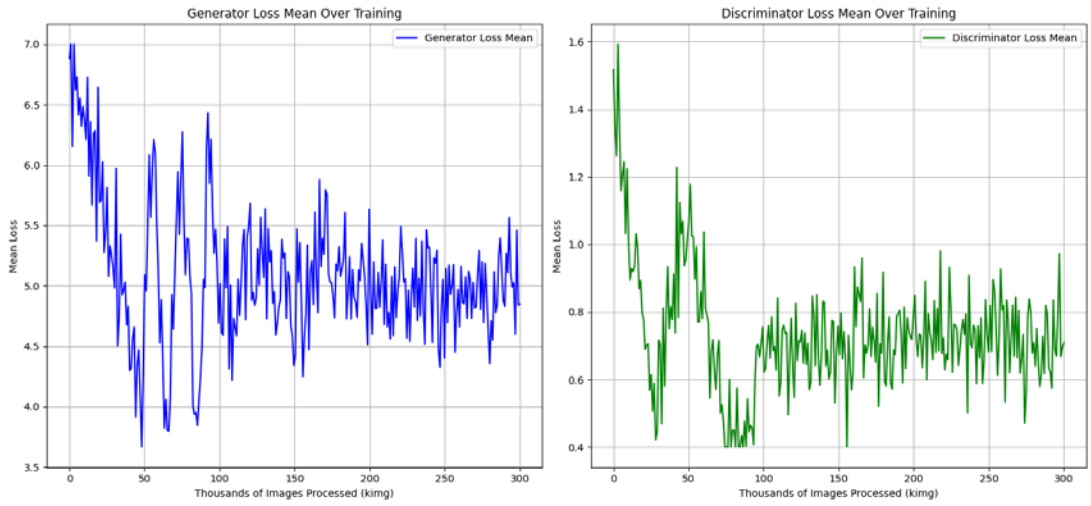

Figure 5.16 StyleGAN G and D Loss Mean

Source: Author

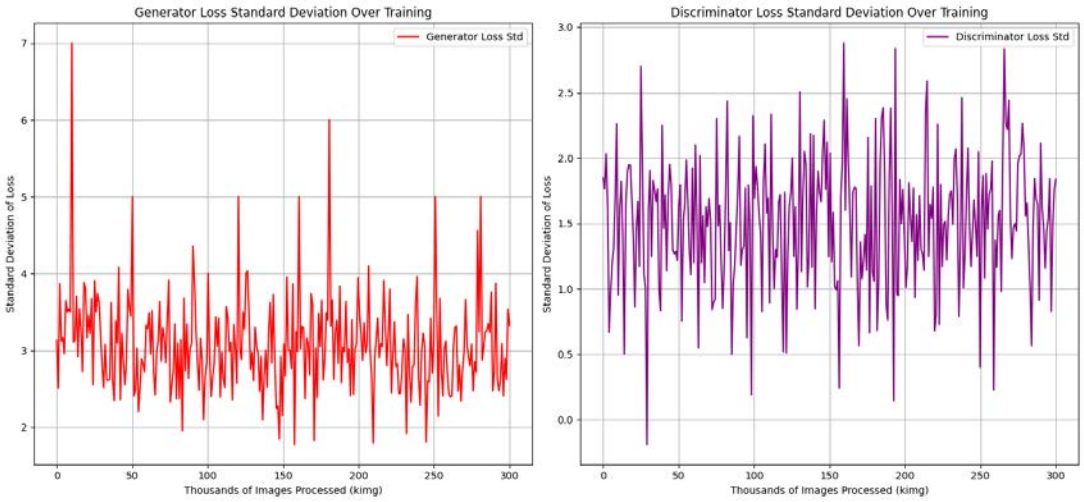

Figure 5.17 StyleGAN G and D Loss Standart Deviation

Source: Author

For this experiment, the image format is organized in a 15x8 image grid format, where all the outputs are fake images produced by StyleGAN. The results can be observed through Figures 5.13, 5.14, and 5.15. The results are not particularly strong, as the generated images appear blurry, distorted, and lack clear definition. The shapes of the garments are generally recognizable, but the details are muddled, and the overall quality of the images is low. The training graphs in Figures 5.16 and 5.17 suggest that while the discriminator maintains stable performance, the generator exhibits significant instability, particularly in the later stages of training. This instability can be indicative of mode collapse, where the generator fails to produce a diverse set of outputs.

5.7 StyleGAN approach: (Maskedbackground)

Table 5.7 StyleGAN masked parameters

Number of epochs of training	30000
Model	StyleGAN
Learning rate	0.25
Batch size	32
Total kimgs	2500
Optimizer	Adam optimizer
Size of image height	512
Size of image width	512
Number of image channels	4
GAN criterion loss method	FID-FID50k

Source: Author

In this experiment, a StyleGAN approach using a masked background technique is used. The model undergoes extensive training with the given parameters in Table 5.7. Over 30,000 epochs with a learning rate of 0.25. During training, a batch size of 32 is employed, processing a total of 2,500 thousand images (kimgs). The Adam optimizer is utilized to adjust the model parameters. The generated images have a resolution of 512x512 pixels and contain four channels, including an alpha channel for the masked background. The quality and performance of the GAN are evaluated using FID as the criterion loss method, ensuring the generated images are closely aligned with the distribution of real images.

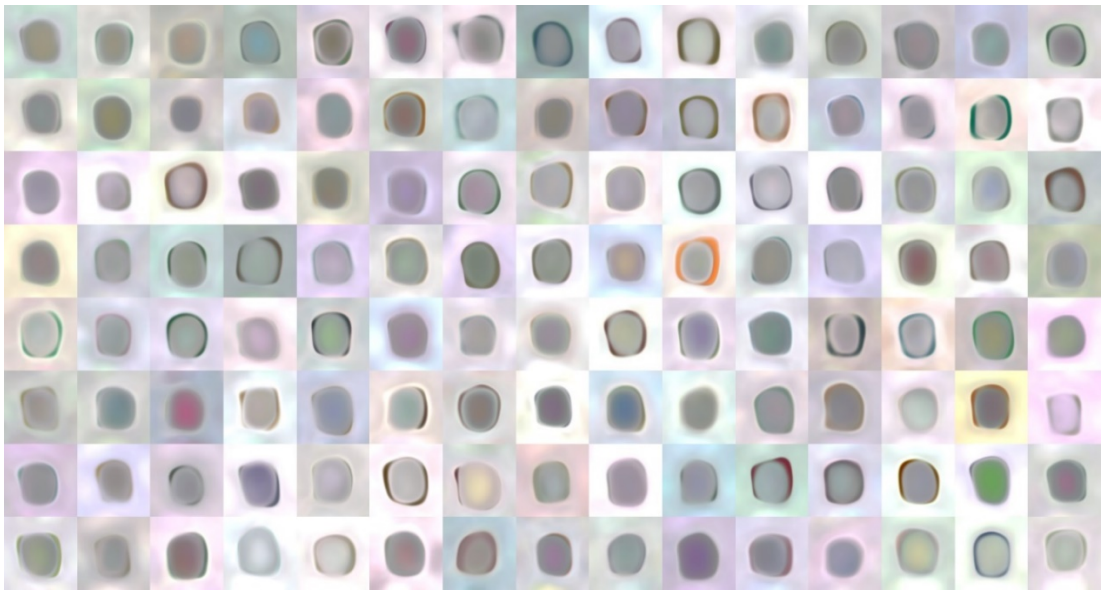

Figure 5.18 StyleGAN masked Epoch: 0

Source: Author

The network has started with a Fid50k value of 268.1456, which is reasonable for the initial value.

Figure 5.19 StyleGAN masked Epoch: 15000

Source: Author

In Epoch 15000, the Fid50k value is 38.6801.

Figure 5.20 StyleGAN masked Epoch: 30000

Source: Author

The final Fid50k value for the StyleGAN network (Epoch 30000), which result can be seen in Figure 5.19, is 8.8481.

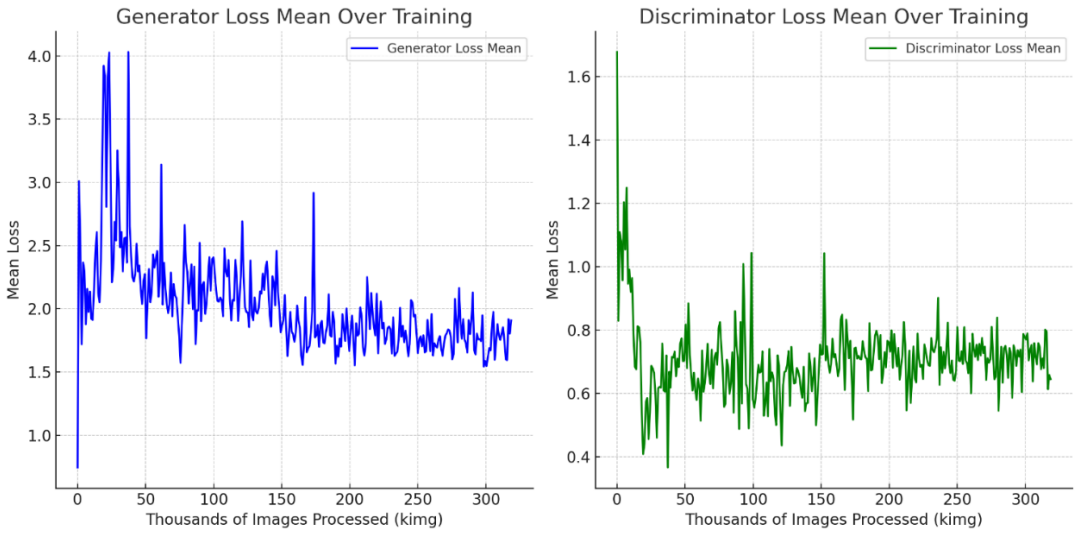

Figure 5.21 StyleGAN masked G and D Loss Mean

Source: Author

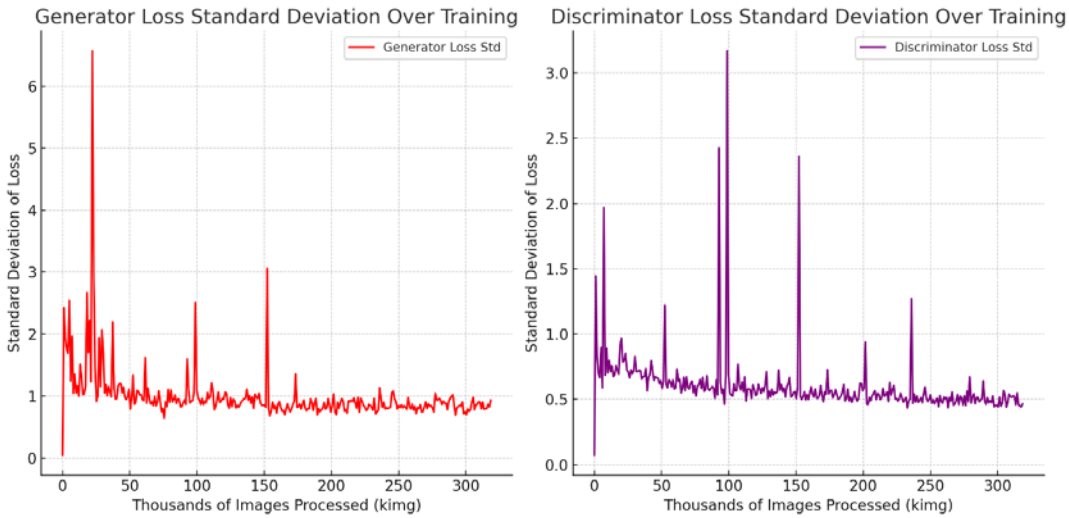

Figure 5.22 StyleGAN masked G and D Loss STD.

Source: Author

For this experiment, the same image format was used as for the previous StyleGAN experiment. The results can be observed through Figures 5.18, 5.19, and 5.20. Final outputs illustrate the improved output of a StyleGAN model when properly segmented images are used as input. Unlike previous attempts with unsegmented data, these results are much clearer and more detailed, with the generated garments showing well-defined shapes, textures, and features. The segmentation allows the model to focus on the essential elements of each garment, resulting in higher-quality and more realistic images. The training graphs in Figures 5.21 and 5.22 for this Dataset of segmented images indicate a more stable and effective training process than the previous Dataset. The generator's loss decreases and stabilizes at a lower value with fewer fluctuations, suggesting improved performance in generating segmented images. The discriminator maintains its effectiveness throughout the training, further supporting the improved performance of the generator. This indicates that the StyleGAN model is producing better results with this segmented image dataset.

Table 5.8 FID50k Comparison Table for GAN Results

GAN Architecture	FID50k Start Epoch	FID50k Halftime Epoch	FID50k Final Epoch
CycleGAN	386.8562	114.5621	97.5456
CycleGAN masked	313.2896	73.7568	61.3201
StyleGAN	420.5244	351.4369	321.0567
StyleGAN masked	268.1456	38.6801	8.8481

Source: Author

Table 5.8 compares the performance of used GAN architectures with and without image segmentation as measured by the FID50k across different stages of training (Start Epoch, Halftime Epoch, and Final Epoch). Lower FID50k scores indicate better performance. These results underscore the critical role of image segmentation in enhancing GAN performance, particularly with StyleGAN. The experiments demonstrate that both CycleGAN and StyleGAN benefit from using masked images, which focus the model's attention on the relevant features, leading to better and faster convergence.

6. CHAPTER: CONCLUSION AND FUTURE WORK

In this thesis, the performance of CycleGAN and StyleGAN models was rigorously assessed across different configurations to determine their effectiveness in managing images with varying background complexities. Initially, both models were given the job of analyzing images that included background elements, which resulted in higher FID scores and noticeable differences from the actual image distributions.

The removal of backgrounds from the images resulted in significant enhancements, allowing the models to focus more on the primary subjects and reducing the complexity of their operations. This, in turn, improved their ability to generate content.

When comparing the two models under these modified conditions, the StyleGAN model demonstrated a clear advantage, as indicated by significantly lower FID scores compared to the CycleGAN model. This demonstrates the improved capacity of StyleGAN to accurately capture and reproduce intricate image details and textures, leading to outputs that are more lifelike and visually appealing.

Although CycleGAN demonstrates promise in unpaired image-to-image translation, StyleGAN proves to be more efficient in producing high-quality synthetic images when background noise is reduced. This result emphasizes the crucial requirement to choose particular GAN structures that are customized to the requirements and conditions of the task. Subsequent inquiries could concentrate on enhancing these models, broadening their utilization in more regulated environments, or optimizing their effectiveness in real-world scenarios.

Furthermore, the experiments with the Neural Transfer approach using a pre-trained VGG19 model with ImageNet weights have shown promising results, especially in the realm of garment style transfer. This model exhibited exceptional capabilities in maintaining the fidelity and intricacies of design patterns during the transfer process, facilitated by strategic training across multiple epochs and controlled learning rate decay.

The success of this Neural Transfer approach underscores its potential as a transformative tool for the creative industries, enhancing the automation of design processes and enabling personalized fashion design. Continued development and refinement of this technique could extend its applications to other domains requiring precise pattern replication or style transfer, thereby boosting both creative potential and production efficiency.

In conclusion, these experiments collectively demonstrate that fashion designers can effectively utilize the outputs generated by these advanced modeling techniques to enrich their creative processes. GANs and Neural Transfer approaches offer designers a robust tool to experiment with new ideas and efficiently visualize concepts by reliably transferring and transforming patterns and styles. This feature not only amplifies the ingenuity of individual designers by broadening the spectrum of visual representations at their disposal but also contributes significant worth by expediting the design process and facilitating more dynamic engagement with the design components. The ongoing advancement of these technologies has the potential to completely transform the fashion industry by changing how designers approach the creative process. This integration could stimulate innovation and potentially redefine the limits of fashion design.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this thesis the author used Quillbot / Paraphraser in order to enhance the clarity and readability of the academic text, ensuring that complex ideas were conveyed more effectively. After using this tool/service, the author and advisor reviewed and edited the content as needed and took full responsibility for the content of the thesis.

7. REFERENCES

Borji, A. (2019). Pros and cons of GAN evaluation measures. Computer Vision and Image Understanding, 179, 41–65. https://doi.org/10.1016/j.cviu.2018.10.009

Gatys, L., Ecker, A., & Bethge, M. (2016). A Neural Algorithm of Artistic Style. Journal of Vision, 16(12), 326. https://doi.org/10.1167/16.12.326

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative Adversarial Networks. http://arxiv.org/abs/1406.2661

Gui, J., Sun, Z., Wen, Y., Tao, D., & Ye, J. (2020). A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications. http://arxiv.org/abs/2001.06937

Hashmi, M. F., Ashish, B. K. K., Keskar, A. G., Bokde, N. D., & Geem, Z. W. (2020). FashionFit: Analysis of Mapping 3D Pose and Neural Body Fit for Custom Virtual Try-On. IEEE Access, 8, 91603–91615. https://doi.org/10.1109/ACCESS.2020.2993574

Hekimoğlu, C. (2017). DEMAND PREDICTION IN CLOTHING INDUSTRY WITH USING NEURAL NETWORKS.

Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., & Hochreiter, S. (2017). GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium. http://arxiv.org/abs/1706.08500

Hsiao, W. L., & Grauman, K. (2017). Learning the Latent "Look": Unsupervised Discovery of a Style-Coherent Embedding from Fashion Images. Proceedings of the IEEE International Conference on Computer Vision, 2017-October, 4213–4222. https://doi.org/10.1109/ICCV.2017.451

Isola, P., Zhu, J.-Y., Zhou, T., Efros, A. A., & Research, B. A. (2018). Image-to-Image Translation with Conditional Adversarial Networks. https://github.com/phillipi/pix2pix.

Karras, T., Laine, S., & Aila, T. (2018). A Style-Based Generator Architecture for Generative Adversarial Networks. http://arxiv.org/abs/1812.04948

Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., & Aila, T. (2019). Analyzing and Improving the Image Quality of StyleGAN. http://arxiv.org/abs/1912.04958

Kato, N., Muramatsu, N., Osone, H., Ochiai, Y., & Sato, D. (2018). DeepWear: A case study of collaborative design between human and artificial intelligence. TEI 2018 - Proceedings of the 12th International Conference on Tangible, Embedded, and Embodied Interaction, 2018-January, 529–536. https://doi.org/10.1145/3173225.3173302

Kato, N., Osone, H., Oomori, K., Ooi, C. W., & Ochiai, Y. (2019, March 11). Gans-based clothes design: pattern maker is all you need to design clothing. ACM International Conference Proceeding Series. https://doi.org/10.1145/3311823.3311863

Kim, J., & Forsythe, S. (2008). Adoption of virtual try-on technology for online apparel shopping. Journal of Interactive Marketing, 22(2), 45–59. https://doi.org/10.1002/dir.20113

Kobayashi, S. (1991). Color Image Scale. Kosdansha International.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. Communications of the ACM, 60(6), 84–90. https://doi.org/10.1145/3065386

Lao, B., & Jagadeesh, K. (2014a). Convolutional Neural Networks for Fashion Classification and Object Detection.

Lao, B., & Jagadeesh, K. (2014b). Convolutional Neural Networks for Fashion Classification and Object Detection. 7. http://cs231n.stanford.edu/reports/BLAO_KJAG_CS231N_FinalPaperFashionClassification.pdf

Lecun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation Applied to Handwritten Zip Code Recognition.

Li, Y., Wang, N., Liu, J., & Hou, X. (2017). Demystifying Neural Style Transfer. http://arxiv.org/abs/1701.01036

Liu, L., Zhang, H., Ji, Y., & Jonathan Wu, Q. M. (2019). Toward AI fashion design: An Attribute-GAN model for clothing match. Neurocomputing, 341, 156–167. https://doi.org/10.1016/j.neucom.2019.03.011

Liu, Z., Luo, P., Qiu, S., Wang, X., & Tang, X. (2016). DeepFashion: Powering Robust Clothes Recognition and Retrieval with Rich Annotations. https://www.google.com/imghp

Liu, Z., Yan, S., Luo, P., Wang, X., & Tang, X. (2016). Fashion Landmark Detection in the Wild. http://arxiv.org/abs/1608.03049

Luan, S., Chen, C., Zhang, B., Han, J., & Liu, J. (2018). Gabor Convolutional Networks. IEEE Transactions on Image Processing, 27(9), 4357–4366. https://doi.org/10.1109/TIP.2018.2835143

Lucic, M., Kurach, K., Michalski, M., Gelly, S., & Bousquet, O. (2017). Are GANs Created Equal? A Large-Scale Study. http://arxiv.org/abs/1711.10337

Mirza, M., & Osindero, S. (2014). Conditional Generative Adversarial Nets. 1–7. http://arxiv.org/abs/1411.1784

Parmar, G., Zhang, R., & Zhu, J.-Y. (2021). On Aliased Resizing and Surprising Subtleties in GAN Evaluation. http://arxiv.org/abs/2104.11222

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., Facebook, Z. D., Research, A. I., Lin, Z., Desmaison, A., Antiga, L., Srl, O., & Lerer, A. (2017). Automatic differentiation in PyTorch.

Qin, X., Zhang, Z., Huang, C., Dehghan, M., Zaiane, O. R., & Jagersand, M. (2020). U^2-Net: Going Deeper with Nested U-Structure for Salient Object Detection. https://doi.org/10.1016/j.patcog.2020.107404

Raffiee, A. H., & Sollami, M. (2020). Garmentgan: Photo-realistic adversarial fashion transfer. Proceedings - International Conference on Pattern Recognition, 3923–3930. https://doi.org/10.1109/ICPR48806.2021.9412908

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. http://arxiv.org/abs/1505.04597

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., & Chen, X. (2016). Improved Techniques for Training GANs. https://github.com/openai/improved-gan.

Sun, W., Bappy, J. H., Yang, S., Xu, Y., Wu, T., & Zhou, H. (2019). Pose Guided Fashion Image Synthesis Using Deep Generative Model. http://arxiv.org/abs/1906.07251

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 07-12-June, 1–9. https://doi.org/10.1109/CVPR.2015.7298594

Yang, C., Zhou, Y., Zhu, B., Yu, C., & Wu, L. (2021). Emotionally intelligent fashion design using CNN and GAN. Computer-Aided Design and Applications, 18(5), 900–913. https://doi.org/10.14733/cadaps.2021.900-913

Yildirim, G., Jetchev, N., Vollgraf, R., & Bergmann, U. (2019). Generating High-Resolution Fashion Model Images Wearing Custom Outfits. http://arxiv.org/abs/1908.08847

Yu, Y., Wang, C., Fu, Q., Kou, R., Huang, F., Yang, B., Yang, T., & Gao, M. (2023). Techniques and Challenges of Image Segmentation: A Review. In Electronics (Switzerland) (Vol. 12, Issue 5). MDPI. https://doi.org/10.3390/electronics12051199

Yuan, C., & Moghaddam, M. (2020). Attribute-Aware Generative Design with Generative Adversarial Networks. IEEE Access, 8, 190710–190721. https://doi.org/10.1109/ACCESS.2020.3032280

Zhang, R., Isola, P., Efros, A. A., Shechtman, E., & Wang, O. (2018). The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 586–595. https://doi.org/10.1109/CVPR.2018.00068

Zhu, J.-Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. http://arxiv.org/abs/1703.10593

APPENDIX

APPENDIX 1.RANDOM OUTPUTS

Figure 0.1 Random seed between 10000-10060

Source: Author

Figure 0.2 Random seed between 20000-20060

Source: Author

APPENDIX 2. CURATED CLOSE UPS

Figure 0.3 Masked StyleGAN30000 Seed1409

Source: Author

Figure 0.4 Masked StyleGAN30000 Seed6257

Source: Author

Figure 0.5 Masked StyleGAN30000 Seed6452

Source: Author

Figure 0.6 Masked StyleGAN30000 Seed1832

Source: Author