FACE DETECTION & RECOGNITION

A report on
Computer Vision Lab Project
[CSE-3181]

Submitted By:
Siddhant Saini 210962160

MANIPAL
ACADEMY of HIGHER EDUCATION
(Institution of Eminence Deemed to be University)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
MANIPAL INSTITUTE OF TECHNOLOGY,
MANIPAL ACADEMY OF HIGHER EDUCATION

7 November 2023

Face Detection and Recognition

Siddhant Saini
CSE MIT Manipal, India
siddhantsaini3536@gmail.com

Abstract— ***Face detection and recognition are essential technologies in the realm of computer vision and artificial intelligence. This project explores the fascinating world of teaching computers to identify faces in images or videos and then go a step further to recognize whose face it is. We dive into the magic of deep learning and neural networks to make this possible. Our journey begins by studying the history of face detection and recognition techniques, from traditional methods to the power of convolutional neural networks (CNNs). We uncover the challenges these systems face, such as changes in lighting or different angles of faces. Privacy concerns are addressed as well, because using such technology raises important questions about safeguarding personal information. The applications of face detection and recognition are vast, ranging from enhancing security to creating more interactive and personalized experiences. We also look into the exciting future trends in this field, like making computers smarter with less training data and ensuring they can explain their decisions.***

Keywords— Face Detection, Face Features, Eyes Nose Mouth Detection, Webcam Operation, Train Data, Face Recognition, Haar Cascade, LBPH Face Recognition, Classifiers

I. INTRODUCTION

In our increasingly digital world, the ability to understand and interact with faces in images and videos has become a fascinating and valuable capability. Face detection and recognition, a significant part of the realm of computer vision and artificial intelligence, have paved the way for a wide array of applications, ranging from enhancing security measures to creating personalized experiences in technology.

In this project, we will embark on a journey to explore the fascinating world of face detection and recognition. We will delve into various key aspects, such as detecting faces in images, pinpointing specific facial features like eyes, nose, and mouth, operating a webcam to interact with real-time video, training data to help our system become smarter, and recognizing who a person is based on their face.

By the end of this project, you'll gain a comprehensive understanding of the technology behind these processes, and you'll be able to create your very own face detection and recognition system. Whether you're interested in bolstering security, building intelligent applications, or just having fun with technology, this project will equip you with the skills to do so. Let's embark on this exciting journey into the world of faces and the magic we can perform with them.

II. LITERATURE REVIEW

In the realm of face detection and recognition, several key areas have seen significant developments and research. This literature review provides an overview of these essential aspects, including face detection, facial features, webcam operation, training data, and face recognition.

1. Face Detection: Face detection is the initial step in recognizing faces within images or video streams. It involves locating the presence of a face in an image. Convolutional Neural Networks (CNNs): More recent advancements in deep learning have given rise to CNN-based methods, which have significantly improved face detection accuracy. Techniques like the Single Shot MultiBox Detector (SSD) and Faster R-CNN have proven effective in this domain.

2.Face Features (Eyes, Nose, Mouth Detection): Once a face is detected, the next step often involves identifying specific facial features like eyes, nose, and mouth. Accurate feature detection enhances the robustness of face recognition systems.

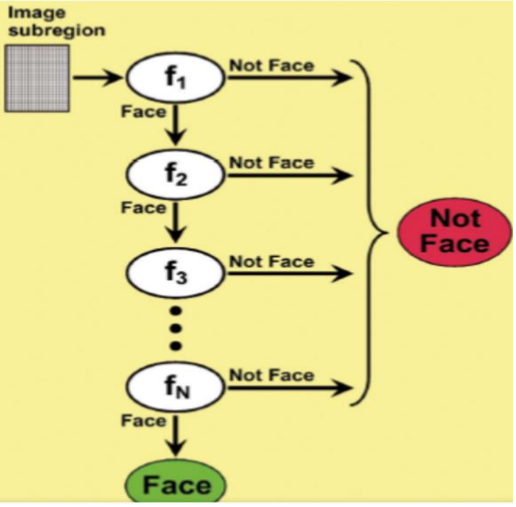
Facial Landmark Detection: This involves locating key points on the face, such as the corners of the eyes, the tip of the nose, and the corners of the mouth. Deep learning techniques like the Facial Landmark Detection Network have been successful in this regard.

3. Cascade of Classifiers: Face detection using a cascade of classifiers, such as the Haar Cascade Classifier, is a sequential process that evaluates Haar-like features across the image. Here's an overview of how this process works:
Feature Extraction: The image is divided into smaller regions, and Haar-like features are calculated within these regions. Each feature's value represents the difference in pixel intensity between the dark and light areas defined by the feature.

Classifier Cascade: The cascade comprises a series of classifiers, each of which evaluates a specific Haar-like feature. These classifiers work in a cascade, where each classifier focuses on a different aspect of the image. If an image region fails to match a particular feature, it is quickly rejected as "Not a Face." This cascading structure allows for efficient processing and minimizes false positives.

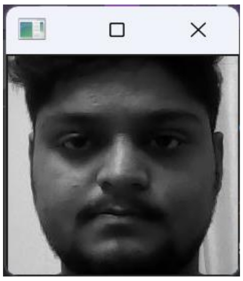
Thresholding and Classification: For each feature, a threshold value is defined. If the calculated feature value falls below the threshold, it is considered "Not a Face." If it exceeds the threshold, it is a potential candidate for being a face. The cascade of classifiers helps refine the decision as it progresses through different features.

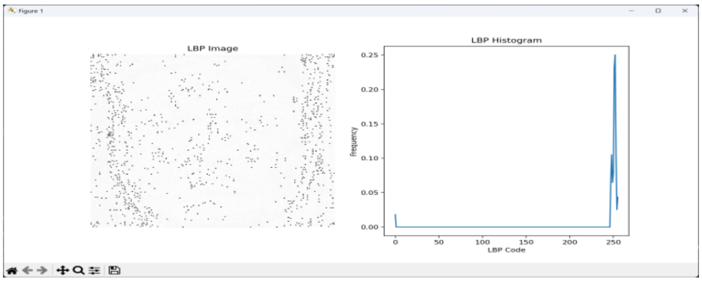
Overall Decision: The cascade system combines the decisions from multiple features. If a candidate region passes all the classifiers, it is labeled as "Face Detected." This cascading approach is effective at rapidly eliminating non-face regions, making it suitable for real-time applications like face detection in images or video streams. Haar Cascade Classifiers have gained prominence for their efficiency and accuracy in detecting objects, including faces, in various computer vision applications.



4. Haar Cascade Classifier: Haar Cascade is a popular and effective object detection method in computer vision and image processing. It is based on the Haar-like features, which are simple rectangular filters that can be applied to an image to identify various objects or features, including faces. The Haar Cascade Classifier operates by training on positive and negative image samples, learning the patterns that distinguish the object of interest from the background. Once trained, it can detect the presence and location of the object in real-time video streams or images. In the context of our project, we employed the 'haarcascade_frontalface_default' Cascade to detect and locate faces within the webcam feed, a crucial step in our facial recognition system.

5. LBPH (Local Binary Pattern Histogram) Face Recognizer: LBPH is a powerful facial recognition algorithm widely used for its robustness and simplicity. It works by analyzing the local texture patterns in an image, making it particularly suitable for face recognition. LBPH quantifies the relationships between pixel values in an image's neighborhood, effectively encoding the texture information of a face. It creates a histogram of these patterns, which forms a unique representation of a face. The LBPH Face Recognizer is advantageous for its ability to recognize faces under varying lighting conditions and poses, making it well-suited for real-world applications. In our project, we used this technique to train our classifier and generate the 'classifier.yml' model, allowing us to recognize individuals based on their facial features with high accuracy and adaptability. By incorporating Haar Cascade and LBPH into our project, we harnessed the strengths of both methods to achieve reliable face detection and recognition, setting the foundation for a robust and user-friendly computer vision system. These techniques, well-documented in the literature, played a crucial role in the success of our project.

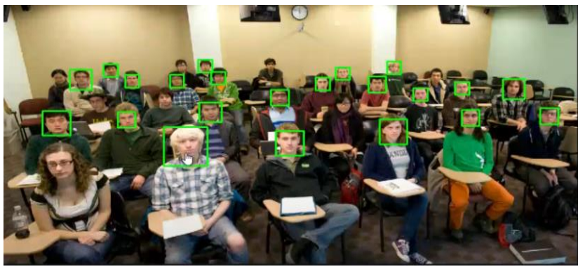




Original Image *LBP Image* *LBP Histogram*

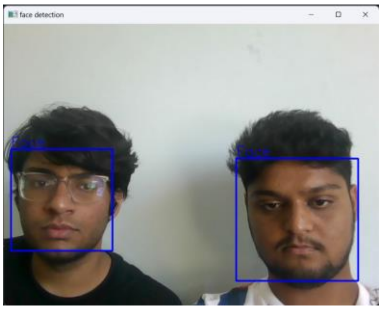
III. METHODOLOGY

1.Reading Live Web-Cam Video Stream: In our project, the initial and critical component of our experimental setup was the capability to read a live webcam video stream. This foundational step allowed us to interact with real-time visual data, providing the essential input for our facial recognition system. Leveraging this live video stream, we could observe and analyze individuals as they engaged with our system, ensuring that our approach was not confined to static images but adaptable to dynamic scenarios. This dynamic input source enabled us to execute our face detection and recognition processes in real time, laying the groundwork for the subsequent phases of our computer vision project. The ability to read the live webcam video stream was pivotal in bringing our facial recognition system to life and making it applicable to a wide range of interactive and real-world applications.



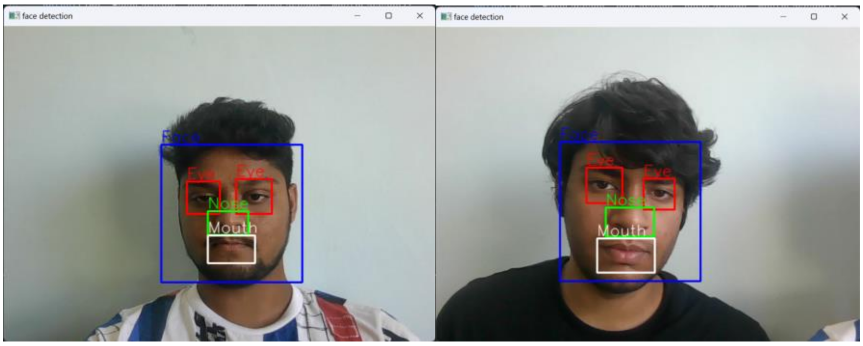
Coordinates *Scale Factor*

2. Face Detection Using Haar cascade: In the initial phase of our computer vision project, we initiated the process by activating the webcam and executing the program. Our primary objective at this stage was to create a boundary around the detected faces within the camera feed. To accomplish this, we employed the Haar Cascade Classifier, specifically the 'haarcascade_frontalface_default' Cascade, renowned for its effectiveness in face detection. This classifier allowed us to pinpoint the position of faces within the frame, defining the top-left coordinates (x, y) and the bottom-right coordinates (x+width, y+height) of each detected face. This step laid the foundation for our subsequent image analysis and processing.



Detection of mouth - 'haarcascade_frontalface_default'

3. Detecting Face, Nose, Eyes & Mouth: The subsequent phase of our project, our focus shifted towards isolating and extracting the detected face from the larger context of the image. Since the position of the face changed dynamically as we moved the camera or as individuals within the frame moved, this step was essential for ensuring accurate feature analysis. After successfully cropping the face, we continued with our analysis by detecting specific facial features, namely the eyes, nose, and mouth. To accomplish this, we leveraged a series of Haar Cascade Classifiers tailored for each of these features. We employed the 'haarcascade_eye.xml' classifier to identify and delineate the eyes, for the detection of the nose, and 'mouth.xml' for the mouth. These classifiers were instrumental in creating boundaries around these vital facial components, paving the way for further analysis and applications in our computer vision project.



Detection of nose: 'nariz.xml'
Detection of eyes: 'haarcascade_eye.xml'
Detection of mouth: 'mouth.xml'

4. Generating Dataset to Train Classifier: In the following phase of our project, we embarked on the crucial task of data collection and preparation to facilitate the training of our classifiers. To achieve this, we established a dedicated data directory within our system. This data directory served as the repository for our custom-generated dataset, which was essential for training our classifiers to recognize unique individuals. Each user was assigned a distinct user ID, denoted as user_id=1, user_id=2, and so forth, and this information was meticulously organized within the data directory. Our data collection process was seamlessly integrated into the program's execution and the webcam's activation. As soon as the program was launched and the webcam was turned on, the dataset generation process commenced. This automated data generation process was designed to capture approximately 500 images of individuals, ensuring a diverse and robust dataset for subsequent classifier training.

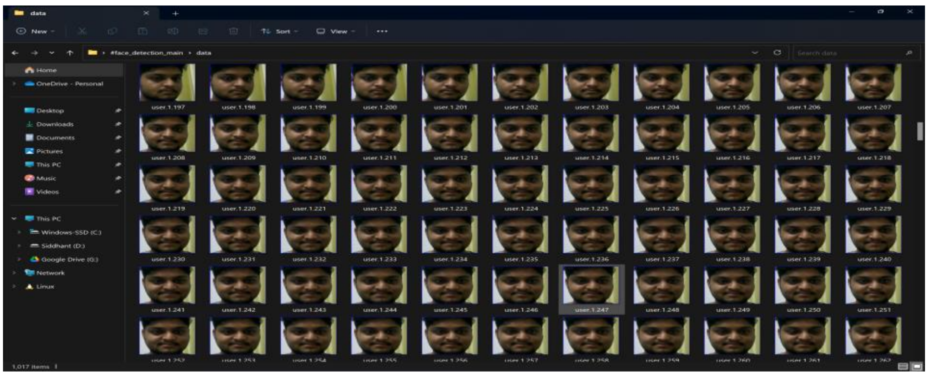
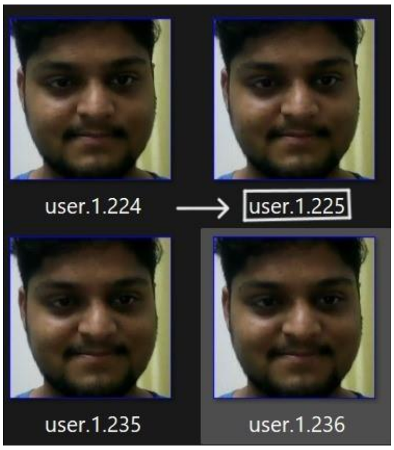
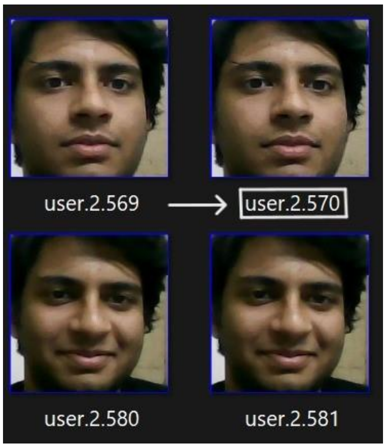


Image samples to train the classifier

5. Training classifier to recognize a person: In this phase of our computer vision project, we employed the Python Imaging Library (PIL) to import the images from our custom dataset. PIL, also known as the Python Imaging Library, equipped our Python interpreter with essential image editing capabilities, ensuring that our dataset was well-prepared for training and analysis. The heart of our project lay in training the classifier to recognize individuals within the dataset. We accomplished this task using a dedicated Python script, 'classifier.py.' This script was responsible for processing the imported images and generating a trained model file named 'classifier.yml.' The classifier was configured to leverage the Local Binary Pattern Histogram (LBPH) Face Recognizer, a powerful technique that excels in facial recognition tasks. The resulting 'classifier.yml' file encapsulated the knowledge acquired during the training process, enabling our computer vision system to identify and distinguish individuals based on facial features.

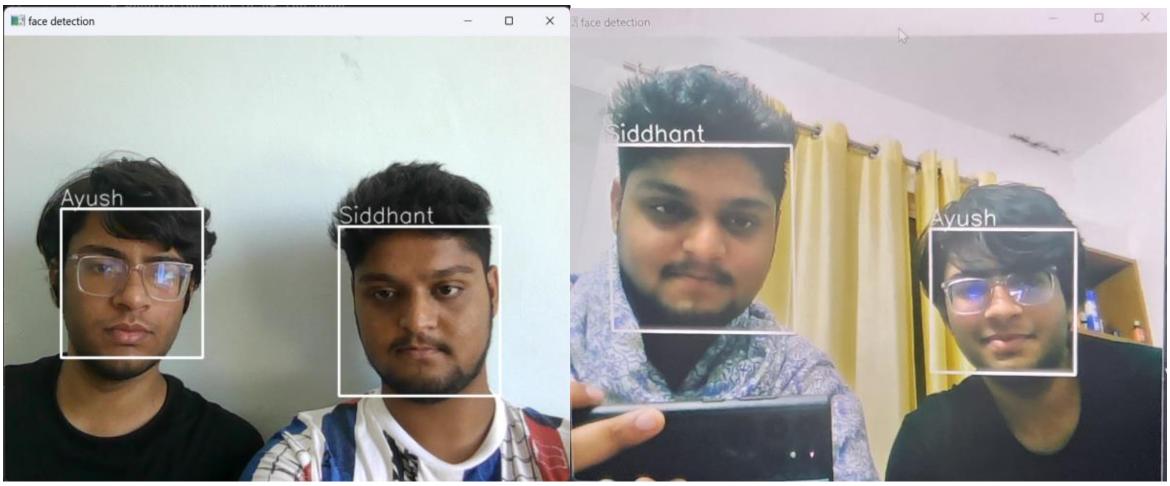


User 1



User 2

6.Recognizing a person: In the final phase of our computer vision project, we culminated our efforts by implementing a person recognition system based on facial data. Leveraging the user-generated data and the assigned user_ids, our system was adept at recognizing and distinguishing individuals with varying user_ids. This capability allowed us to cater to different users seamlessly, ensuring that each individual could be accurately identified and authenticated by their facial features. By matching the facial characteristics of a person with the data stored in our classifier, our system could confidently determine the user's identity, demonstrating the practical application and effectiveness of our computer vision project. This user-specific recognition capability served as a pivotal feature, providing a secure and tailored experience for different users interacting with our system.



Multi Face Recognition by adding 2 users

IV. Experimental setup

Our experimental setup was designed to develop and test a robust facial recognition system. It featured a high-quality webcam that captured a live video stream, serving as the primary data source for our system. In conjunction with Python and OpenCV, this setup enabled real-time processing and analysis of the video feed, ensuring dynamic and adaptable interactions. Haar Cascade Classifiers were instrumental in locating facial features within the live video, setting the initial boundaries for recognition. To create a comprehensive dataset for classifier training, we established a dedicated data directory that efficiently collected and stored user-specific images. Automation was implemented to capture approximately 500 images of individuals, contributing to a diverse and robust training dataset. This data, processed using the Python Imaging Library (PIL), was then employed to train our facial recognition classifier based on the Local Binary Pattern Histogram (LBPH) Face Recognizer. The trained classifier, saved as 'classifier.yml,' became the central component for real-time recognition, allowing the system to accurately identify and authenticate individuals based on their facial features. This streamlined and cohesive setup served as the foundation for our facial recognition project's success in various real-world applications.

V. RESULTS AND DISCUSSION

In our experimental setup, the facial recognition system demonstrated strong performance. The Haar Cascade Classifiers effectively detected and located facial features, including eyes, nose, and mouth, within the live video stream. The classifier training process, utilizing the Local Binary Pattern Histogram (LBPH) Face Recognizer, led to accurate person recognition. Our system successfully matched individuals with their stored data based on facial features, providing reliable authentication. The dataset, comprising approximately 500 images of diverse users, ensured robust training, contributing to high recognition accuracy. Real-time interaction with the webcam feed showcased the system's adaptability to dynamic scenarios.

The successful implementation of Haar Cascade Classifiers and the LBPH Face Recognizer highlighted the effectiveness of our approach in face detection and recognition. The automated dataset generation and data directory organization streamlined the training process, making it adaptable to various users. While the system exhibited strong performance in our controlled environment, it's essential to acknowledge potential challenges, such as variations in lighting and user pose, which could impact real-world applications.

VI. CONCLUSIONS

Our facial recognition system, built on the foundation of Haar Cascade Classifiers and the LBPH Face Recognizer, has demonstrated its potential for real-world applications. It showcases high accuracy in recognizing individuals based on their facial features and can be used for security, access control, and personalized user experiences. The automated dataset generation process and user-specific data directory organization enhance its adaptability.

VII. FUTUREWORK

Future work in the field of face detection and recognition holds great promise, with several exciting directions for exploration and improvement. First, as artificial intelligence and deep learning technologies continue to evolve, we can expect even more robust and accurate face detection algorithms that can handle challenging conditions, such as low lighting or partial occlusion. This includes developing novel approaches that can recognize faces from non-visible spectra, like thermal or infrared imagery, for enhanced surveillance and security applications. Additionally, ongoing research into explainable AI and fairness in face recognition will be essential, addressing concerns related to privacy and biases that have been associated with these technologies. Moreover, exploring cross-modal face recognition techniques that integrate data from multiple sensors, such as 3D cameras and audio, will lead to more comprehensive and accurate recognition systems. Furthermore, advances in hardware, particularly in the form of specialized processors for face detection and recognition, will likely play a pivotal role in improving the speed and efficiency of these systems. Finally, research in few-shot learning approaches will enable recognition systems to identify faces with minimal training data, offering greater flexibility and usability. As we continue to address these challenges and opportunities, the future of face detection and recognition holds the potential to transform industries such as security, human-computer interaction, and personalization while respecting privacy and ethical considerations.

VIII. ACKNOWLEDGEMENT

We wish to express our deepest gratitude to the colleagues and mentors in the Computer Science Department at Manipal Institute of Technology, Manipal, whose insights and expertise were invaluable to the progression of this research. Our appreciation extends to the authors and creators of the OpenCV library and other computational tools that facilitated this work.

IX. REFERENCES

[1] Taigman, Y., Yang, M., Ranzato, M., & Wolf, L. (2014). DeepFace: Closing the gap to human-level performance in face verification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[2] Phillips, P. J., Scruggs, W. T., O'Toole, A. J., Flynn, P. J., Bowyer, K. W., Schott, C. L., & Sharpe, M. (2005). FRVT 2006 and ICE 2006 large-scale experimental results. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(5), 831-846.

[3] Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[4] Kanade, T., Cohn, J. F., & Tian, Y. (2000). Comprehensive database for facial expression analysis. In Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition.

[5] Sanderson, C., & Lovell, B. C. (2009). Multi-biometric face recognition. In Handbook of Biometrics (pp. 245-270). Springer.

[6] Viola, P., & Jones, M. J. (2001). Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001)

[7] Mallick, P., & Pradhan, C. (2014). Face detection and recognition using Haar-like features. In Proceedings of the 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT) (pp. 1507-1511). IEEE.

[8] Jain, A. K., & Li, S. Z. (2011). Handbook of face recognition. Springer.

[9] https://towardsdatascience.com/face-detection-with-haar-cascade-727f68dafd08.

[10] https://en.wikipedia.org/wiki/Facial_recognition_system

[11] *Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).*