MANIPAL INSTITUTE OF TECHNOLOGY
MANIPAL
A Constituent Institution of Manipal University

Department of Computer Science and Engineering

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

Face Detection and Emotion Recognition

Submitted By
D TARUN KUMAR REDDY
Reg no. 210962082

AKULA KARTHIK
Reg no. 210962134

Fifth Semester

[Jul – Nov 2023]

Department of Computer Science and Engineering
Manipal Institute of Technology, Manipal.
November 2023

Face Detection and Emotion Recognition

1st Tarun Kumar Reddy D
Dept of Computer Science Engineering with AI & ML.
Manipal Institute of Technology
Manipal.
tarunreddydanduga@gmail.com

2nd Karthik Akula
Dept of Computer Science Engineering with AI & ML.
Manipal Institute of Technology
Manipal.
advk1234mkarthik@gmail.com

***Abstract*— This project is dedicated to the development of a real-time system for face detection and emotion recognition. Face detection, a fundamental component, involves the precise identification and localization of human faces within images and video streams. This technology serves as the cornerstone for various applications, including security systems, access control, and user experience enhancement. Emotion recognition complements face detection by categorizing emotional states based on facial expressions. The methods utilized encompass HOG, deep learning using Convolutional Neural Networks (CNN) and make use of computer vision libraries such as OpenCV, dlib, and face_recognition. The practical applications of this technology are far-reaching, from bolstering security and surveillance systems to creating personalized user experiences in human-computer interaction scenarios.**

Keywords—Face Detection, HOG, Emotion Detection, OpenCV-python, CNN.

I. INTRODUCTION (*HEADING 1*)

This research project is centred on the development of an efficient facial detection and recognition system for both images and videos, empowered by two pivotal methodologies: Convolutional Neural Networks (CNN) and the Histogram of Oriented Gradients (HOG) method.

Convolutional Neural Networks (CNN): CNNs are deep learning models that have revolutionized computer vision tasks, especially in image recognition. These neural networks are inspired by the human visual system and are adept at learning and recognizing patterns in images. They employ layers of convolutional operations to extract hierarchical features, enabling the model to identify complex structures and patterns in the data. In our project, CNNs play a central role in facial recognition, enabling the system to learn and differentiate facial features, expressions, and characteristics.

Histogram of Oriented Gradients (HOG): HOG is a feature descriptor used for object detection, especially in the case of pedestrian detection and facial recognition. HOG operates by dividing an image into small cells and calculating the gradient orientation in each cell. These gradient orientations are then used to construct histograms that capture the distribution of edge directions. HOG is instrumental in encoding the shape and structure of objects, making it highly relevant in the precise localization and delineation of facial features. In this project, HOG is an integral component, contributing to accurate facial detection and recognition.

The project aims to address the fundamental goals of any facial detection system, including the identification of faces in images and the precise localization and delineation of those faces. The complexity arises from various factors, such as background variations, lighting conditions, poses, and individual attributes.

The contemporary landscape emphasizes the need for autonomous systems capable of monitoring critical areas. In this research article, we introduce a facial recognition model that utilizes videos and images for training, which can serve as a key component in a decision support system for classifying images and video content.

The primary objective is to introduce a real-time facial detection and recognition system that is highly reliable and cost-effective. This system is designed to integrate into a variety of applications seamlessly. Our utilization of deep learning techniques, including CNN, in tandem with the HOG method, for facial recognition ensures that the system can accurately identify and classify facial features. We leverage readily available components and libraries, such as Dlib, the Face Recognition Library, and the Open Source Computer Vision Library (OpenCV), and explore various facial recognition machine learning algorithms.

Our results showcase the system's capacity to deliver superior performance in real-time applications, even under the constraints of limited CPU and GPU processing power. In summary, our work aims to contribute to the development of accessible and robust facial recognition technology applicable across a wide range of domains.

II. Literature Review

Image-based Face Detection and Recognition" presents a comprehensive review of image-based face detection and recognition techniques, providing a thorough analysis of the state-of-the-art methods in the field. The authors discuss the evolution of facial recognition technology, emphasizing its significance in various applications, from security systems to social media platforms. They delve into the fundamental components of face detection, such as Viola-Jones, deep learning-based methods, and their variants, exploring their strengths and limitations. The paper also highlights advances in face recognition algorithms, including Eigenfaces, Fisherfaces, and modern deep neural networks. Additionally, it delves into the challenges posed by factors like occlusion and illumination, as well as the ethical concerns associated with face recognition technology, offering a comprehensive overview of the current landscape of image-based face detection and recognition research.

"Face Detection & Recognition from Images & Videos Based on CNN & Raspberry Pi" presents an innovative study at the intersection of computer vision and embedded systems, focusing on face detection and recognition using Convolutional Neural Networks (CNNs) in conjunction with the Raspberry Pi platform. The authors delve into the contemporary landscape of deep learning in computer vision, emphasizing the prominence of CNNs in handling complex tasks such as face detection and recognition. They discuss the portability and energy-efficient aspects of Raspberry Pi, making it an appealing choice for real-time applications. The paper outlines the practical implementation of this system, discussing methodologies for training CNN models on labelled datasets and deploying them on Raspberry Pi devices. Additionally, it addresses challenges related to real-time processing, resource constraints, and accuracy in face detection and recognition tasks. The study underlines the potential for low-cost, portable, and efficient solutions in the field of image and video-based facial analysis, offering a comprehensive overview of this promising and evolving research area.

"Emotion Detection Through Facial Feature Recognition" explores various methodologies and techniques employed in the domain, ranging from traditional machine learning approaches to the more recent advancements in deep learning and computer vision. The review encompasses studies focusing on the extraction of facial features, such as expressions, micro-expressions, and action units, and their association with emotional states. Additionally, it covers datasets, algorithms, and evaluation metrics used in emotion recognition systems, highlighting the challenges, including issues related to varying cultural contexts, non-universal facial expressions, and the impact of noise in real-world scenarios. The paper aims to offer a holistic view of the current state of research in the field of emotion detection through facial feature recognition.

III. Methodology

The project works mainly focused on face detection in images or in real-time through a webcam and detects faces, recognizes a person based on the provided database, detects emotion based on facial expression and predicts the age and gender of the person.

1. Face Detection: It can detect faces in images or real-time video streams from a webcam using the face recognition library.

2. Person Recognition: The algorithm matches detected faces with a provided database of known individuals to recognize and label them.

3. Emotion Detection: It analyses facial expressions to recognize and label the emotional states of individuals.

4. Age and Gender Prediction: It predicts the age and gender of the detected individuals based on their facial features.

A. Face Detection

- The Algorithm captures real-time video from the default camera, detects faces, and displays their locations with rectangles. It uses OpenCV and the face recognition library for this purpose. The process includes importing libraries, initializing video capture, face location storage, and real-time loop. Each frame is processed with face detection, and the detected faces are displayed on the original frame. The Algorithm is suitable for human tracking through cameras to detect and continuously mark human detection.

B. Face Detection and Person Recognition

1. Prepare arrays: Initialize arrays to store face locations, encodings, and names.

2. Real-time recognition loop:
    - Continuously process frames from the video stream.
    - Resize the frame to 1/4 size for faster processing.

3. Face detection and encoding:
    - Detect faces in the resized frame using face recognition.
    - Compute face encodings for the detected faces.

4. Recognition:
    - Loop through detected face locations and encodings.
    - Adjust face position magnitude.
    - Compare face encodings with known encodings to find matches.
    - Assign recognized names or "Unknown faces" based on matching results.

C. *Face Detection and Age and Gender prediction*

It uses pre-trained models to estimate the gender (Male/Female) and age group (e.g., '(0-100) yrs.') of each detected face, displaying the results alongside the face rectangles in the video stream.

1. Video Capture and Face Detection Loop:
    - Captures video from the default camera.
    - Enters a loop to process each frame.
2. Face Detection:
    - Detects faces in the frame using face recognition.
    - Adjusts face locations for the actual frame size.
3. Gender and Age Estimation:
    - Estimates gender and age for each detected face.
    - Displays gender and age information on the frame.
4. Real-Time Display:
    - Displays the video frame with detected faces and information in real-time.

D. *Face detection with emotion recognition:*

The Algorithm performs real-time facial expression recognition using a pre-trained model. It detects faces in the webcam feed, classifies the detected facial expressions (e.g., 'happy', 'angry') and displays the recognized emotion labels. Alongside the detected faces.

1. Real-time emotion recognition loop:
    - Continuously process frames from the video stream.
    - Resize the frame to 1/4 size for faster processing.
2. Face Detection:
    - Detects faces in the frame using face recognition.
    - Loop through the detected face locations.
3. Face preprocessing:
    - Extract the detected face from the frame.
    - Convert the face to grayscale and resize it to 48x48 pixels.
    - Normalize pixel values to the [0, 1] range.
4. Emotion prediction:
    - Use the pre-trained model to predict the emotion from the preprocessed face.
    - Determine the emotion label with the highest prediction score.

IV. EXPERIMENTAL SETUP

The dataset was curated using various photographs featuring distinct individuals, primarily for the purpose of person recognition in both image and real-time scenarios. The dataset encompasses images of approximately 10 different individuals. To set up the environment, it's essential to install specific packages, including OpenCV-python, NumPy, dlib, and face_recognition.

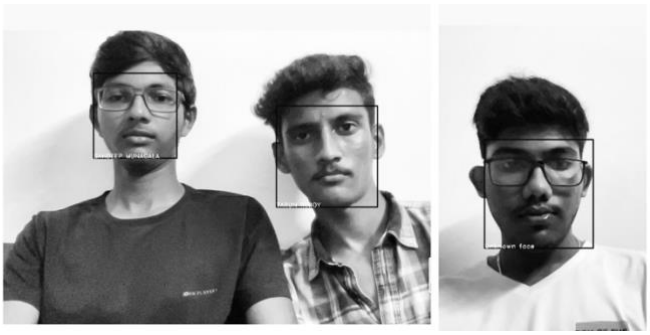

Fig a. Face recognition based on database.

Fig b. Age and gender prediction.

Fig c. Emotion detection based on facial expression.

V. RESULTS AND DISCUSSION

The accuracy of this system is influenced by the quality of images, and in real-time scenarios, it's subject to various factors such as lighting conditions, the person's distance from the camera, and the presence of any obstacles. Face detection performs optimally when individuals are at a proximity of 1 to 2 meters, although some variations may be observed at greater distances. Additionally, the effectiveness of person recognition is dependent on the reference images available in the dataset. While age and emotion detection functions, there is room for improvement to enhance their accuracy.

VI. Conclusion

The project successfully implements a comprehensive system for face detection, emotion recognition, age prediction, and gender identification using computer vision techniques. It offers real-time face detection on webcam feeds and pre-recorded videos, enhancing human-computer interaction and security applications. The integration of OpenCV, face recognition libraries, and deep learning models has facilitated efficient facial analysis.

Future works

1. Improved Accuracy: Future work can focus on enhancing the accuracy of emotion recognition and age prediction models through fine-tuning and larger training datasets.

2. Real-World Applications: The technology can be applied in diverse real-world scenarios, including personalized advertising, smart surveillance, and emotion-aware user interfaces.

3. Privacy and Ethics: Considerations for privacy and ethical use of facial analysis technologies should be addressed, including robust data protection measures and consent mechanisms.

4. Adaptability: The system can be further adapted to recognize specific facial expressions, gestures, and complex emotions, allowing for more nuanced interaction.

Reference

[1] Article: Face Detection & Recognition from Images & Videos Based on CNN & Raspberry Pi, Muhammad Zamir, Department of Software Engineering, Mirpur University of Science & Technology, Mirpur 10250, Pakistan

[2] Shivam Singh, Prof. S. Graceline Jasmine: Face Recognition System Department of CSE, Vellore Institute of Technology, Chennai Tamil Nadu, India, International Journal of Engineering Research & Technology (IJERT) (This work is licensed under a Creative Commons Attribution 4.0 International License.) Published by Vol. 8 Issue 05, May 2019

[3] Bradski, G.; Kaehler, A. Learning OpenCV: Computer Vision with the OpenCV Library; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2008.

[4] Wang, J.; Li, Z. Research on face recognition based on CNN. In Proceedings of the IOP Conference Series: Earth and Environmental Science, Banda Aceh, Indonesia, 26–27 September 2018; Volume 170, p. 032110.

[5] Dalal, N.; Triggs, B. Histograms of oriented gradients for human detection. In Proceedings of the 2005 IEEE ComputerSociety Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–26 June 2005; Volume 1, pp. 886–893.

[6] Ge, H.; Zhu, Z.; Dai, Y.; Wang, B.; Wu, X. Facial expression recognition based on deep learning. Comput. Methods Programs Biomed. 2022, 215, 106621. [CrossRef] [PubMed]

[7] Faizan Ahmad, Aaima Najam: Image-based Face Detection and Recognition, Department of Computer Science & Engineering, Beijing University of Aeronautics & Astronautics, Beijing, 100000, China

[8] G. L. Marcialis, F. Roli, Chapter: Fusion of Face Recognition Algorithms for Video-Based Surveillance Systems, Department of Electrical and Electronic Engineering- University of Cagliari- Italy.

[9] Emotion Detection Through Facial Feature Recognition: James Poa, jpao@stanford.edu

[10] Face Detection and Recognition Using Face Mesh and Deep Neural Network, Shivalila Hangaragia, Department of Electronics and Communication Engineering, Amrita School of Engineering, Bengaluru, Amrita Vishwa Vidyapeetham, India, International Conference on Machine Learning and Data Engineering,

[11] P. Yaswanthram and B. A. Sabarish. (2022) "Face Recognition Using Machine Learning Models - Comparative Analysis and Impact of Dimensionality Reduction," IEEE Fourth International Conference on Advances in Electronics, Computers, and Communications (ICAC): pp. 1-4,