International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 4 April 2023, pp: 850-856 www.ijaem.net ISSN: 2395-5252

FRIV: Face Detection in Video

Bhavin Rajesh Gaikwad, Swati Umashankar Pal, Dhananjay Budhiram Prajapati, Lakshit Vijay Reddy, Prof. Silviya D' monte

Department of Computer Engineering Universal College of Engineering Mumbai, India
Department of Computer Engineering Universal College of Engineering Mumbai, India
Department of Computer Engineering Universal College of Engineering Mumbai, India
Department of Computer Engineering Universal College of Engineering Mumbai, India
Department of Computer Engineering Universal College of Engineering Mumbai, India

---

Date of Submission: 10-04-2023 Date of Acceptance: 20-04-2023

ABSTRACT—Face Detection in a live video is a crucial task for ensuring public safety and security. It can be also used for various places including institutions, organizations, offices, and social places, etc. Finding particular person who is present in a CCTV/ live video is a major task when the person is in the crowded place and this task is much difficult with the help of the human intervention, So in the existing system this task is implemented which has the less accuracy and it may not detect the particular person who is at a distant so in this case it needs the strong processor to run the algorithm in the backend. So to solve this problem we have come up with a way in which the admin will just upload the particular person's photo in image format and after uploading the system will detect the particular person in that live CCTV or a video which has the better accuracy compared to previous existing systems. This system is much portable that it can run on any low-end devices as the admin has to only upload the one image of the particular person so it will take less time to process and the system will work efficiently in this case it will reduce the production cost, also the system can detect the particular person at a long distance.

Keywords— CCTV / Video, face detection, Suspect, HOG Algorithm, face Comparison, distance, Image processing, Image Encoding, image comparison.

I. INTRODUCTION

Real-time video surveillance is an increasingly significant aspect of the intelligent world. It enables sensory connections worldwide and serves as a link between the digital and real worlds, acting as a catalyst for the digital transformation of surveillance applications. Machine learning, deep learning, and image processing techniques have opened up new research possibilities in this field, with deep learning allowing for automated information extraction and analysis from images and videos. Traditional face detection methods rely on fixed surveillance cameras that capture similar-sized face targets; however, this approach falls short in crowded environments where moving cameras capture faces of varying sizes.

In light of this, let us consider authorities who are currently on the lookout for a criminal suspect who has evaded capture for several weeks. The suspect is known to frequent crowded areas such as train stations and airports, making it difficult for authorities to locate them. Additionally, a child has been reported missing in a busy shopping mall that has numerous CCTV cameras installed. However, given the high volume of people in the area, locating the child manually has proved to be challenging. So, we implemented a system that can aid in finding a missing person or identifying a criminal in crowded areas by utilizing live CCTV camera feeds. The system is capable of analyzing live video streams and monitoring individual's movements to determine if they match the missing person or suspect's description.

Our advanced detection process is highly effective, which is enabling us to precisely and accurately identify individuals from both images and video footage. The process commences by taking a single image as an input and uploading it to our system. The resulting value is stored in our system for future use. In the next phase of our process, we take either live surveillance camera feeds or pre-recorded video footage and feed it into our system. Our system then breaks the video down into individual frames, and proceeds to detect any faces that may be present in each frame. Our system then compares the detected face with the face values stored in our system. In cases where a match is found, a bounding box is created around the face, and the individual's name is displayed, making it easy to identify them. This detection process is highly beneficial, particularly in real-time, where quick identification of a particular individual or suspect may be required.

The paper is organized into different sections to present the research work in a systematic and coherent way. In Section II, a comprehensive literature review is presented, which provides an in-depth analysis of previous research studies related to the topic. This section aims to identify the research gaps and highlight the significance of the current study. Section III outlines the implemented approach of the project, which includes the methodology, techniques, and tools used for data collection, analysis, and interpretation. The section provides a detailed explanation of the research design, experimental setup, and procedures used to carry out the study. The results of the research are presented in Section IV, along with a comparative analysis of the findings. This section presents the key findings of the study and discusses their implications for the research topic. The results are compared with the existing literature and other related studies to provide a broader perspective. Section V concludes the paper by summarizing the main findings, contributions, and limitations of the research. It also provides recommendations for future research studies and suggests potential areas of improvement for the current project. Finally, Section VI presents the future work that can be undertaken to upgrade the project, which includes adding more features and improving its efficiency. This section outlines the potential research directions that can be pursued to enhance the project's capabilities and broaden its scope.

II. LITERATURE REVIEW

[1] SamitShirsat, Aakash Naik, have tried to survey the existing technologies as well as propose a new system for criminal Detection & Recognition using Cloud Computing and Machine Learning, which if used by our Crime Agencies would help them to find criminals from CCTV footage. The proposed system can help to find missing children and people from the CCTV footage available from the respective site.

[2] Jee Hun Kang, John Grossi, Son Tung Do, in their system the Cameras can capture thousands of hours of footage, making manual review an arduous task. By creating a model that can detect if a crime is taking place in a still image, we can assist in reducing this review process. In this research project, we investigated how to apply state-of-the art computer vision models to identify anomalous images among 14 different classes.

[3] Surbhi Gupta, Kutub Thakur, Munish Kumar, they have presented the feature-based method for 2D face images. speeded up robust features (SURF) and scale invariant feature transform (SIFT) are used for feature extraction. Various combinations of SIFT and SURF features with two classification techniques, namely decision tree and random forest, have experimented in this work. A maximum recognition accuracy of 99.7% has been reported by the authors with a combination of SIFT (64-components) and SURF (32-components).

[4] JEREMIAH R. BARR, KEVIN W. BOWYER, PATRICK J. FLYNN, and SOMA BISWAS, conducted a comprehensive survey on facial recognition research, organizing recent methods for addressing challenges in unconstrained settings. Their paper explores the similarities and differences between human and algorithmic facial recognition, and also highlights important research challenges and opportunities for the future of the field.

[5] Imran Ahmed and Gwanggil Jeon proposed various approaches to person detection and tracking using image processing and artificial intelligence, including machine and deep learning. However, their focus was primarily on frontal view camera perspectives. The authors introduced a real-time person tracking and segmentation system that uses an overhead camera perspective and a deep learning-based algorithm called SiamMask. The algorithm has demonstrated excellent performance, outperforming other real-time tracking algorithms, with a tracking accuracy of 95%. The paper concludes with a comparison of SiamMask to other tracking algorithms.

[6] Deng-Yuan Huang, Chao-Ho Chen, Tsong-Yi Chen, Jian-He Wu, and Chien-Chuan Ko presented a real-time face detection system that can be used with a moving camera. Their system consists of three modules: (1) face candidate detection, which generates candidates using skin color, edges, and face area information; (2) face candidate verification, which uses HOG features and a two-class C-SVM classifier with pretrained face samples to determine whether candidates are real faces; and (3) face tracking, which estimates the overlap between face targets in current and previous frames to determine whether tracking will be continuous. The authors report that the system successfully detects most faces in open spaces,

which can help quickly identify specified individuals and prevent potential criminal events. Additionally, the paper notes that the proposed system significantly improves face detection accuracy.

[7] Zhenfeng Shao, Jiajun Cai, and Zhongyuan Wang introduced a new solution for processing and utilizing large amounts of surveillance video data. Their approach is based on event detection and alarming messages generated by smart cameras. The method consists of three parts: intelligent pre-alarming for abnormal events, smart storage for surveillance videos, and rapid retrieval of evidence videos. The authors highlight that their approach explores temporal-spatial association analysis with respect to abnormal events in different monitoring sites.

[8] Shu Liao, Dinggang Shen, and Albert C.S. Chung proposed a new framework for addressing the face recognition problem. The authors' contributions include: (1) representing each pixel in a facial image using an anatomical signature derived from the most salient scale local region, determined by the survival exponential entropy (SEE) information-theoretic measure; (2) using the anatomical signature of each pixel, introducing a novel Markov random field-based groupwise registration framework to formulate the face recognition problem as a feature-guided deformable image registration problem, with similarity between different facial images measured on a nonlinear Riemannian manifold based on the deformable transformations; and (3) addressing the generalizability problem that commonly arises in learning-based algorithms. The proposed method was extensively evaluated on four publicly available databases, including FERET, CAS-PEAL-R1, FRGC ver 2.0, and LFW. Experimental results demonstrated that the proposed method consistently achieved the highest recognition rates among all the methods under comparison.

[9] Lamiaa A. Elrefaei, Alaa Alharthi, Huda Alamoudi, Shatha Almutairi, and Fatima Al-rammah proposed a criminal detection framework aimed at helping police officers recognize the face of a criminal or a suspect. This paper focuses on developing the client side of the framework, specifically face detection and tracking using Android mobile devices. To detect faces, the authors employed a robust Viola-Jones algorithm that is not affected by illuminations. For face tracking, the Optical Flow algorithm was used with two feature extraction methods: Fast Corner Features and Regular Features. The proposed face detection and tracking method was implemented using Android studio and OpenCV library, and the Sony Xperia Z2 Android 5.1 Lollipop Smartphone was used for testing. Experimental results show that face tracking using Optical Flow with Regular Features achieves a higher level of accuracy and efficiency than Optical Flow with Fast Corner Features.

[10] Xingxian Zhang, Chuanchang Liu and ZhiyuanSu, With the rise of electronic information and computer science technology, the ease of video capture and decreasing storage costs has led to a significant increase in the volume of available video. However, manual annotation and processing are unable to handle such a massive amount of data, resulting in a need for automated video analytics and interpretation technologies. The fundamental basis for these video analytics applications is the automatic detection algorithm for video targets. This paper examines the popular face detection algorithm and its applications in video streaming. Based on the circumstances, a face detection system is implemented using video streaming.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 4 April 2023, pp: 850-856 www.ijaem.net ISSN: 2395-5252

III. IMPLEMENTED APPROACH

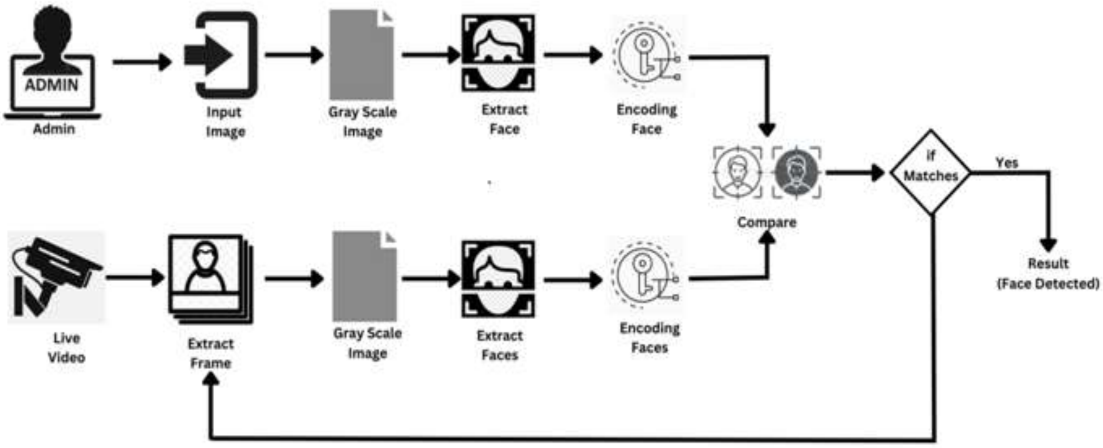

Fig.1. System architecture of Face Detection in Video

The flowchart of the proposed system is shown in Fig.1. In our proposed system the user/admin will upload the photo of the suspect through our portal. Then that image will get encoded using OpenCV package and get uploaded into the media folder and the path of that particular image will get saved into the PostgreSQL database. After this process, the system will detect that particular suspect into a live streaming video e.g. CCTV/ Video. For detecting a particular suspect in a video we have used CNN algorithm and if the suspect is found then the system will return a message as suspect match found. Then the system will draw a box on a suspect face and it will display the name of a suspect below that particular area in which the box is plotted.

3.1. Encoding and Loading Image

Traditionally, multi-scale face detection involves constructing an image pyramid and searching for face targets at different scales using a sliding window approach. While effective, this method requires a significant amount of computational time. In this study, we propose a more efficient algorithm for detecting potential regions where face targets may be present, which reduces the computational complexity of multi-scale face detection in crowded scenes.

3.1.1. Detection of skin tone

To enhance the performance of face detection by eliminating the impact of variations in illumination, color space conversion is often used. The RGB color space is not ideal for color analysis since the three components are highly correlated between channels. Therefore, in this study, the HSV color space was employed with a skin color model to detect face regions. An efficient skin color detection method is essential for face detection in complex scenarios with diverse skin colors and lighting conditions.

3.1.2 Detection of Edges

Sobel operators aim to implement first derivatives in image processing using the magnitude of the gradient for detecting edges. The first derivative of an image can be expressed as $\nabla f=[G_x, G_y]=[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}]$, and the magnitude as well as angle of the gradient can be formalized as $M(x, y) =$

$$\sqrt{G_x^2 + G_y^2} \quad \text{and} \quad \theta(x, y) = \tan^{-1}\left(\frac{G_y(x,y)}{G_x(x,y)}\right)$$

respectively. Therefore, horizontal and vertical edges can be effectively detected to form a binary image, i.e., $Edge_H(x, y)$ and $Edge_V(x, y)$, by the proposed formula, where θ_M , θ_H and θ_V are thresholds that are determined empirically.

3.1.3 Estimation of Face area

The HaarCascade classifier is used to detect the face in the preprocessed image. This classifier is trained to recognize specific patterns in the image that are characteristic of a face, such as the shape of the eyes, nose, and mouth. Once the classifier detects a face, it returns a bounding box that encloses the face. The size of the bounding box returned by the face detection step is used to estimate the face area. The area of the bounding box can be computed by multiplying its width and height. Since the bounding box tightly encloses the face, the area of the bounding box provides a good estimate of the actual face area.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 4 April 2023, pp: 850-856 www.ijaem.net ISSN: 2395-5252

3.1.4 Encoding

Face encoding is a process in computer vision that involves analyzing a face image and generating a numerical representation, or encoding, of the face that can be used to identify the person in the image. Once the face is detected, the next step is to align the face in a standardized manner. This involves rotating and scaling the image so that the face is centered and aligned with the eyes and mouth. Once the face is aligned, a deep neural network is used to generate a numerical representation of the face. The neural network is trained on a large dataset of faces and learns to extract features that are unique to each face.

3.2 Detection of Face from CCTV

This involves connecting the CCTV camera to a computer or server and configuring the camera settings. The next step is to capture the live video feed from the CCTV camera using OpenCV's video capture module. Once the live video feed is captured, the next step is to detect faces in the video stream. This can be done using pre-trained deep learning models such as Haarcascades. The video will be further divided frame by frame and from each frame to detect face and encode it, it undergoes previous step (1) Detection of skin tone , (2) Detection of Edges , (3) Estimation of Face area and (4) Encoding.

3.3 Comparing Faces

The compare faces module in OpenCV is a function that compares two face encodings to determine if they belong to the same person. The first step is to generate the face encodings for the two faces that need to be compared. Once the face encodings are generated, the next step is to calculate the distance between the two encodings using the Euclidean distance metric in OpenCV. The distance between the two face encodings is then compared to a predefined threshold value. If the distance is below the threshold, the two face encodings are considered a match and are assumed to belong to the same person. If the distance is above the threshold, the two face encodings are considered different and are assumed to belong to different people. The compare faces function returns a boolean value indicating whether the two face encodings belong to the same person or not. If the encodings are matched, it will draw a bounding box around the face. If it doesn't find a face or doesn't match the encoding value, it will keep on checking the next frame.

IV. RESULT

Algorithm	Accuracy	Speed	Distance Range
CNN	> 99%	Slow	Long
HOG	> 95%	Fast	Medium
Viola - Jones	> 90%	Fast	Short
SIFT	> 90%	Slow	Short
SURF	> 90%	Slow	Short
SiamMask	> 95%	Fast	Short
SSD	> 95%	Fast	Short

Table.1. Algorithm Comparison

After conducting a comparative analysis between our implementation of the HOG algorithm and previous projects or existing systems, we found that our implementation achieves high accuracy, speed, and range. This implies that our implementation has outperformed the previous systems in terms of the desired outcomes.

Furthermore, our implementation has met our project requirements and preferences. The desired outcomes were determined based on the specific objectives of the project and the user's needs. Our implementation of the HOG algorithm was able to achieve these desired outcomes effectively, which indicates that it has been successful in meeting the project's goals.

V. CONCLUSION

Our advanced system is designed to efficiently identify individuals, even in crowded places, using only a single image. With the help of cutting-edge technology and powerful algorithms, our system can quickly scan a large area and accurately detect and identify a particular individual based on their facial features. This can be incredibly valuable in situations where there is a need to locate a person of interest, such as in the case of a missing person or a criminal suspect. Additionally, our system can also detect a particular person from a faraway distance, making it ideal for use in large-scale surveillance applications. Thanks to our system's impressive detection speed, we can quickly and easily identify individuals from a distance, even in complex and challenging environments, whether it's a crowded city street or a remote wilderness area. Our system provides fast and reliable identification that can help keep people safe and secure.

The system we have developed is an efficient tool that aids law enforcement authorities in identifying criminal suspects in densely populated areas. It provides an accurate and reliable method of identifying individuals who may be of interest to law enforcement agencies, enabling them to take appropriate action to prevent or solve crimes. Not only does the system assist in identifying suspects, but it can also assist in locating missing persons using surveillance cameras, without the need for manual intervention. Furthermore, if a criminal suspect changes their appearance, such as environments, by wearing a fake beard or mustache, our system is still able to accurately identify them by comparing their image with previous records, thus making it easier for authorities to track them down and bring them to justice. Overall, our advanced facial recognition system is a powerful tool that can help law enforcement agencies and security personnel to identify individuals quickly and accurately, even in challenging ultimately improving public safety and security.

VI. FUTURE WORK

Integration with existing surveillance systems: The facial recognition system can be integrated with existing surveillance cameras in public places such as airports, shopping centers, and public transportation. This will enable law enforcement agencies to quickly identify and track down suspects, potentially reducing crime rates and improving public safety.

Improving the accuracy of the system: Although the current system is accurate, there is always room for improvement. Future work can focus on improving the system's accuracy, especially in challenging and complex environments such as low light or noisy areas.

Addressing privacy concerns: As with any facial recognition system, privacy concerns are always present. Future work can focus on addressing these concerns by implementing strict privacy policies and regulations and ensuring that the system is only used for legitimate purposes.

Database management: The accuracy and effectiveness of a facial recognition system depends on the quality and diversity of the database it uses for comparison. Researchers could work on developing efficient database management systems that can handle large volumes of images, ensure data privacy and security, and allow for the easy addition of new data as it becomes available. This could help to improve the accuracy and effectiveness of facial recognition systems over time.

Developing a mobile application: The facial recognition system can be developed into a mobile application that can be used by law enforcement agencies and security personnel in the field. This will enable them to quickly identify individuals using their mobile devices, potentially reducing response times and improving public safety.

REFERENCES

[1]. SamitShirsat, Aakash Naik, Darshan Tamse, Jaysingh Yadav, Pratiksha Shetgaonkar and Shailendra Aswale. Proposed System for Criminal Detection and Recognition on CCTV Data Using Cloud and Machine Learning.

[2]. Jee Hun Kang, John Grossi and Son Tung Do. Crime Detection in Security Camera Footage.

[3]. Surbhi Gupta, Kutub Thakur and Munish Kumar. 2D-human face recognition using SIFT and SURF descriptors of face's feature regions.

[4]. JEREMIAH R. BARR, KEVIN W. BOWYER, PATRICK J. FLYNN and SOMA BISWAS.FACE RECOGNITION FROM VIDEO: A REVIEW.

[5]. Imran Ahmed and Gwanggil Jeon. A real-time person tracking system based on SiamMask network for intelligent video surveillance.

[6]. Deng-Yuan Huang, Chao-Ho Chen, Tsong-Yi Chen, Jian-He Wu and Chien-Chuan Ko. Real-Time Face Detection Using a Moving Camera

[7]. Zhenfeng Shao, Jiajun Cai and Zhongyuan Wang. Smart Monitoring Cameras Driven Intelligent Processing to Big Surveillance Video Data.

[8]. Shu Liao, Dinggang Shen, and Albert C.S. Chung. A Markov Random Field Groupwise Registration Framework for Face Recognition.

[9]. Lamiaa A. Elrefaei, Alaa Alharthi, Huda Alamoudi, Shatha Almutairi and Fatima Al-rammah. Real-time Face Detection and Tracking on Mobile Phones for Criminal Detection.

[10]. Xingxian Zhang, Chuanchang Liu and ZhiyuanSu. Face Detection System Based on Video Stream.