International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 Impact Factor: 8.176 ISSN: 2582-3930

Human Detection in Surveillance Video

Ajay Dandge[1], Ajit Patil[2], Mayur Nehare[3], Hrushikesh Shukla[4], Prof. Pallavi Bute[5]

1234, Final Year Student, Department of Information Technology, SSGMCE, Shegaon, Maharashtra
5, Assistant Professor, Department of Information Technology, SSGMCE, Shegaon, Maharashtra

---

Abstract - In recent years, there has been a lot of focus on the crucial task of human detection in surveillance footage. Security, public safety, traffic monitoring, and healthcare are just a few areas where it might be used. This paper provides an overview of the technique, as well as its uses and problems. Several processes are involved in the human detection process, including object detection, face detection, and tracking. The Face-recognition Library is a versatile tool that contains face detection and recognition algorithms that may be used to identify people in photos and videos. It can be used for duties such as criminal investigations and traffic monitoring for safety infractions, allowing for precise and efficient human detection in surveillance video.

Key Words: human detection, surveillance video, face detection

1.INTRODUCTION

This Human detection in surveillance video is the problem of locating and identifying humans in video streams captured by cameras. Human detection in surveillance video is important for various domains and applications, such as security, safety, public services, healthcare, entertainment, etc. Human detection in surveillance video can help to identify and track criminals or suspects, verify and authenticate civilians or citizens, detect and alert abnormal or suspicious events, analyze and understand crowd characteristics and dynamics, re-identify and recognize persons of interest, estimate and classify gender or age groups, characterize and model human gait or posture, detect and prevent falls or injuries, etc.

There are three main groups of methods for human detection in surveillance video: pixel-based methods, edge-based methods, and part-based methods. Pixel-based methods use low-level features such as color, intensity, texture, or motion to segment the foreground pixels from the background pixels, and then group them into human regions or blobs. Edge-based methods use edge features such as gradients, contours, or corners to detect human silhouettes or shapes. Part-based methods use high-level features such as body parts or landmarks to model the human appearance and structure.

Pixel-based methods are fast and simple, but they are sensitive to noise, illumination changes, occlusion, and background clutter. Edge-based methods are robust to illumination changes and color variations, but they are sensitive to noise, occlusion, and pose variations. Part-based methods can handle pose variations and partial occlusion, but they require more computational resources and prior knowledge of human anatomy.

Some examples of pixel-based methods are background subtraction, optical flow, and spatio-temporal filtering. Some examples of edge-based methods are shape discrimination using Fourier descriptors, human detection based on a probabilistic assembly of robust part detectors, and detecting human motion with support vector machines. Some examples of part-based methods are histograms of oriented gradients, monitoring of cyclists and pedestrians, and pedestrian detection using boosted features.

There are some open challenges and limitations for human detection in surveillance video research, such as scalability, robustness, privacy, etc. Scalability refers to the ability to handle large-scale and high-resolution data efficiently and effectively. Robustness refers to the ability to handle various variations and uncertainties in human appearance, camera settings, environmental conditions, and scene complexity. Privacy refers to the ability to protect the privacy and security of the personal data and biometric information of the civilians or citizens who are detected and recognized by the surveillance systems.

There are some possible future directions and opportunities for improving human detection in surveillance video research, such as multimodal fusion, domain adaptation, explainable AI, etc. Multimodal fusion refers to the use of complementary and supplementary information from multiple sources and modalities, such as thermal cameras, infrared cameras, depth cameras, audio sensors, etc. Domain adaptation refers to the adaptation of learned models or features from one domain or scenario to another, such as indoor vs outdoor, day vs night, crowded vs sparse, etc. Explainable AI refers to the provision of explainable and interpretable results by revealing the underlying logic or rationale behind the detection decisions or predictions.

2.LITERATURE REVIEW

Human Identification Using Human Body Features Extraction, Martino C. Khuangga & Dwi H. Widyantoro [1]

This work offers a human-identification-system that uses human body feature extraction and can monitor a room's occupants' presence. To eliminate the requirement for particular action, the system employs a camera as an input device. Features on the human body are choose because they are often simpler to identify and serve as a strong identification to label people who enter a place. This system uses two primary processes. The system first notices when someone enters a room. Additionally, it notices when the same individual leaves the room. Utilising image processing methods including HOG descriptor-based person recognition, HSV colour conversion, and template matching, this application was created.

Human Identification Recognition In Surveillance Videos, Kai Jin, Xuemei Xie , Fangyu Wang, Xiao Han, Guangming Shi [2]

Despite having superior performance, users still requires to cooperate with the camera while using facial-recognition. Due to the blurred and warped settings in surveillance footage, human identify recognition works poorly. In this study, we provide a fresh approach to identify human identity by combining global and local structural data. First, in order to enhance the recognition of faces with occlusions and facial deformities, researchers integrate pedestrian-detection and tracking with face recognition. Additionally, selective-recognition method based on person's trajectory that may be used to recognize a person approaching a webcam are offered. Compared to techniques that just use facial information, our strategy performs better.

Human Detection and Tracking for Video Surveillance: A Cognitive Science Approach, Vandit Gajjar, Ayesha Gurnani, Yash Khandhediya [3]

Video surveillance is becoming increasingly vital as a result of the growth in crimes worldwide. Since there aren't enough people to manually monitor all of these cameras, new computer vision algorithms are being created to handle both simple and complex jobs. To identify people in video sequences, we have created a novel technique that combines the most renowned Histograms of Oriented Gradients, the theory of Visual Saliency, and the Deep Multi-Level Network saliency prediction model. Additionally, we used the k-Means technique to cluster the HOG feature vectors of the positively identified windows and identify a person's route in the video.

Real-Time Criminal Identification System Based On Face Recognition, Mr.R.Prashanth Kumar, Abdul Majeed, Farhan Pasha, A Sujith[4]

This facial recognition-based real-time criminal detection system uses a facial recognition technology that is completely automated. For face identification and recognition, OpenCV LBPH (Local Binary Pattern Histograms) Algorithms and the Haar feature-based cascade classifier are both employed. This technology will be able to recognise faces in real time and automatically detect faces. It's still difficult to pinpoint the face with accuracy. The Viola-Jones framework has been widely utilised by academics to identify faces and other objects in a picture. Public communities like OpenCV share face detection classifiers.

3. METHODOLOGY

We make use of the Face-recognition Library which is an open-source library for face recognition and detection tasks. It is written in Python and is built on the dlib-library, a contemporary C++ toolkit that includes machine learning methods and resources.

The working of the Face-recognition Library can be divided into three main parts: face-detection, face-alignment, and face-recognition.

A. Face Detection:

The crucial step in the Face-recognition Library is the face detection. The library uses an object detection algorithm to identify the location of faces in the input image. The library supports two different methods of face detection: the HOG-based method and the CNN-based method.

The HOG-based method uses a Histogram of Oriented Gradients (HOG) feature descriptor and a linear support vector machine (SVM) classifier to identify the presence of faces in an image. This method is faster but less accurate than the CNN-based method.

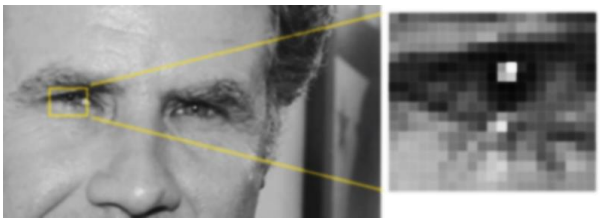

Figure.1 Pixel level representation of image

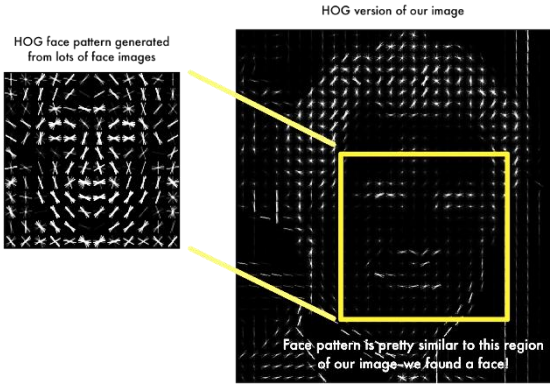

Figure.2 HOG pattern

The CNN-based method uses a deep convolutional neural network to detect faces in an image. This method is more accurate but slower than the HOG-based method.

B. Face Alignment:

Once the faces have been detected, the library uses face alignment to normalize the faces to a common coordinate system. The face alignment process involves identifying the facial landmarks, like the eyes, nose, and mouth, and then applying a geometric transformation to warp the image to a canonical face representation. This step helps to remove the variations in pose, scale, and orientation of the face images, making them more suitable for face recognition.

C. Face Recognition:

The final step in the Face-recognition Library is face recognition. The library uses a deep neural network to extract features from the aligned face images. These features are then compared to the database of known faces to identify the individual in the image. The library supports several different face recognition algorithms, including the classic Eigenfaces method, Fisherfaces method, and Local Binary Patterns Histograms (LBPH) method.

This library also provides additional features, such as the ability to train custom face recognition models using user-supplied training data, and the ability to perform face clustering to group similar faces together.

Overall, the Face-recognition Library is a powerful tool for face recognition and detection tasks, allowing for the accurate identification of individuals in images and videos. Its ability to detect and align faces, as well as its support for multiple recognition algorithms, make it a versatile tool for a wide range of applications.

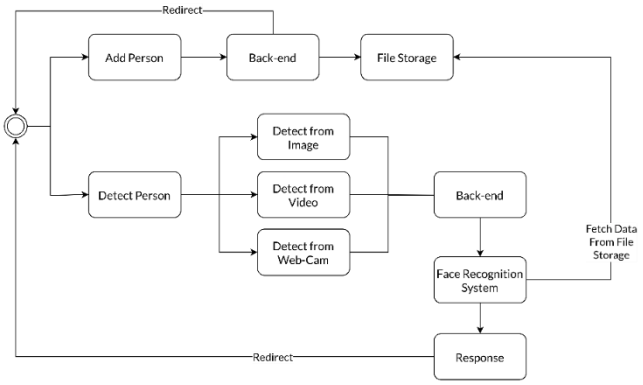

Figure 4. Block diagram of the System

*D. **Desired Output***

The Human Detection in Surveillance Video System will be provided to the organization, where user will be able to perform different available tasks. After login into the portal it asks to upload picture, upload video and detect using webcam through which user can detect the wanted person, where webcam is used to model a CCTV as shown in fig. 5.

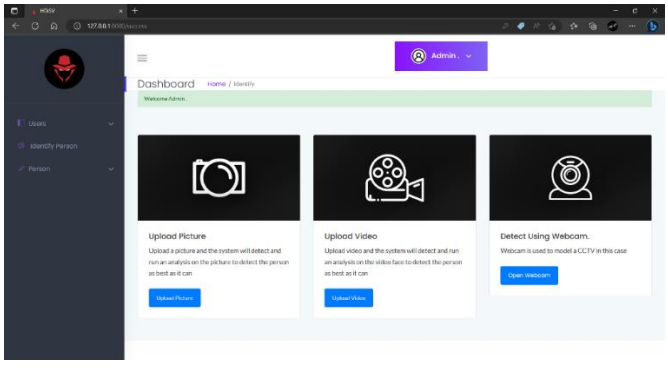

Figure 5. Proposed System

After selecting particular option, the system detects the face of that individual and matches it with the one already present in the database. When Upload Video option is selected, the system will perform analysis on the video, and if the person is present, it will show the frame with that person's information and timestamp of when matched as shown in fig. 6.

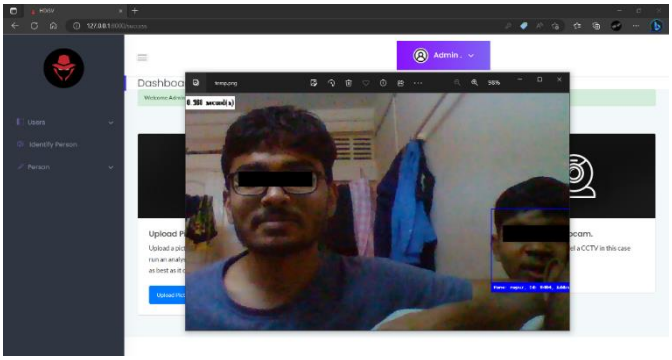

Figure 6. Output

The system provides the functionality to mark particular person as wanted, which will help in avoiding unwanted labelling of other individuals. The action tab will be used to mark person wanted or free as shown in fig. 7.

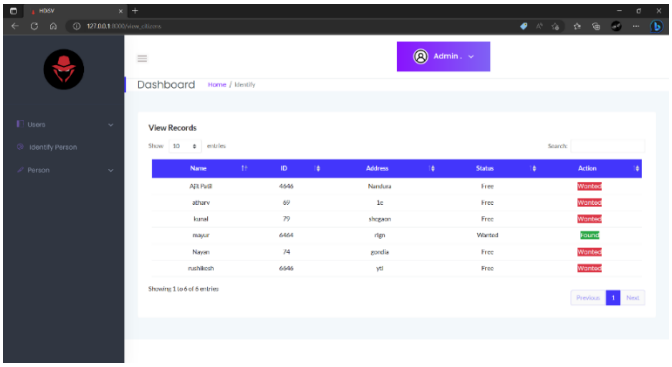

Figure 7. View Records

With the help of this system it becomes easier to detect individuals by matching it with the available in the database. At the same time, we can reduce human efforts required to go through large dataset of images and videos.

4. RESULTS

We used a dataset of pictures videos and live broadcasts to evaluate our technology the system performed well in live video processing and it was able to alert the user in real-time if a certain person is present in the video it was also able to identify and recognize faces effectively and return the frame in which a person first appears in the video at the same time providing the timestamp.

5. CONCLUSION

This approach effectively identifies people in surveillance footage despite challenges such occlusions, changes in illumination, and variations in facial expressions, especially for applications involving criminal detection. It is essential to take ethical issues into account when utilising face recognition technology for such reasons. The accuracy and dependability of human detection utilising facial recognition libraries are anticipated to increase with sustained study and development, helping law enforcement authorities to uphold public safety.

Overall, finding techniques that are precise, reliable, scalable, and privacy-preserving, while also being open and understandable, is crucial to the future of human identification in surveillance video research. The route forward can be paved by developments in multimodal fusion, domain adaptability, explainable AI, and privacy-preserving techniques.

REFERENCES

[1] Martino C. Khuangga & Dwi H. Widyantoro, “Human Identification Using Human Body Features Extraction”, ICACSIS 2018 978-1-7281-0135-4/18/S31.00, 2018 IEEE
[2] Kai Jin, Xuemei Xie, Fangyu Wang, Xiao Han, Guangming Shi, “Human Identification Recognition In Surveillance Videos”, 2019 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), Shanghai, China
[3] Vandit Gajjar, Ayesha Gurnani, Yash Khandhediya, “Human Detection and Tracking for Video Surveillance: A Cognitive Science Approach”, 2017 IEEE International Conference on Computer Vision Workshops, DOI 10.1109/ICCVW.2017.330
[4] Mr.R.Prashanth Kumar, Abdul Majeed, Farhan Pasha, A Sujith, “Real-Time Criminal Identification System Based On Face Recognition” V26, Num 05, May/2020| Advanced Science Letters | E-ISSN:1936-7317S.
[5] V. Tathe and S. P. Narote, “Real-Time Human Detection and Tracking”, in Annual IEEE India Conference, Mumbai, 2013.
[6] Dr. Asif Ali, Radhika Mandhanya, Shraddha Birla, Ujjwal Mandloi and Vipul Jain’s “Automatic Face Recognition Attendance System using Python and OpenCV”, GRD Journals- Global Research and Development Journal for Engineering | Volume 6 | Issue 4 | March 2021
[7] R. Rathi, M. Choudhary & B. Chandra, “An Application of Face Recognition System using Image Processing and Neural Networks”, International Journal Computer Technology Application, 3:1, (2012), pp. 45-49.
[8] I. Haritaoglu, D. Harwood and L. S. Davis, “Real-time surveillance of people and their activities”, IEEE Trans. Pattern Anal. Mach. Intell. 22 (8) (2000) 809–830.
[9] A. Mahapatra, T.K. Mishra, P.K. Sa, B. Majhi, “Background subtraction and human detection in outdoor videos using fuzzy logic”, in Proc. IEEE Int. Conf. Fuzzy Systems, July 2013, pp. 1–7.
[10] Kim, Jiwon, Jung Kwon Lee, and Kyoung Mu Lee, “Accurate image super-resolution using very deep convolutional networks.” in IEEE conference on computer vision and pattern recognition(CVPR),2016.
[11] Liu, Weiyang, “Shereface:Deep hypersphere embedding for face recognition.” in IEEE conference on computer vision and pattern recognition(CVPR), 2017.
[12] Parkhi, Omkar M., Andrea Vedaldi, and Andrew Zisserman. “Deep face recognition” bmvc. vol. 1 no. 3 pp. 41.1-41.12,2015.
[13] Viola, Paul, and Michael J. Jones. "Robust real-time face detection." International journal of computer vision vol. 57 no. 2 pp. 137-154,2014
[14] K. Mikolajczyk, C. Schmid, A. Zisserman, “Human detection based on a probabilistic assembly of robust part detectors”, in Proc. European Conf. Computer Vision (2004), pp. 69–81.
[15] Y.-T. Tsai, H.-C. Shih and C.-L. Huang, “Multiple Human Objects Tracking in Crowded Scenes,” in 18th International Conference on Pattern Recognition, Hong Kong, 2006.