Advances in Parallel Computing Algorithms, Tools and Paradigms
D.J. Hemanth et al. (Eds.)

doi:10.3233/APC220054

Smart Surveillance System Using Face and Optical Character Recognition for Secure Environment

Lakshmi HarikaPalivela[a,1], P. M. Ashok Kumar[b] and V.V. Rama Krishna[c]
[a] *Department of IT, M IT, Anna University, India*
[b]*Department of CSE, KoneruLakashmaiah Education Foundation, Vaddeswaram*
[c]*Department of ECE, Lakireddy Bali Reddy College of Engineering, Mylavaram, India*

Abstract. One of the important applications in the smart surveillance system is to identify unknown suspicious persons. Instead of manually and tediously monitoring the cameras continuously, this system can be used to identify and recognize suspected individuals and consequently send a warning message on the occurrence of such recognition. In this work, we propose a novel methodology, in which the system is connected to a database containing images of the Aadhar cards of the malfunctioning individuals. Since SQL database cannot directly store images, the Aadhar card pictures will have to be stored in a NoSQL database. The proposed system was built utilizing the Python language and the PyCharm IDE. The input video stream is processed frame by frame using the OpenCV library. The system will process each frame to check for the presence of a face. If a face is found, it matches the detected face with each of the faces present in the Aadhar card photos stored in the connected database. This process is carried out using a face recognition algorithm. If a match is found, the system uses OCR to isolate and recognize the text present in the image in order to get the name and Aadhar card number of the offender, and subsequently an alert will be displayed to the corresponding moderators with the name and Aadhar card number of that individual, in order to facilitate the necessary action.

Keywords. Aadhar card, NoSQL, OpenCV, Optical Character Recognition

1. Introduction

With the rates of crime increasing all over the world, there is a strong need for a powerful and effective surveillance system. Most areas are fitted with surveillance cameras in the form of either traffic cameras or street cameras. However, these surveillance systems are not automated. They are monitored by humans. As a result, the entire process is prone to errors, as is human nature. Continuously monitoring the surveillance videos requires the utmost attention and the individuals must work expeditiously so that the necessary actions are taken in a timely manner and the delinquent individuals are apprehended. Even the slightest delay could aid the escape of a suspected individual. Therefore, human intervention hinders the overall performance of the surveillance system. As a result, there is a strong need to automate

[1] P.M.Ashokkumar, Department of Computer Science and Engineering, Koneru Lakashmaiah Education Foundation, Vaddeswaram, India. Email: profpmashok@gmail.com

the entire process. This is possible with the assistance of artificial intelligence and machine learning systems.

The proposed smart surveillance system can record the critical event, detect and even recognize the person and send an alert message to the respective authorities. Performance evaluation of systems that require continuous video analysis requires significant amounts of annotated data. The technical challenges include finding a suspicious person in crowded environments or other environments in which identifying a person could prove to be cumbersome.

The suggested system's major goal is to boost intelligence in video surveillance and thereby minimize reliance on human assistance in surveillance-related tasks. By diminishing human interactions in the system, the errors and delays associated with the mentioned interactions can be precluded and thus the performance of the overall system is improved. Additionally, the awareness of security personnel can also be enhanced. The surveillance system is implemented by firstly detecting faces from the input video using OpenCV and Haar Cascade, then face comparison is performed with the already known suspicious individuals stored in the NoSQL database. Finally, extract the respective person's Aadhar card number using OCR and an alert message is sent to the corresponding authorities so as to facilitate action.

2. Related Works

The strategy provided in this research by Kai Jin et al.,[1] seeks to highlight human identification utilizing not only overall contextual information, but also local structural information. The selective recognition algorithm based on pedestrian trajectory is used to identify the pedestrian whose motion is directed towards the installed webcam. Detection is done using YoloV3 and recognition of faces is performed using Multi-Task Cascaded CNN (MTCNN). This strategy outperforms previous systems that rely solely on facial information to accomplish recognition tasks. However, the accuracy is only 77.4 %.

The system proposed by Fu et al.,[2] uses the Eigen Face algorithm, PCA, along with the Local Binary Patterns (LBP) algorithm for face recognition, and the user interface for this system is designed with the help of PyQt. However, when the training sets are short, Eigen Face's performance suffers marginally. However, when the amount of training samples is small, the efficiency of LBP suffers dramatically. So, for training samples lacking in size, the system does not show its accurate results because the performance is not as expected as it is.

The robustness of Umara Zafar et al. [3] face recognition approach is heavily reliant on the quality of derived features and the capacity to cope with reduced face photos. Deep convolutional neural networks (DCNNs) are appealing for face identification due to their ability to learn robust features from raw face photos. However, face-alignment for 3D face data and a Bayesian deep convolutional neural network for face identification are not included in their research. As a result, this approach cannot meet its set of expectations.

Sathyavathi et al.[4] suggested a system based on both the method to detect abnormal detection in common places through the cctv videos. To decrease the face dimensional space of facial characteristics, the Eigen technique employs Principal Component Analysis (PCA). The drawback of this method is that the number of Eigen faces to be used in the Eigen Face method and the Fisher Face method is restricted due

to the PCA transformation and, as a result of that, the system did not have an accuracy of more than 90% for both manual face recognition as well as automatic face recognition.

Saibal Manna et al., [5] introduced a face recognition system that enables scanning for faces simply and fast by decreasing the time required in the procedure and to successfully help to recognize faces. Face recognition is performed using the OpenCV and FaceNet model models, which uses a deep neural network architecture to map faces from a Euclidean space. The limitation here is that the proposed system was not implemented for live camera video feed. Only pre-taken video can be used.

According to Khan et al.[6], the PCA approach is used to decrease the enormous quantity of data storage that is related to the amount of the feature set that is necessary to effectively represent the data. However, because face projection is not performed in this technique, the accuracy of the neural network classifier is lower.

The remaining sections are organized as follows. Section 3 deals with the proposed work and architecture. Section 4 is about the datasets and experimental results. In Section 5, the paper is concluded and future works are discussed.

3. Proposed Work

The architecture diagram of the proposed smart surveillance system is depicted in Figure 1. The input is in the form of live video and the extracted frames are pre-processed and then passed on to the face detection module. Then the features are extracted from the face in the form of face encodings and these are passed onto to the face comparison module, OCR module, and finally the alert module.

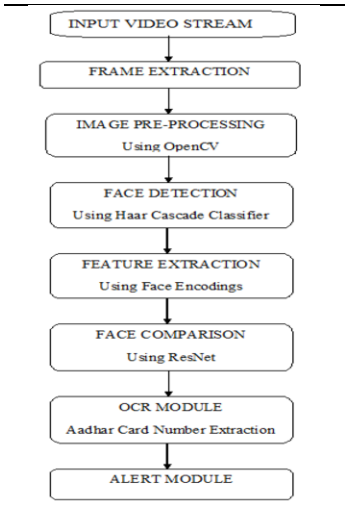

Figure 1. Smart surveillance system

The following are the steps involved in the proposed model in order to detect suspicious individuals, identify the detected individual and subsequently extract the Aadhar card number and alert the authorities on that particular identification if the

detected person is a part of the list of malfunctioned individuals stored in the NoSQL database, which is a condensation of what has been covered in the previous sections.

Table 1. ProposedSmart Surveillance Algorithm

Input: Given Training_Samples (Ts), Ground_Truth(gt)

Step1: First, the input video is dissected and individual frames are extracted for further processing
Step 2: Speed of the proposed system is of utmost importance, and hence, to avoid a performance bottleneck, every 5th frame extracted is further used for processing
Step 3:The selected frame is scrutinized for the presence of a face using the Haar Cascade classifier
Step 4: If a face is detected in the selected frame go to step 5. Otherwise go back to step 1.
Step 5:Extract all the Aadhar card images previously stored in the NoSQL database
Step 6: For every image present in the database, compare the face present in the Aadhar card to the face detected in the selected frame
Step 7: If the face detected in the selected frame matches the face present in the Aadhar card image, then go to step 8. Otherwise go back to step 1.
Step 8:Perform OCR on the matched Aadhar card image to extract the text present in the image
Step 9:Isolate the Aadhar card number from the extracted list of texts by identifying the string which consists of a 12-digit number
Step 10: Check if this particular suspected individual has been identified by the system within the last three minutes. This ensures that if an individual is continuously detected over a prolonged period of time, the respective authorities are not inundated with emails fired by the alert module
Step 11:If the individual has been identified within this given time frame, go to step 1). Otherwise got to step 12)
Step 12: Update the last seen column of the respective individual in the NoSQL database
Step 13: Send an alert email to the corresponding authorities to inform them that a particular suspected individual has been spotted. Include the Aadhar card number as well as the time at which the suspect was spotted.

Once the suspected individual has been recognized by the face comparison module, and identified by the OCR module, the corresponding authorities need to be informed about the identification of that particular individual. The Aadhar card number extracted by the OCR module is sent to the authorities as an email along with the time at which the suspect was last identified. The mail is sent with the help of the SMPT library available in Python and it is sent from a pre-defined email ID specifically created for this purpose. Since only a single camera is being used, the body of the email only consists of the Aadhar card number of the identified individual along with the time at which he or she was spotted by the camera. By including the time at which the identified individual was detected, the immediate responders can take the required action.

If multiple cameras are used, additionally, the area or location at which the suspected individual was identified can also be included so as to better facilitate the authorities to apprehend that particular suspect. If a suspect is identified for a continuous protracted period of time, back-to-back emails will be sent to the authorities, which will lead to a spamming problem.

To avoid this problem, the time at which a suspect is last seen is also stored in the MongoDB. A threshold of 3 minutes is fixed, i.e., the minimum amount of time between successive emails regarding the identification of the same individual will be 3 minutes. So, if an individual is continuously detected and identified by the system, only if the difference between the current time and the time at which he or she was last seen is greater than 3 minutes, the system will fire an email to the corresponding authorities.

If the time gap does not exceed 3 minutes, the system will detect and recognize the suspected individual, but it will not send an alert email to the authorities.

4. Experimental Results

The experiments performed to verify the performance of the proposed approach are discussed in this section. The face comparison model developed with the help of the state-of-the-art residual network provided by dlib for face recognition produced an output of 99.38 % while comparing two images containing faces. The accuracy was determined in accordance with the LFW face detection benchmark. The LFW face detection dataset consists of 1680 images of people with two or more images. The main advantage of comparing faces rather than recognizing them is that each time a new suspected individual needs to be added to the list, or an existing suspected individual needs to be removed from the list, the model need not be retrained each time. The face encodings extracted from each image encapsulate the different distinct characteristics of faces so as to be able to distinguish different people or in other words, identify different images of the same person. By comparing only the facing encodings and not performing any recognition, the model remains dynamic and flexible to any change. The comparison of the overall accuracy of the system with the existing systems in Fig 2 below:

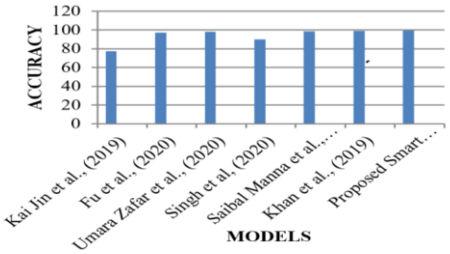

Table 2: Accuracy Comparison

Papers	Models	Accuracy
Kai Jin et al.,[1]	YoloV3 and Multi-Task Cascaded CNN	77.4 %
Fu et al., [2]	Local Binary Patterns	97 %
Umara Zafar et al., [3]	Deep Convolutional Neural Network	98.1 %
Singh et al, [4]	EigenFace method and FisherFace method	90 %
SaibalMannaet al., [5]	OpenCV and FaceNet	98.47 %
Khan et al., [6]	Principal Component Analysis and Local Binary Patterns	98.8 %
ProposedMethodology	ResNet and OCR	99.38 %

The proposed smart model is compared with various traditional modes whose accuracy is shown in Table 2. The OCR module performed with the help of Tesseract, is used to isolate the text from images so as to identify the suspected individual detected by the comparison module. An alert module is used to send an email to the corresponding authorities. This module is fired each time a suspected individual is identified in the live video stream. If the individual is repeatedly detected over a prolonged period of time, then the email is sent to the authorities every 3 minutes so as to not spam them, as each frame is scrutinized for suspected individuals. This time, indicating the time at which the suspect was last seen, is maintained and continuously updated using the NoSQL database. If the difference between the current time and last seen time is less than 3 minutes, then the alert module is not launched. The time taken to set up an SMTP connection is 5 seconds, while the time taken to send the actual email is only 1 second.

5. Conclusion

The proposed smart surveillance system is vital to today's society due to the increasing crime rates and the necessity for people to feel safe and comfortable in their social environments. The system takes the live video feed from the surveillance camera as input and it extracts every 5th frame for further processing. The selected frames are then pre-processed to make them compatible with the further modules used in the system. The frames are then checked for the presence of a single or even multiple faces. If a face is detected, feature extraction is performed by extracting the distinguishing and unique features from the face in the form of face encodings. The same feature extraction is performed on the images of the faces of the list of suspected individuals from their respective Aadhar cards which are stored in a NoSQL database. Each of the face encodings extracted from the images in the database is compared with the encodings of the detected face by the face comparison module using the dlib library. If the encodings match, then the OCR module is applied to that particular Aadhar card image in order to extract the Aadhar card number to identify the detected individual. Finally, an alert mail is sent to the immediate responders, which contains the Aadhar card number and the time at which the individual was spotted.

References

[1] K. Jin, X. Xie, F. Wang, X. Han and G. Shi, "Human Identification Recognition in Surveillance Videos," 2019 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), 2019, pp. 162-167.
[2] L. Fu and X. Shao, "Reseach and Implementation of Face Detection, Tracking and Recognition Based on Video," 2020 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS), 2020, pp. 914-917.
[3] Zafar, U., Ghafoor, M., Zia, T. et al. Face recognition with Bayesian convolutional networks for robust surveillance systems. J Image Video Proc. 2019, 10 (2019).
[4] Sathiyavathi, V., M. Jessey, K. Selvakumar, and L. SaiRamesh. "Smart Surveillance System for Abnormal Activity Detection Using CNN."Advances in Parallel Computing Technologies and Applications 40 (2021): 341..
[5] S. Manna, S. Ghildiyal and K. Bhimani, "Face Recognition from Video using Deep Learning," 2020 5th International Conference on Communication and Electronics Systems (ICCES), 2020, pp. 1101-1106.

[6] M. Khan, S. Chakraborty, R. Astya and S. Khepra, "Face Detection and Recognition Using OpenCV," 2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), 2019, pp. 116-119.

[7] A. Bharadwaj K H, Deepak, V. Ghanavanth, H. Bharadwaj R, R. Uma and G. Krishnamurthy, "Smart CCTV Surveillance System for Intrusion Detection With Live Streaming," 2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), 2018, pp. 1030-1035.

[8] A. Hampapur, L. Brown, J. Connell, S. Pankanti, A. Senior and Y. Tian, "Smart surveillance: applications, technologies and implications," Fourth International Conference on Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint, 2003, pp. 1133-1138 vol.2.