AI Virtual Wardrobe

Submitted By

Smiti Kothari

20BCE280

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INSTITUTE OF TECHNOLOGY

NIRMA UNIVERSITY

AHMEDABAD-382481

May 2024

AI Virtual Wardrobe

Project Report

Submitted in partial fulfillment of the requirements

for the degree of

Bachelor of Technology in Computer Science and Engineering

By

Smiti Kothari

(20BCE280)

Guided By

Dr. Jaiprakash Verma

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INSTITUTE OF TECHNOLOGY

NIRMA UNIVERSITY

AHMEDABAD-382481

May 2024

Certificate

This is to certify that the major project entitled **"AI Virtual Wardrobe"** submitted by **Smiti Kothari (20BCE280)**, towards the partial fulfillment of the requirements for the award of the degree of Bachelor of Technology in Computer Science and Engineering, Nirma University, Ahmedabad, is the record of work carried out by him/her under my supervision and guidance. In my opinion, the submitted work has reached the level required for being accepted for examination.

Dr Jaiprakash Verma
Associate Professor
Department of Computer Science & Eng.
Institute of Technology,
Nirma University, Ahmedabad.

Dr Madhuri Bhavsar
Professor and Head,
Department of Computer Science & Eng.
Institute of Technology,
Nirma University, Ahmedabad

Statement of Originality

I, **Smiti Kothari**, Roll. No. **20BCE280**, give an undertaking that the major project entitled **"AI Virtual Wardrobe"** submitted by me, towards the partial fulfilment of the requirements for the degree of Bachelor of Technology in **Computer Science and Engineering**, Nirma University, Ahmedabad, contains no material that has been awarded for any degree or diploma in any university or school in any territory to the best of my knowledge. It is the original work carried out by me and I give assurance that no attempt of plagiarism has been made. It contains no material that is previously published or written, except where reference has been made. I understand that in the event of any similarity found subsequently with any published work or any dissertation work elsewhere; it will result in severe disciplinary action.

Signature of Student
Date:
Place:Ahmedabad

Endorsed by
Dr Jaiprakash Verma
(Signature of Guide)

Acknowledgements

I want to extend my gratitude to everyone who contributed to the success of this project and made it an incredible learning journey.

Firstly, a heartfelt thank you to Dr. Jaiprakash Verma, our guide. His continuous guidance, feedback, and unwavering support were instrumental in shaping our project. Dr. Verma's expertise challenged me to perform at my best and helped me grow professionally.

I also want to express my appreciation to our external guides, Mr. Vivek Solanki and Ms. Honey Jani, for their constant monitoring and valuable advice throughout the project.

Furthermore, I am deeply thankful for my teammate, Jayneel Shah. Her insightful contributions and collaborative spirit significantly improved the quality of our work. Together, we navigated challenges and capitalized on opportunities, creating a synergy that made our project impactful.

To all involved, thank you for making this journey so memorable. Working alongside such dedicated and talented individuals has been a privilege.

- Smiti Kothari

20BCE280

Abstract

The AI-Virtual Wardrobe project presents an approach to use computer vision techniques to transform the process of trying on clothes and accessories. Using computer vision techniques and artificial intelligence, the project seeks to improve and automate the virtual try-on experience. The main goals are to achieve high accuracy in garment detection and segmentation, placement of accessories, build an intuitive interface for seamless interaction, and apply posture estimation to precisely adjust clothing items to the wearer's body. The scope of the project extends to include accessories such as necklaces, earrings, lipstick, hair wigs, and glasses, ensuring a comprehensive virtual try-on experience. This comprehensive approach encompasses preprocessing images, identifying objects and positions, warping garments and accessories, and creating a user-friendly web application that integrates all elements seamlessly. The AI-Virtual Wardrobe project represents a significant advancement in personal style exploration and clothing customization, with the potential to enhance user experience and streamline workflows in the fashion industry.

Abbreviations

BI	Business Intelligence
YOLO	You Only Look Once
ACGPN	Adaptive Content Generating and Preserving Network
GMM	Geometric Matching Module
TOM	Try On Module
API	Application Programming Interface
IUV	Index for Unwarping Vertices
TPS	Thin Plate Spline

–

Contents

Certificate iii

Statement of Originality iv

Acknowledgements v

Abstract vi

Abbreviations vii

List of Tables x

List of Figures xi

1 Introduction 1
1.1 About The Company 1
1.1.1 Company Details 1
1.1.2 Quality Policy 2
1.1.3 Communication 2
1.1.4 Resources 3
1.2 Project Profile 3
1.2.1 Project Title 3
1.2.2 Scope of Project 3
1.2.3 Project Team 4
1.2.4 System Requirements 5
1.3 The System 5
1.3.1 Definition of the System 5
1.3.2 Purpose and objectives 5
1.3.3 About present system 7
1.3.4 Proposed system 7

2 System Analysis 10
2.1 Feasibility Study 10
2.1.1 Operational Feasibility 10
2.1.2 Financial and Economical feasibility 11
2.1.3 Schedule Feasibility Study 11
2.1.4 Legal and Ethical Feasibility Study 12
2.1.5 Handling Infeasible Projects. 12
2.2 Approaches Used for each Try on 12

2.2.1 Methods of Warping Used in Cloth Try on Module 12
2.2.2 Necklace Try on Module 14
2.2.3 Earring Try on Module 15
2.2.4 Glasses Try on Module 16
2.2.5 Lipstick Try on Module 17
2.2.6 Hairwig Try on Module 18

3 System Design 20
3.1 System Architecture for Cloth Try on 20
3.2 System diagram used in Warping Stage 22
3.2.1 U2net architecture used in Warping Model 23
3.3 Dlib and Haarcascade Architecture used in Accesory Try on 25
3.4 System architecture for Hairwig Try on 26

4 User Manual 29
4.1 Prerequisites 29
4.2 Tools and Technology Stack 29
4.3 Installation & Usages 30
4.4 Features 31

5 Results and Discussion 32
5.1 Results 32
5.1.1 Virtual Cloth Try On 32
5.1.2 Virtual Necklace Try On 33
5.1.3 Virtual Earring Try On 33
5.1.4 Virtual Glass Try On 34
5.1.5 Virtual Lipstick Try On 34
5.1.6 Virtual Hairwig Try On 34
5.1.7 Flask Web App 35
5.2 Discussion 37

6 Conclusion and Future Work 39
6.1 Conclusion 39
6.2 Future Work 39

Bibliography 41

List of Tables

1.1 Hardware Requirements 5
1.2 Software Requirements 5
1.3 Present System 9

List of Figures

3.1 System Diagram 20
3.2 Warping Diagram [1] 22
3.3 U2net Diagram 24
3.4 Haarcascade Features 26
3.5 Flowchart for face and eye detection 26
3.6 Flowchart for neck and lip detection 27
3.7 System diagram for hairwig try on 28

5.1 Cloth Try On Output 32
5.2 Necklace Try On Output 33
5.3 Earring Try On Output 33
5.4 Glasses Try On Output 34
5.5 Lipstick Try On Output 35
5.6 Hairwig Try On Output 35
5.7 Home Page of Application 36
5.8 Try On Page of Application 36
5.9 Steps to use the website 37

Chapter 1

Introduction

1.1 About The Company

1.1.1 Company Details

Established and founded in 2015, F(x) Data Labs Pvt. Ltd., now known as FxisAi, is recognized as India's first dedicated data science company[2]. Based in the city of Ahmedabad, Gujarat, the company has been at the forefront of technological innovation in data science and artificial intelligence.

With a focus on Cryptocurrency, Technology, Analytics & BI, Machine Learning, and Full Stack Development, F(x) Data Labs has carved out a niche for itself as a leader in these domains. The company operates on a versatile business model that includes on-demand developers, project outsourcing, and tech/talent partnerships.

The company boasts a global presence, having successfully spearheaded over 400 projects and serving more than 200 clients worldwide, including Fortune 50 companies [3]. The company's clientele spans various industries, from startups to large enterprises, reflecting its ability to adapt and cater to a diverse range of needs.

The company is backed by a robust team of over 75 skilled professionals, including developers, data scientists, and creative experts [3]. This cross-functional team is committed to delivering top-notch technical, business, and creative solutions that are tailored to the specific needs of each client.

1.1.2 Quality Policy

The company is committed to delivering the highest quality of work, ensuring that each project not only meets but surpasses client expectations. The company's dedication to excellence is reflected in its meticulous approach to project delivery and its unwavering focus on client satisfaction.

At the heart of Fx Data Lab's quality policy is a client-centric approach that prioritizes the needs and goals of the clients. The company engages with clients to understand their unique challenges and tailors solutions that are both innovative and effective.

To maintain its high standards, the company employs rigorous quality assurance practices throughout the project lifecycle. This includes regular reviews, testing, and feedback loops that ensure the final deliverables are of the highest caliber.

The quality of work is second to none, as evidenced by the glowing testimonials from global partners and clients [3]. The company has earned accolades for its exceptional project delivery and the trust it has built with esteemed clients like Johnson & Johnson and Guinness World Records.

Technological Expertise With a team of highly skilled professionals, the company leverages cutting-edge technologies to deliver cost-effective solutions. The company's expertise in data science, machine learning, artificial intelligence, and full-stack development positions it as an industry leader in delivering quality solutions[4].

1.1.3 Communication

Fx Data Labs prioritizes understanding client needs and fostering collaboration. With a robust feedback mechanism and seamless communication channels, the team ensures continuous improvement and client satisfaction, ensuring efficient project delivery. The key points are:

- Client Engagement: The company places immense value on engaging with clients to understand their needs and objectives1. The company fosters a collaborative environment that encourages open dialogue and transparency.

- Feedback Mechanism: A robust feedback mechanism is in place, allowing clients to share their experiences and suggestions. This feedback is integral to the company's continuous improvement and client satisfaction [5].

- Effective Communication: The team ensures seamless communication through virtual meetings and prompt responses, facilitating a smooth workflow and timely project delivery[5].

1.1.4 Resources

The company maintains its position at the forefront of data science and AI innovation by leveraging cutting-edge technology and a strong commitment to research and development, which underpins its excellence and industry advancement.

- Talent Pool: The company has a huge talent pool of skilled professionals, each bringing their expertise to the table. This diverse team is the backbone of the company's innovative solutions.

- Technological Infrastructure: The company utilizes latest technologies to develop and deliver state-of-the-art solutions[3].

- Research & Development: The R&D efforts contribute significantly to the company's growth and the advancement of the industry.

1.2 Project Profile

1.2.1 Project Title

AI Virtual Wardrobe: Automated Clothing Change using Computer Vision: The program intends to accurately and efficiently automate the selection and fitting of clothing on individuals using computer vision and artificial intelligence.

1.2.2 Scope of Project

- **Objectives**
  - Create a user-friendly interface that allows for easy navigation and interaction with the application.
  - Apply computer vision techniques to achieve high levels of accuracy and precision in clothing detection, segmentation, and posture estimation.
  - Allow users to virtually try on garments and accessories and see how they look on their bodies to make more informed decisions.

- Adjust clothing to match the user's position, increasing the realism of the virtual try-on experience.

- **Deliverables**
  - User-friendly application interface with intuitive navigation and interaction features.
  - Efficient results for garment detection, segmentation, and pose estimation, utilizing models such as U-Net, YOLO, DeepLab V3, and MoveNet.
  - Virtual try-on feature allowing users to visualize clothing items on their bodies realistically.

- **Features and Requirements**
  - Development of computer vision algorithms for clothing change automation, including image segmentation, object detection, landmark detection, and geometric transformations.
  - Utilization of U-Net and YOLO architectures for garment detection and segmentation.
  - Implementation of DeepLab V3 and U-Net models for semantic segmentation to isolate individuals and garments from backgrounds.
  - Application of Thin Plate Spline transformation techniques, ACGPN, and GANs for garment warping and adaptation.

1.2.3 Project Team

The project team was made up of six people, whose details are mentioned below:

- Parth Maru (FS - Software Engineer II) - Superstandup Tech Lead
- Vivek Solanki (ML - Software Engineer II) - Tech Lead and Project Guide
- Honey Jani (Data Scientist) - Tech Lead and Project Guide
- Vishvdeep Dasadiya (Data Science Engineer and Manager) - Growth Manager
- Jayneel Shah (Machine Learning Intern) - Developer
- Smiti Kothari (Machine Learning Intern) - Developer

1.2.4 System Requirements

The hardware and software requirements for the project are listed in the tables below.

Processor	Intel i5 @ 2.20GHz
Hard Disk	40 GB(min)
RAM	16 GB

Table 1.1: Hardware Requirements

Operating System	Windows 8 and above
Language	Python
Framework	Flask
Tools	Tensorflow
Editors	Google Colab and VS Code

Table 1.2: Software Requirements

1.3 The System

1.3.1 Definition of the System

The system is an AI-driven virtual makeover platform that leverages cutting-edge computer vision technology, particularly employing Generative Adversarial Networks (GANs), deep convolutional neural networks (CNNs), and various other techniques to provide users with an immersive and personalized virtual try-on experience for fashion accessories and clothing items.Whether the user is curious about a new pair of earrings or want to test out a different hairstyle, this system lets the user see how it all looks on the user before making any decisions. This not only saves time but also reduces the typically high return rates associated with online shopping.

1.3.2 Purpose and objectives

Purpose:

The primary purpose of the system is to revolutionize the traditional fashion retail experience by offering users a seamless and highly interactive virtual try-on process. By using the power of AI and computer vision, the system aims to provide users with the ability to virtually experiment with different styles, hairstyles, lipsticks, accessories like necklaces and earrings, and even clothing items. Instead of the traditional, sometimes cumbersome process of browsing through racks of clothing or endless web pages, this system introduces a seamless and highly interactive virtual try-on process. This system isn't just about browsing static images or descriptions—it's about experiencing fashion in a whole new way. With lifelike virtual simulations, users can see exactly how different clothing items and accessories look on themselves in real-time.this system aims to make the online shopping experience more efficient and enjoyable. By providing users with a highly user friendly and intuitive platform, it reduces the guesswork and uncertainty often associated with purchasing clothing and accessories online. This, in turn, can help to minimize the likelihood of returns, saving both time and hassle for shoppers. Moreover, the users without leaving their homes can curate their own unique fashion experiences, all from the comfort of their own device.

Objective:

- Enhanced User Experience: The system endeavors to provide an intuitive and engaging platform for users to explore various fashion choices, thereby enhancing their overall shopping experience.

- Efficiency through Automation: By automating the try-on process using AI algorithms, the system aims to streamline the selection process, saving users time and effort typically spent in physical try-on sessions.

- Personalization: The system aims to empower users by allowing them to customize their wardrobe choices based on their preferences and individual style, thereby catering to diverse tastes and preferences.

- Seamless Virtual Try-On Capabilities: Through the integration of advanced computer vision techniques such as haarcascade classifier algorithms, dlib predictors, thin plate spline transformation, and GMM and TOM models, the system seeks to provide users with realistic and accurate virtual representations of the chosen fashion items on their body.

1.3.3 About present system

In this section, an overview of the research conducted by other researchers in the field is provided. The findings from various studies are included in Table 1.3, to offer a comprehensive understanding of the subject matter.

1.3.4 Proposed system

The proposed system employs a combination of sophisticated AI models and computer vision techniques to achieve its objectives:

- Virtual Makeover with GAN Technology: The system utilizes advanced GAN technology to generate lifelike virtual representations of hairstyles, lipsticks, necklaces, and earrings, allowing users to experiment with different looks virtually. From trying out a new hairstyle to testing bold lipstick shades, the possibilities are virtually endless.

- Accurate Trying with KeyPoint Based Warping and GMM/TOM Models: To ensure that each virtual try-on experience is tailored to the user's unique body shape and pose, the system employs keypoint-based warping and GMM/TOM models, the system ensures accurate alignment and customization of clothing items to match the user's body shape and pose, thereby enhancing accuracy.

- Efficient Try-On Process with Automation: Automation techniques such as image segmentation, landmark detection, and resizing enable the system to make the try-on process more efficient, enhancing convenience for users. Whether they're browsing through tops or accessories, the system automates the process, saving users time and effort.

- Seamless Integration of Accessories: Through techniques like facial landmark detection and HaarCascade classifiers, the system seamlessly integrates accessories such as necklaces, earrings, and glasses onto the user's image, ensuring a natural and realistic virtual try-on experience.

- Enhanced Visual Quality with Image Processing: Utilizing methods like image cropping, transformation, fast marching, and inpainting, the system enhances the visual quality of virtual try-ons by filling in missing parts and ensuring seamless transitions between different fashion items.

- Interactive User Interface: Users can simply upload their image and select the fashion item they wish to try on by clicking on the respective icon. Upon selection, the system swiftly processes the image and applies the chosen item, providing users with a very fast preview of their new look. Furthermore, users have the option to download and save the augmented image directly and then share it on various social media platforms with just a few clicks.

Reference	Objective	Methodology	Findings	Limitations
X. Han et al. [6]	Introduces the VITON system, combining image-based rendering and neural networks to generate realistic clothing pictures on users.	Utilizes image-based rendering and neural networks to create realistic clothing images on users.	Demonstrates the effectiveness of the VITON system in producing high-quality virtual try-on outcomes.	Lacks real-world applicability analysis.
Kedan Li et al. [7]	Introduces quantitative evaluation methods for virtual try-on using a novel dataset.	Utilizes learned embedding to automatically select target models and trains multiple specialized warpers. Combines warps with a U-Net architecture.	Demonstrates significant improvements in virtual try-on results, encompassing outline, texture shading, and garment details.	Quantitative evaluation may not fully capture subjective aspects of try-on quality.
E. Yanmaz et al. [8]	Provides an overview of virtual try-on methods and covers difficulties such as precise body measurements and fabric simulations.	Literature review summarizing virtual try-on methods including 3D modeling, texture mapping, and image-based rendering.	Offers insights into various virtual try-on techniques.	Lacks detailed analysis on specific methodologies
Y. R. Cui et al. [9]	Introduces an end-to-end virtual garment display method based on cGAN for efficient showcasing of garment designs.	Proposes FashionGAN, requiring user input of a fashion sketch and fabric image to generate virtual garment images automatically.	Demonstrates FashionGAN's effectiveness in generating high-quality virtual garment images.	Limited discussion on potential challenges in real-world implementation.
Sangho Lee et al. [10]	Introduces a method to enhance clothing detail in virtual try-on through an additional fitting step after geometric warping.	Implements a fitting step post-geometric warping to separate clothing from the wearer, preserving clothing structure and details. .	Empirically confirms effective disentanglement of clothing from the wearer and detail preservation. Compares with existing methods using various metrics.	Limited exploration of discussion on computational complexity.

Table 1.3: Present System

Chapter 2

System Analysis

2.1 Feasibility Study

2.1.1 Operational Feasibility

Operational feasibility assessment involves evaluating whether the proposed project aligns effectively with the operational processes and objectives of the organization. This evaluation considers various factors, including user needs and identifies the intended users of the tool, such as fashion enthusiasts, shoppers, or retailers. It is essential to gain an understanding of their requirements, challenges, and expectations regarding virtual try-on experiences. Moreover, defining the essential features and functionality needed for the AI Virtual Wardrobe tool is important. This encompasses functionalities such as virtual try-on capabilities for clothing items and accessories and real-time visualization of different looks. Furthermore, evaluating the tool's capability to access and handle the necessary data, such as clothing or accessory images, user images and preferences is crucial. This involves assessing the feasibility of integrating computer vision technologies, AI algorithms, and data retrieval methods to gather and analyze huge relevant data effectively. Considerations must also be made for data security, privacy, and compliance with regulatory requirements, ensuring the protection of user information and sensitive data. Additionally, the assessment includes evaluating the scalability of the tool and assessing the tool's performance to ensure optimal functionality. Moreover, the user interface and user experience design of the tool plays a significant role in its operational feasibility. The tool should offer an intuitive and visually appealing interface that allows users to navigate seamlessly through different features and functionalities.

2.1.2 Financial and Economical feasibility

An assessment of project schedule feasibility determines whether the software project can be completed within the set time frame. This evaluation considers factors like the project's scope, available resources, chosen development approach, and potential risks. It involves identifying and calculating the expenses linked to the project, covering costs for research materials, equipment, software licenses, employee salaries, and other relevant expenditures. Both one-time and recurring costs are taken into account. Furthermore, sustainability and viability are crucial aspects of this assessment, focusing on the project's long-term feasibility and viability. This includes examining ongoing maintenance costs, resource needs, and the potential for growth or expansion. Additionally, a comprehensive risk analysis is performed to identify potential project risks and uncertainties. The impact of these risks on the project's financial feasibility is carefully evaluated, and strategies are developed to address them. Risks could include inadequate funding, limited resources, or unexpected challenges that may hinder progress. An assessment of project schedule feasibility determines whether the software project can be completed within the set time frame. This evaluation considers factors like the project's scope, available resources, chosen development approach, and potential risks. It involves identifying and calculating the expenses linked to the project, covering costs for research materials, equipment, software licenses, employee salaries, and other relevant expenditures. Both one-time and recurring costs are taken into account. Furthermore, sustainability and viability are crucial aspects of this assessment, focusing on the project's long-term feasibility and viability. This includes examining ongoing maintenance costs, resource needs, and the potential for growth or expansion. Additionally, a comprehensive risk analysis is performed to identify potential project risks and uncertainties. The impact of these risks on the project's financial feasibility is carefully evaluated, and strategies are developed to address them. Risks could include inadequate funding, limited resources, or unexpected challenges that may hinder progress.

2.1.3 Schedule Feasibility Study

Assessing schedule feasibility involves determining if the software project can be completed within the given time frame. This evaluation considers factors such as the project's scope, resource availability, development approach, and potential risks. It helps determine whether the project can meet the deadline as planned or if adjustments are needed to ensure timely completion.

2.1.4 Legal and Ethical Feasibility Study

The objective of this review is to evaluate whether the project complies with legal and ethical standards. This includes examining potential legal concerns related to data privacy, intellectual property rights, and other relevant regulations. Moreover, ethical considerations are considered to ensure that the tool upholds ethical norms and protects user privacy. This feasibility study enables the project to confirm its compliance with relevant laws and ethical guidelines.

2.1.5 Handling Infeasible Projects.

Examining how to handle infeasible projects involves a thorough assessment and decision-making process. The main goal is to identify the specific reasons why the project is considered unfeasible, such as technical limitations, limited resources, or the unavailability of necessary data or APIs. Exploring alternative approaches or solutions becomes essential to achieve the desired outcome within the given limitations. This may include considering different technologies, frameworks, or methodologies that are more suitable or readily available. Assessing the possibility of adjusting the project's scope or requirements is also important. This could involve prioritizing essential features, simplifying functionalities, or narrowing down the target audience to reduce complexity and resource demands. It's crucial to determine if any compromises or trade-offs can be made to salvage the project, such as accepting certain limitations or finding workarounds to overcome specific obstacles.

2.2 Approaches Used for each Try on

This section outlines the techniques employed for each try-on to achieve the desired outcome.

2.2.1 Methods of Warping Used in Cloth Try on Module

- Index for Unwarping Vertices(IUV): It basically means aligning clothing onto a user's image by assigning indices to vertices or points on the garment and the corresponding points on the user's body. These indices correspond to corresponding points on the user's body, allowing for precise placement of clothing items during the virtual try-on process.

- Thin Plate Spline Transformation(TPS): It is used to deform the shape of the clothing to match the shape of the user's body. It involves defining control points on both the clothing and the user's body. This is essential for achieving a natural fit of clothing items to the user's body shape.

- VGG for Warping: It Utilizes the VGG architecture for warping which enhances the visual quality of virtual try-ons. The VGG model is a deep convolutional neural network capable of extracting detailed features from clothing and user images. By leveraging these features, the model generates transformed clothing images that align with the user's body contours.

- KeyPoint Based Warping Technique: This technique relies on identifying key landmarks or points on both the clothing and the user's image to accurately warp the clothing to match the body's contours. By precisely aligning the clothing with the user's body shape, this technique ensures a more realistic and visually appealing virtual try-on experience.

- SeiveNet Model: The SeiveNet Model utilizes layers of neural networks to manipulate garment textures, shapes, and alignments to better fit the user's body. By dynamically adjusting textures and shapes, the model enhances the overall fit and appearance of clothing items.

- GMM and TOM Model: The combination of Geometric Matching Module (GMM) and Try on Module (TOM) is instrumental in accurately aligning and overlaying clothing onto the user's image. GMM ensures precise alignment of clothing with the user's body shape and pose, while TOM facilitates the overlaying of clothing onto the user's image, resulting in a seamless integration of garments into the virtual environment.

- ACGPN Model: It is a variation of Generative Adversarial Network (GAN) that focuses on controlling attributes within the generated content. The GAN is built on top of the Pix2Pix model,designed for high-resolution image-to-image translation tasks.

Benefits of ACGPN

The ACGPN architecture has demonstrated better accuracy and realism in virtual cloth try-on scenarios. Its advantages, which set it apart from others, include

- ACGPN utilizes a specialized variation of Generative Adversarial Network (GAN) tailored to control attributes within the generated content, focusing on manipulating specific aspects of the target person.

- It creates a controllable latent space, a multi-dimensional representation of data, allowing for precise adjustments and customization of generated images.

- The system is built on top of the Pix2Pix model, renowned for its effectiveness in high-resolution image-to-image translation tasks, ensuring detailed and realistic virtual try-on results.

- The system's architecture is designed for scalability, allowing it to handle large volumes of user data and accommodate future enhancements and updates to the virtual try-on platform.

- A special parsing module architecture is integrated into ACGPN, designed to enhance the accuracy of human parsing. This module improves the system's ability to segment and understand different components of an image, such as clothing, body parts, and background, contributing to more realistic virtual try-on outcomes.

2.2.2 Necklace Try on Module

- Fine-tuning DLIB Shape Predictor:
  - We fine-tune a pre-trained DLIB shape predictor using the iBUG 300-W dataset.
  - The iBUG 300-W dataset comprises 300 face images annotated with 68 facial landmarks, providing detailed key points on the face such as the corners of the eyes, nose, mouth, and chin, as well as points along the eyebrows and jawline.

- Utilization of Fine-tuned DLIB Shape Predictor:
  - The fine-tuned DLIB shape predictor model is utilized to precisely identify facial landmarks, including those in the neck area.

  - By leveraging the detected facial landmarks, we calculate the precise cropping points for the necklace image.

- Resizing and Positioning of Necklace Image:
  - Based on the detected facial landmarks and predefined adjustments, we resize and position the necklace image appropriately.
  - This resizing and positioning ensure that the necklace image aligns accurately with the user's neck region.

- Overlaying Necklace Image onto User:
  - Subsequently, the resized necklace image is overlaid onto the user's image.
  - The overlay process involves utilizing masking and bitwise operations to seamlessly integrate the necklace image with the user's image, ensuring a natural and realistic appearance.

2.2.3 Earring Try on Module

- Fine-tuning Pretrained Haarcascade Classifiers:
  - We fine-tune pretrained Haarcascade classifiers using the Labeled Faces in the Wild (LFW) dataset for face detection and the Nottingham Ear dataset for ear detection.
  - This process involves adjusting the parameters of the classifiers to improve their accuracy in detecting faces and ears in images.

- Preprocessing of Data:
  - The dataset is preprocessed to create positive and negative samples. Positive samples represent the target part (e.g., faces, ears), while negative samples represent the background.
  - All positive and negative samples are resized to a uniform size, matching the input size required by the Haar Cascade classifiers. This ensures consistency in the training data and improves the classifiers' performance.

- Utilization of HaarCascade Frontal Face Classifiers:

  - HaarCascade frontal face classifiers, trained on face images, are employed to detect the user's face in the input image.
  - Once the face is detected, the dimensions of the detected face are determined, after which processing steps take place.

- Resizing and Overlaying Earring Images:

  - The left and right earring images are resized to fit the dimensions of the detected face.
  - Using the dimensions and position of the detected face, the system calculates the precise positions for placing the left and right earrings.
  - Then, the resized left and right earring images are overlaid onto the input image at the calculated positions.
  - This overlay process involves utilizing transparent masks to integrate the earring images with the user's image, ensuring a natural appearance.

2.2.4 Glasses Try on Module

- Fine-tuning Pretrained Haarcascade Classifiers:

  - Transfer learning techniques are applied to fine-tune pretrained Haarcascade classifiers for tasks, such as face and eye detection.
  - The Labeled Faces in the Wild (LFW) dataset is utilized for training the face detection classifier, while the UTFace dataset is used for training the eye detection classifier.

- Usage of Fine-tuned HaarCascade Classifiers:

  - A fine-tuned HaarCascade classifier, trained on frontal face images, is applied to accurately detect the user's face within the input image.
  - Another fine-tuned HaarCascade classifier, trained on eye images, is used to detect the eyes within the detected face region.

- Preprocessing for Positive and Negative Samples:

  - The dataset is preprocessed to generate positive and negative samples.

- Positive samples represent regions containing the target parts (faces, eyes), while negative samples represent background areas.
- Each sample is labeled to indicate its class-positive or negative.

- Uniform Resizing of Samples:
  - All positive and negative samples are resized to a equal dimension that matches the input size required by the Haar Cascade classifiers.
  - This uniform resizing ensures consistency in the dataset and prepares the samples for effective classification.
- Region of Interest (ROI) Identification:
  - Once the eyes are detected within the face region, regions of interest (ROIs) are identified based on the positions of the detected eyes.
- Resizing and Overlaying Glasses onto Eyes:
  - The size of the glasses image is dynamically adjusted to match the width of the detected eye.
  - Subsequently, the resized glasses image is overlaid onto the eyes of the user at the calculated positions, ensuring proper alignment and fit.

2.2.5 Lipstick Try on Module

- Landmark Detection for Lips:
  - The system utilizes a finetuned face detector and shape predictor from the dlib library to accurately detect landmarks corresponding to the person's lips in the input image.
  - These landmarks are key points that define the shape and position of the lips, providing important information for next processing steps.
- Color Application to Lips:
  - Once the lip landmarks are detected, the system applies a selected color to the lips by filling the region bounded by these landmarks.

- This process involves outlining the lip area based on the detected landmarks and filling it with the chosen color, ensuring accurate lip coloration.

- Gaussian Blur for Smoothing:
  - To enhance the blending of the colored lips with the original image and achieve a more natural appearance, Gaussian Blur is applied.
  - Gaussian Blur is a filtering technique used to smoothen the edges and transitions of the colored lips, reducing sharpness and improving overall visual quality.
  - By carefully adjusting the parameters of the Gaussian Blur, the system achieves optimal blending between the colored lips and the surrounding facial features, resulting in a realistic virtual makeup effect.

2.2.6 Hairwig Try on Module

- Image Segmentation using UNet Architecture:
  - The system uses the UNet architecture for image segmentation, dividing the input image into five distinct parts: Hair, Face, Ears, Neck, and Shirt.
  - UNet is a convolutional neural network (CNN) architecture commonly used for semantic segmentation tasks, allowing for precise segmentation of different regions within the image.
- Image Cropping Transformation with Face Landmark Detection:
  - Face landmark detection is performed on the segmented image to generate semantic maps, which serve as guiding point for cropping the regions of interest, particularly the face and hair.
  - By accurately identifying facial landmarks, such as the eyes, nose, and mouth, the system ensures precise cropping and transformation of the face and hair regions, increasing the virtual try-on experience.
- Fast Marching Method for Seamless Transition:
  - The Fast Marching Method is employed to fill in any missing parts of the image resulting from the cropping and transformation process.

  - This method facilitates a seamless transition between the original image and the modified version, ensuring that no inconsistencies remain in the final output.

- Image Replacing Inpainting with SDEdit Method:
  - After determining the source hair and target face regions, the system overlaps them to create a composite image.
  - A difference mask is calculated to identify and complete any missing parts of the face and hair, ensuring a perfect integration of the virtual hairstyle with the user's face.
  - The SDEdit method is utilized for image inpainting, effectively filling in the missing regions while maintaining visual realism.

Chapter 3

System Design

3.1 System Architecture for Cloth Try on

Figure 3.1: System Diagram

Each component in the cloth try-on system diagram (Figure 3.1) is explained below.

- User Interface (UI):
  - This is the front-facing part of the system where users interact with the application. It includes elements such as buttons, menus, and displays where users can upload images of themselves or select clothing items to try on.
  - The UI provides a seamless experience for users to navigate through different functionalities of the system.

- Preprocessing:
  - Before the images are fed into the main algorithms, preprocessing is done to enhance the quality of the input data.
  - This may involve resizing images to a standard resolution, normalization, and noise reduction to improve the performance of subsequent algorithms.
- Clothes Detection and Segmentation using U-Net:
  - U-Net is a convolutional neural network architecture commonly used for image segmentation tasks.
  - In this step, the system detects and segments the clothing items in the input image.
  - The output of this stage is a binary mask where pixels belonging to clothing items are labeled as foreground and background pixels are labeled as background.
- Pose Detection using OpenPose:
  - OpenPose is a popular library for real-time multi-person key-point detection.
  - This component analyzes the posture and pose of the user in the input image.
  - It identifies key points such as joints and body parts, which are crucial for accurately placing clothing items on the user's body.
- Cloth Warping:
  - Cloth warping involves deforming the clothing items detected in the input image to fit the pose and shape of the user.
  - Using the pose information obtained from the previous step, the system adjusts the position, orientation, and size of the clothing items accordingly.
  - Techniques such as geometric transformations and mesh warping may be employed to achieve realistic fitting of the clothes on the user's body.
- Try-On:

– This is the final stage where the warped clothing items are overlaid onto the input image of the user.

– The system generates a composite image showing the user wearing the selected clothing items.

– Users can visualize how the clothes look on them and make informed decisions before making a purchase or selecting an outfit.

3.2 System diagram used in Warping Stage

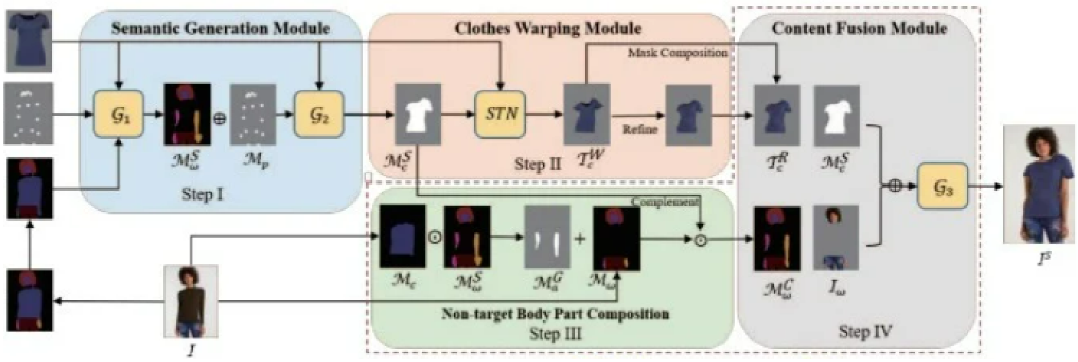


Figure 3.2: Warping Diagram [1]

The ACGPN, as outlined, consists of three modules, as depicted in Figure 3.2.

- Semantic Generation Module (SGM):
  – SGM begins by separating the target clothing region while preserving other body parts, like arms, without altering the pose or other human body details.
  – It adopts a two-stage approach:
  – First, it generates masks for body parts (such as arms) which help retain body part details. This is achieved by training a body parsing GAN (Generative Adversarial Network) to synthesize these masks based on inputs like pose maps, target clothing images, and fused maps indicating body part regions.
  – Second, it combines these body part masks with pose maps and target clothing images to synthesize masks specifically for the target clothing.
  – The module utilizes conditional GANs (cGANs) and pixel-wise cross-entropy loss to train the generator, ensuring accurate semantic segmentation.

- Clothes Warping Module (CWM):
  - CWM aims to fit the clothes to the shape of the target clothing region while retaining clothing characteristics.
  - It employs Thin-Plate Spline transformation (TPS) for warping but introduces a second-order difference constraint to ensure precise transformation, particularly for complex textures and colors [11].
  - By minimizing the loss function that includes geometric constraints, the module stabilizes the warping process and produces natural deformations.
- Content Fusion Module (CFM):
  - CFM integrates information from previous modules to adaptively determine how to generate or preserve distinct human parts in the synthesized image.
  - It addresses the challenge of layout adaptation by ensuring clear rendering of the target clothing region and preserving fine-scale details of body parts.
  - CFM includes:
    * **Non-target Body Part Composition**: Composites body masks to preserve non-targeted body parts while adaptively dealing with different cases, such as transferring short-sleeve clothes to long-sleeve.
    * **Mask Inpainting**: Utilizes masks to randomly remove parts of the arms in body images, then combines semantic information and generates realistic try-on images.

3.2.1 U2net architecture used in Warping Model

The components of U2net model are described below. The model takes images as input for processing. These images typically undergo preprocessing steps, such as resizing and normalization, before being fed into the model. Then comes the parsing branch which is responsible for parsing or segmenting the input images into multiple regions or parts. This process involves assigning a class label to each pixel in the image, indicating which part of the human body it belongs to. The backbone of the model is likely a feature extractor, such as a convolutional neural network (CNN), which is used to extract meaningful features from the input images. Common backbones include VGG, ResNet, and

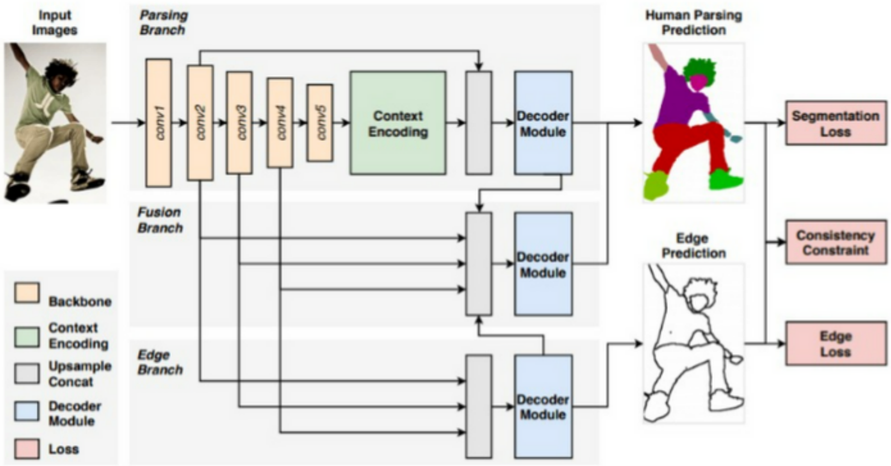


Figure 3.3: U2net Diagram

DenseNet. Context module is responsible for capturing and encoding contextual information from the input images. Contextual information can include the relationship between different parts of the human body, as well as the overall layout of the scene. There are 5 convolutional layers used for feature extraction. Convolutional layers apply a set of filters to the input image, producing a set of feature maps that capture different aspects of the image. Then comes the fusion module which combines features from different levels or sources to create a more comprehensive feature representation. Fusion can involve concatenating features from different layers, or applying attention mechanisms to selectively weight the importance of different features. Encoding branch encodes the features and contextual information into a more manageable format for the decoder. Encoding typically involves reducing the spatial resolution of the feature maps, while increasing the number of channels or features [12]. Edge module is then responsible for detecting and encoding edges or boundaries between different segments in the image. Edge detection can help improve the accuracy of the parsing result, particularly in regions where different parts of the human body are in close proximity. Upsample module increases the spatial resolution of the feature maps, often using techniques like transposed convolutions or unpooling. Upsampling is necessary to recover the original resolution of the input image during the decoding process. Concat operation combines the features from the encoder and upsampled features from the decoder. Concatenation allows the decoder to access both high-level semantic information and low-level spatial information during the decoding process. The decoder is then responsible for reconstructing the original image or generating the final parsing result from the encoded features. The decoder typically applies a series of transposed convolutional layers to upsample the feature maps and recover the original resolution of the input image. Human Parsing Prediction module generates the final parsing or segmentation result for the input images. The output of the model is a set of pixel-wise class labels, indicating which part of the human body each pixel belongs to. Segmentation Loss is the loss function used to train the model, which calculates the difference between the predicted parsing result and the ground truth. Common loss functions for semantic segmentation include cross-entropy loss and dice loss. Edge Prediction module predicts the edges or boundaries between different segments in the image. Edge prediction can help improve the accuracy of the parsing result, particularly in regions where different parts of the human body are in close proximity. Consistency Constraint ensures that the predicted edges are consistent with the overall parsing result.

3.3 Dlib and Haarcascade Architecture used in Accesory Try on

Haar cascade is a method employed for detecting faces within images. It operates by analyzing an image using Haar wavelet analysis, which breaks down the image into small square regions known as Haar-like features. These features represent various patterns of intensity changes, such as edges or textures. Through the application of convolutional kernels, Haar cascade identifies features like horizontal and vertical lines, which are indicative of facial features. These identified features undergo further processing to recognize patterns associated with faces. Central to the operation of Haar cascade is the cascade function, which consists of a series of classifiers trained using examples of both positive (images containing faces) and negative (images lacking faces) instances. Through repeated training iterations, the cascade function learns to differentiate between images with faces and those without. Ultimately, the cascade classifier combines multiple weak classifiers into a robust classifier, facilitating efficient and accurate face detection in images.

Figure 3.5 shows the flowchart which has been used for face and eye detection for glasses try on. Same thing can be used for ear detection which may be useful in earings try on.

Figure 3.6 shows the flowchart which has been used for neck and lip detection using dlib predictor which ultimately helps in necklace and lipstick try on. First it acquires

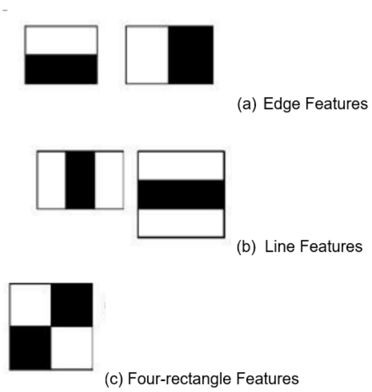


Figure 3.4: Haarcascade Features

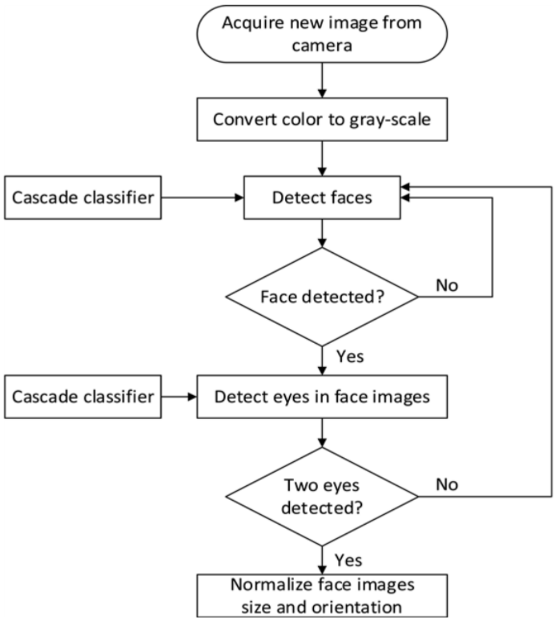


Figure 3.5: Flowchart for face and eye detection

the image and then convert it into gray scale. After which it uses the finetuned dlib face predictor to detect the face. After the face has been correctly detected it moves to dlib shape predictor. If the face has not been predicted properly it runs the dlib face predictor again. Then, the dlib shape predictor extracts the indices for neck and lip and finally draws a contour to represent the neck and lip. If it does not detect properly then we run it again. This ultimately helps in accessory try on such as necklace and lipstick.

3.4 System architecture for Hairwig Try on

The figure 3.7 describes the steps used in hairwig try on. Detailed explanation is described below. The image-to-image translation process in the diagram is a technique used to modify an image (source image) to resemble another image (target image) by accurately

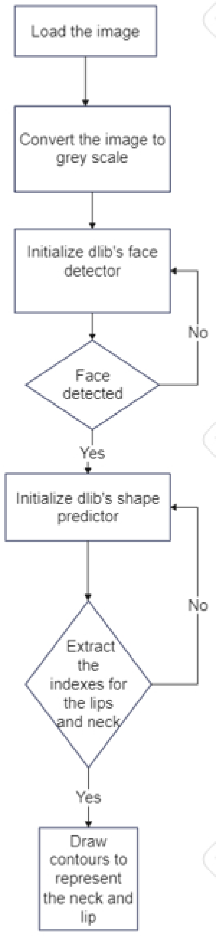


Figure 3.6: Flowchart for neck and lip detection

mapping specific features and structures between them. The primary goal is to maintain the integrity of the source image while including the textural and color information from the target image.

The process begins with image segmentation, which is the process of partitioning an image into multiple segments or regions. In this context, the focus is on segmenting faces and hair structures in both the target and source images. This step is crucial for accurately identifying the features that will be mapped between the images.

Once the images are segmented, the facial features and hair structures from the target image are mapped to the source image. This mapping process provides a structural guide for the image-to-image translation, ensuring that the essential features are accurately represented in the transformed source image.

The Fast Marching Method is a numerical algorithm which is used for solving boundary value problems and is then employed to efficiently compute the shortest path between

Figure 3.7: System diagram for hairwig try on

pixels or regions in the image [13]. This method is particularly helpful in the context of image-to-image translation, as it ensures a smooth transfer of facial features and hair structures from the target image to the source image.

Finally, the SDEdit method is applied to generate the transformed source image. It fills in the missing gaps. SDEdit is a powerful image-to-image translation technique that combines the structural information from the source image with the textural and color information from the target image.

Chapter 4

User Manual

4.1 Prerequisites

Before installing and using the AI virtual wardrobe system, users have to follow these prerequisites:

- A computer with a compatible web browser.
- Stable internet connection for seamless access to the virtual wardrobe platform.
- Basic understanding of how to navigate and interact with web applications.
- Visual Studio Code (VS Code) or any preferred code editor installed on your computer

4.2 Tools and Technology Stack

- Frontend Development: HTML, CSS, JavaScript, and React have been used to create an interactive and user-friendly interface
- Computer Vision Technologies: Advanced computer vision techniques, including image segmentation, landmark detection, and neural network models, have been used for accurate virtual try-on experiences.
- Deep Learning Frameworks: Frameworks like TensorFlow and PyTorch are utilized for training and fine-tuning neural network models for tasks such as face detection, landmark detection, and image inpainting.

- Image Processing Libraries: Libraries such as OpenCV are used for image manipulation, including cropping, resizing, and blending, to ensure realistic integration of virtual clothing and accessories onto user photos.
- Backend Infrastructure: A robust backend infrastructure,built using Python,is used to handle data storage, and communication with external APIs for fetching clothing and accessory data.

4.3 Installation & Usages

To set up the AI virtual wardrobe system locally, users need to follow these steps:

- Download the system's source code as a zip file from the provided source.
- Extract the zip archive to a directory on your computer.
- Open a terminal or command prompt and navigate to the extracted directory.
- Ensure you have Python installed, preferably Python 3.x.
- Install Flask if you haven't already by running the command: pip install Flask

Usage: Once the installation is complete, you can use the AI virtual wardrobe system locally by following these steps:

- Start the Flask server by running the command: flask run app.py
- This command will start the Flask application and serve the frontend on a local server.
- Open a web browser and navigate to http://localhost:5000 (or the specified port if different) to access the locally deployed virtual wardrobe platform.
- Explore the various features and options available, such as virtual try-on for clothing, accessories, and hairstyles.
- Upload a photo to visualize different items on yourself.
- Save and download your favorite looks and then share them to social media sites.

4.4 Features

- Users can upload the cloth image which they want to try on .
- Users can upload the person whose hairstyle they want to try on.
- Users can choose from a variety of lipstick colors and select whichever they want to try on
- Users can browse and select from a variety of accessories such as necklaces, glasses, and earrings to try on virtually
- Users are able to simultaneously experiment with multiple elements, such as changing lipstick colors while also applying necklaces
- Users have the option to download and save images of their virtual try-on sessions, allowing them to keep a record of their favorite looks for future reference
- Users can easily share their virtual try-on sessions and styled looks on various social media platforms such as Facebook, Instagram, Twitter, and Pinterest.

Chapter 5

Results and Discussion

5.1 Results

This section describes the results of our efforts to build and deploy the AI Virtual Wardrobe. The project's major goal was to develop a user-friendly platform that allows people to visualise different clothing goods, improving their shopping experience and allowing them to make informed decisions. Computer vision techniques and algorithms were used to accomplish precise garment detection, seamless garment adaption, and realistic virtual try-on simulations. Here, the results of our system are demonstrated.

5.1.1 Virtual Cloth Try On

The outcomes of the virtual clothing try-on, depicted in Figure 5.1, were obtained using the ACGPN architecture. ACGPN, a derivative of the Generative Adversarial Network (GAN), is specialized in regulating attributes within the generated content. This GAN framework is based on the Pix2Pix model, which is specifically designed for high-resolution image-to-image translation tasks.

Figure 5.1: Cloth Try On Output

5.1.2 Virtual Necklace Try On

Figure 5.2 depicts the results of the virtual necklace try-on. A fine-tuned Dlib shape predictor is used to identify facial landmarks accurately. These detected facial landmarks serve as the foundation for identifying precise cropping points for the necklace image, which is then resized and positioned based on the landmarks and any necessary changes. The overlay of the necklace image is achieved using masking and bitwise operations.

Figure 5.2: Necklace Try On Output

5.1.3 Virtual Earring Try On

The results of virtual earring try-on are shown in Figure 5.3. HaarCascade face classifiers trained on face images are used to identify the user's face. The earring images are then adjusted to fit the dimensions of the face, and the earring position is established appropriately. These scaled earring images are then used to overlay the input image with transparency masks.

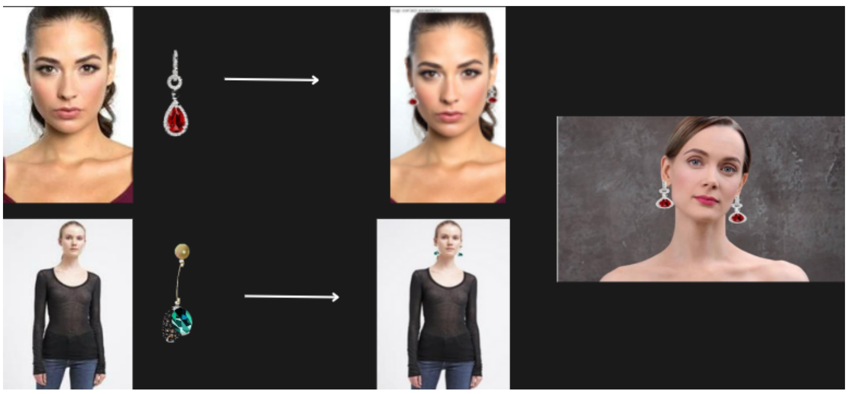


Figure 5.3: Earring Try On Output

5.1.4 Virtual Glass Try On

Figure 5.4 shows the results of the virtual glasses try on. The method uses a two finetuned HaarCascade classifier to identify the face, and another to detect the eyes inside the observed facial region. The identified eyes are used to determine regions of interest (ROIs). The glasses are modified to match the width of the detected eye and then overlaid onto the person's eyes.

Figure 5.4: Glasses Try On Output

5.1.5 Virtual Lipstick Try On

The results of the virtual lipstick try on are depicted in Figure 5.5. The method uses a pre-trained face detector and shape predictor from the dlib package to recognise the person's lips. Once the lips have been detected, the chosen lipstick colour is applied by filling the area defined by the lip landmarks. Gaussian Blur is used to smooth down the appearance of the coloured lips and make them merge better with the source image.

5.1.6 Virtual Hairwig Try On

The outcomes of the virtual hairwig try on are displayed in Figure 5.6. UNet architecture was used to divide the face into five sections: hair, face, ears, neck, and shirt. Following segmentation, facial landmark detection is conducted on the segmented image to generate

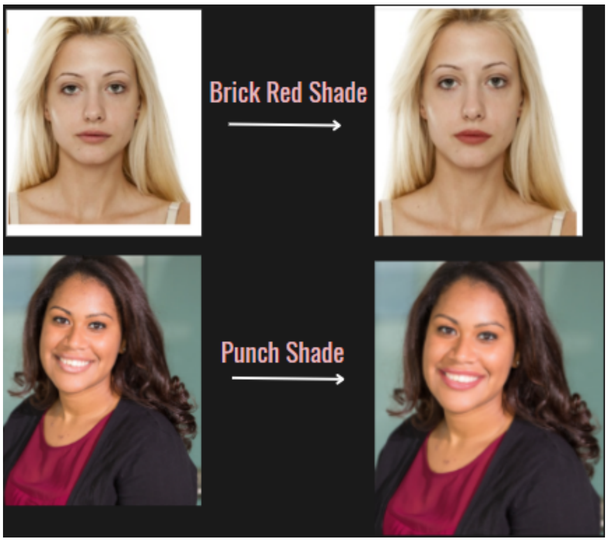


Figure 5.5: Lipstick Try On Output

semantic maps, which are then used to crop the face and hair. The source hair and target face results are overlapped, and a difference mask is created to fill in the missing sections of the face and hair. The SDEdit method is used for overlapping and creating a difference mask. To achieve a smooth transition, the image's missing sections are filled by the fast marching approach.

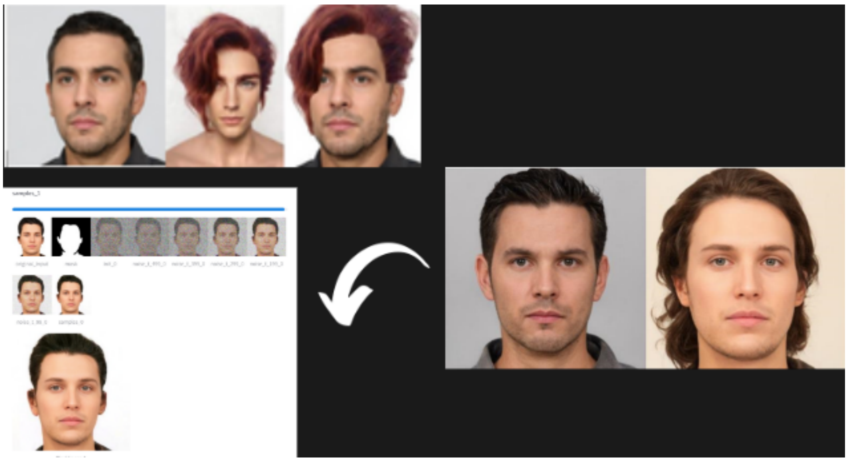


Figure 5.6: Hairwig Try On Output

5.1.7 Flask Web App

All virtual try-on systems were tested using Google Colab before being incorporated into a unified Flask application. This Flask application provides a unified, user-friendly platform for accessing all functions independently. The interface is simple: customers upload an image on the first page and then select their preferred try-on choice. Depending on their choice, users then upload the appropriate accessory. After rendering, the finished result is displayed, along with options for sharing and downloading it. The home page of the web application is shown in figure 5.7.

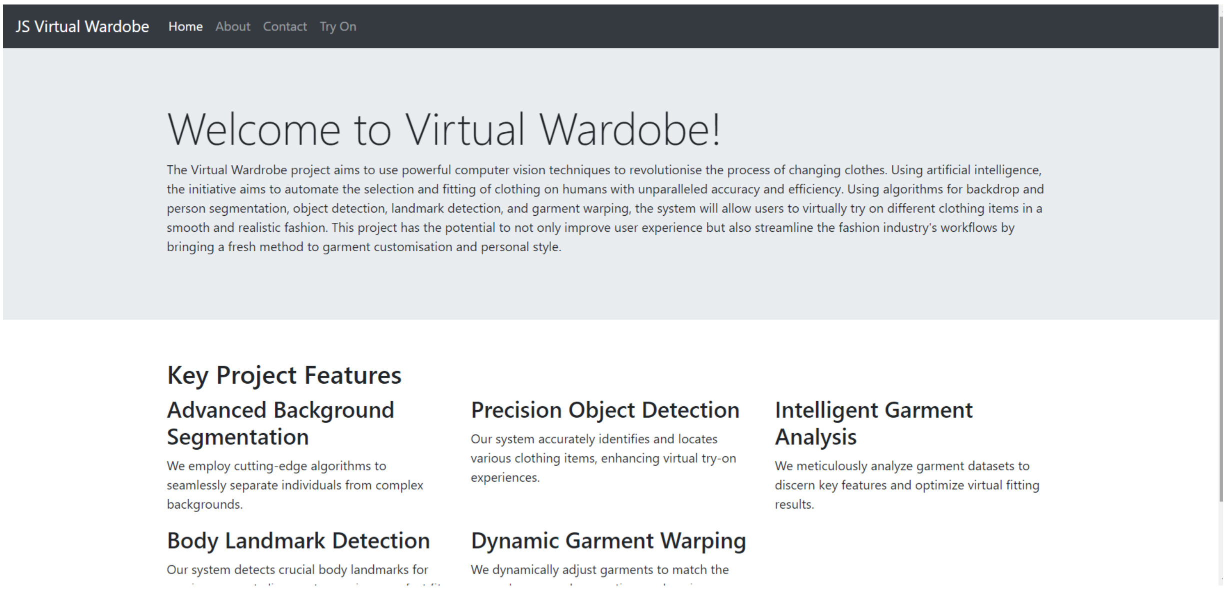


Figure 5.7: Home Page of Application

The Try On Page is shown below in figure 5.8. The user can select desired option for try on and they will be redirected to that try on page.

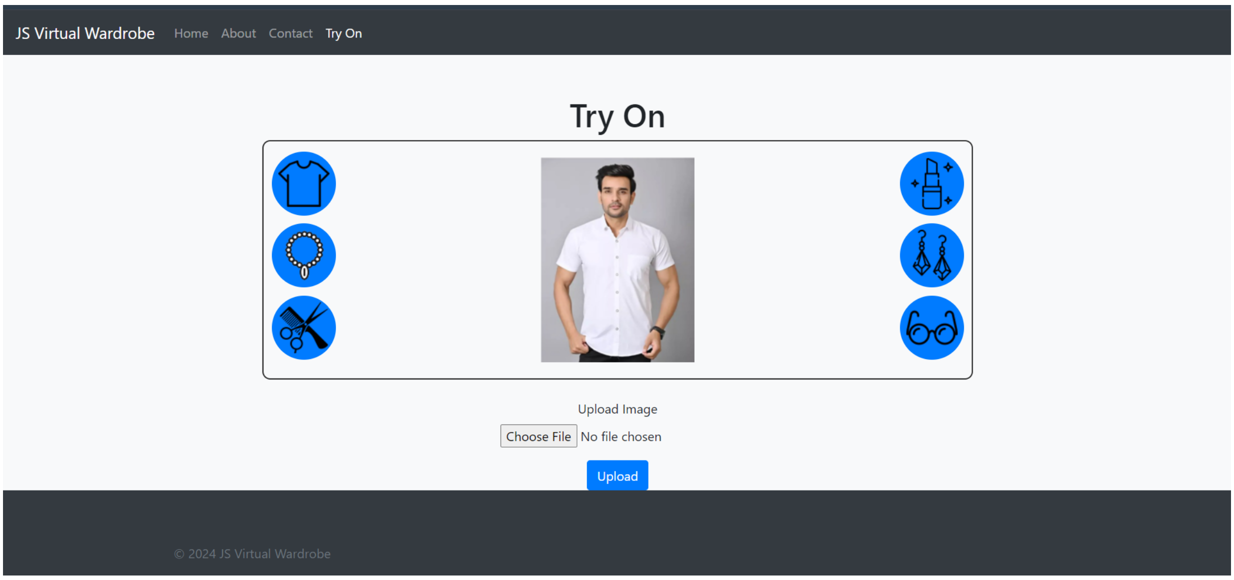
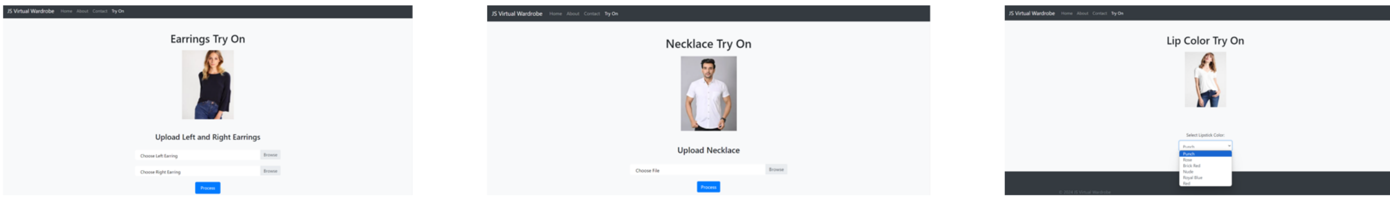


Figure 5.8: Try On Page of Application

The steps to use the web application are demonstrated in figure 5.9. Briefly, the steps are as follows:

- Upload your own image to try on accessories or test the functionality using a pre-defined model image.
- Depending on the selected try-on feature, upload the necessary accessory or choose from the available options.
- View the final result. Download it using the provided button or share it on social media.

Step 1: Upload the image and choose the desired option



Step 2: Upload the Accessory

Step 3: View the result

Figure 5.9: Steps to use the website

5.2 Discussion

The project results provide vital insights into the virtual try-on system's effectiveness in improving consumers' online purchasing experiences. The application has the potential to greatly increase user engagement and satisfaction levels. Users who utilise the virtual try-on tool are more likely to spend more time exploring product options and show greater interest than those who do not use it, so improving the shopping experience and increasing user engagement. Extensive research has consistently shown that virtual try-on features are associated with higher conversion rates, lower product returns, and increased consumer loyalty [14]. Our system is consistent with the ideas of the Technology Acceptance Model (TAM), which states that perceived usefulness and simplicity of use are key factors of consumer acceptance of technology [15]. In practical terms, these findings have significant implications for e-commerce companies looking to improve their online buying platforms. Businesses that incorporate virtual try-on capabilities can differentiate their goods, attract a larger audience, and eventually boost sales figures. However, it is critical to recognise that, despite the positive outcomes, the system is not without flaws. Improvements could be made by taking into account other elements influencing the online retail landscape. Although the results are commendable, additional refinement could be done by using more advanced hardware and larger datasets.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

The internship provided an invaluable opportunity to delve into the realm of software development and its diverse stages, offering exposure to a wide array of technologies and tools. Throughout the 19-week duration, a comprehensive exploration of numerous tools and technologies significantly enhanced overall proficiency and coding standards. The initial phase of training included learning the fundamentals of Python, VS Code, Git, Flask, NLP and Deep learning techniques.

Following the training period, each intern was entrusted with an individual project, overseen by the tech lead and super standup lead. This meant developing an entire application from the start, utilizing the knowledge and skills acquired during training. Engaging in this project afforded us a hands-on understanding of the complete software development lifecycle—from conceptualization and research to design, prototyping, and eventual implementation and finally testing.

6.2 Future Work

- Integration Extension Development: Expanding on the current functionality, the next step involves creating a browser extension to integrate the try-on feature across different shopping websites such as Myntra, Flipkart or Amazon. This extension aims to simplify the user experience, allowing them to visualize how clothing items look on them without leaving their preferred online stores. By leveraging common web technologies like browser extensions and APIs, this integration extension will enhance user convenience and accessibility across various e-commerce platforms. This way it may also help to earn some money through this. So, to acheive this first we would need to deploy our flask webiste.

- Size Recommender Feature: With a focus on developing personalized recommendations, there's a need to have the size recommender feature. The goal is to develop a system that can suggest and automatically select the perfect fit for users based on their individual body measurements and preferences. By utilizing machine learning and data analysis techniques, this enhancement will not only enhance user satisfaction but also reduce returns and exchanges. Moreover, after suggesting the perfect fit it should allow the user to select that size of shirt and visualise on him.

- Using advance hardware: To improve the accuracy as well as reduce the startup time of the try-on system, upgrading hardware components, especially the GPU (Graphics Processing Unit), is essential. By investing in high-performance servers with advanced GPUs, the system's processing capabilities can be significantly improved. This upgrade will enable faster image processing and rendering, resulting in quick and realistic try-on experiences for users.

Bibliography

[1] S. Lee, S. Lee, and J. Lee, "Towards detailed characteristic-preserving virtual try-on," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2236–2240, 2022.

[2] "F(x) data labs." https://yourstory.com/companies/fx-data-labs. Accessed: May 8, 2024.

[3] "F(x) is ai." https://fxis.ai/. Accessed: May 8, 2024.

[4] "F(x) is ai." https://themanifest.com/company/fx-ai. Accessed: May 8, 2024.

[5] "The original ai company." https://clutch.co/profile/fx-ai#highlights. Accessed: May 8, 2024.

[6] X. Han, Z. Wu, Z. Wu, R. Yu, and L. S. Davis, "Viton: An image-based virtual try-on network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7543–7552, 2018.

[7] K. Li, M. J. Chong, J. Liu, and D. Forsyth, "Toward accurate and realistic virtual try-on through shape matching and multiple warps," *arXiv preprint arXiv:2003.10817*, 2020.

[8] D. Song, X. Zhang, J. Zhou, W. Nie, R. Tong, and A.-A. Liu, "Image-based virtual try-on: A survey," *arXiv preprint arXiv:2311.04811*, 2023.

[9] Y. R. Cui, Q. Liu, C. Y. Gao, and Z. Su, "Fashiongan: display your fashion design using conditional generative adversarial nets," in *Computer Graphics Forum*, vol. 37, pp. 109–119, Wiley Online Library, 2018.

[10] H. Yang, R. Zhang, X. Guo, W. Liu, W. Zuo, and P. Luo, "Towards photo-realistic virtual try-on by adaptively generating preserving image content," *arXiv preprint arXiv:2003.05863*, 2020.

[11] G. Liu, D. Song, R. Tong, and M. Tang, "Toward realistic virtual try-on through landmark guided shape matching," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, pp. 2118–2126, 2021.

[12] H. Zhang, K. Dana, J. Shi, Z. Zhang, X. Wang, A. Tyagi, and A. Agrawal, "Context encoding for semantic segmentation," in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pp. 7151–7160, 2018.

[13] J. A. Sethian, "Fast marching methods," *SIAM review*, vol. 41, no. 2, pp. 199–235, 1999.

[14] H. Hwangbo, E. H. Kim, S.-H. Lee, and Y. J. Jang, "Effects of 3d virtual "try-on" on online sales and customers' purchasing experiences," *IEEE Access*, vol. 8, pp. 189479–189489, 2020.

[15] S. Zaineldeen, L. Hongbo, A. L. Koffi, and B. M. A. Hassan, "Technology acceptance model'concepts, contribution, limitation, and adoption in education," *Universal Journal of Educational Research*, vol. 8, no. 11, pp. 5061–5071, 2020.