IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-4, April 2019 E-ISSN: 2347-2693

An innovative methodology for automated ATM surveillance system using skeleton-based action recognition neural networks and IoT

Jaimon Jacob[1*], Sudeep Ilayidom[2], V.P. Devassia[3]

[1]Dept. of Computer Science and Engineering, Govt. Model Engineering College, Thrikkakara, Ernakulam, Kerala, India
[2]Division of Computer Engineering, School of Engineering, Cochin University of Science and Technology, Thrikkakara, Ernakulam, Kerala,, India
[3]Govt. Model Engineering College, Thrikkakara, Ernakulam, Kerala, India

[*]*Corresponding Author: jaimonl@mec.ac.in, Tel.: +91-484-2301430*

DOI: https://doi.org/10.26438/ijcse/v7i4.949953 | Available online at: www.ijcseonline.org

Accepted: 17/Apr/2019, Published: 30/Apr/2019

Abstract— The criminal offences in the ATM kiosk are happening very commonly in recent days. A fully automated ATM surveillance system is the need of present era intended for detecting suspicious actions in the surveillance system and to trigger the proactive steps before the incident to occur. An innovative methodology proposed in this paper, which deals an automation of video surveillance in ATM kiosk and detect any type of potential criminal activities. In this system, an innovative methodology is proposed for automated ATM surveillance System using skeleton-based action recognition neural networks and IoT sensors. Multiple layers of detection techniques used to confirm the activity as suspicious. Skeleton-based action recognition by part-aware graph convolutional networks is used for detecting suspicious human action using the NTU RGB-D data set. Aadhar enabled finger print scanner which is integrated with ATM is used to fetch the demographic information from aadhar server. IoT proximity sensor is used to recognize any trial to block the vision of surveillance camera. Similarly, any physical attack made on ATM will be identified using IoT pressure/gas sensors. Suspicious sound generating during the criminal offence is also considered to confirm the activity as suspicious. Once, the activity is confirmed as suspicious, demographic information of the suspect will be fetched from aadhar server maintain by unique identification authority of India (UIDAI) and initiate the proactive steps and warning procedures.

Keywords—ATM, part-aware graph, convolutional networks , NTU RGB-D data set, Surveillance System

I. INTRODUCTION

Automated Teller Machines (ATM) today have become areas of target due to their easy and readily available cash at everyone's convenience. The attacks on ATM's are steadily rising and this is a serious problem for law enforcement and banking sectors. So there has to be a system developed and put into place that will make sure the ATM is safeguarded and also gives customers the confidence when using the ATM. There are a variety of ATM attacks because it is such an attractive target for burglars. Basically, there are three basic types of ATM attacks which can be as follows.

1) Physical attack: Brute force attack to ATM machines with intention of gaining access to cash within the safe
2) ATM Fraud: Theft of bank card information
3) Software Attack: Theft of sensitive information

An innovative approach is proposed to address the Physical attack in ATM kiosk in this work. No effective technics realized for real time recognition of such criminal offences and trigger the proactive steps before the incident to occur. In this paper, an State-of-the-art techniques are proposed for Automated ATM Surveillance System in which Skeleton-based action recognition neural networks is used for recognizing any suspicious human action by analyzing the video content and audio content. Various IoT sensors (pressure sensor, and Gas sensor) are used for recognizing any ATM Brute force attack to ATM machines with intention of gaining access to cash within the chamber and IoT proximity sensor is used for recognizing any trial to block the vision of surveillance camera. Immediately after the confirmation from the control unit, it is possible to disable the unlocking facility in the door of ATM kiosk itself. Aadhar enabled finger print scanner is also integrated with ATM to fetch the demographic information from the Aadhar server maintain by Unique Identification Authority of India (UIDAI) and can initiate the procedure to freeze further movements of the suspect. Since, aadhar is linked with the mobile number, it is easy to track and locate the suspect by using the mobile tower location, Also, the face image can be send to various public transport key stations to freeze his further schedules.

This paper is organised as follows. Section I describes an introduction, which describes an overview of work proposed in this paper. In section II, various works already done related with the area, automated ATM surveillance system. It also ensured that no similar works has been done as presented in this paper. The proposed work described in detail with the support of block diagram in Section III. Result of various partial implementations of this proposed work described in section IV. Section V describes the conclusion and future scope of this proposed system.

II. RELATED WORK

In [1], proposes a smart system which uses embedded technology using various sensors to monitor suspicious activities. It works on lively procedures to respond the robbery attempt. The sensors work to recognise suspicious activities as first stage of defence in this proposed model. Real time suspicious activity detection including Human detection, Camera tampering, Collision of Human, suspicious voice, and long term tracking is proposed in [2]. Python library functions are used to detect and recognize human faces, recognize objects, categorise human actions in videos, and finally concluding with the detection and initiate of the necessary action for the prevention of such type of activities.

PIR sensor is used to observe the suspicious movements of human in [3]. ARM controller based embedded system process real time data from PIR sensor and alarm warning messages. In [4] two separate channels are used for action detection. One channel is used as a feature extraction network and second channel as frame-wise action detection network. Finally, combine these channels to generate spatiotemporal action detections.

Estimate the human poses at each frame and train different presence models using the superpixels inside the pose bounding boxes in [5]. Since the pose estimation per frame is inherently noisy, the conditional probability of pose hypotheses at current time-step (frame) is computed using pose estimations in the current frame and their consistency with poses in the previous frames. Both the superpixel and pose-based foreground likelihoods are used to infer the location of actors at each time through a Conditional Random Field enforcing spatio-temporal smoothness in color, optical flow, motion boundaries and edges among superpixels.

In [6], a Biometric person identification systems has been designed which identify the person by determining the authenticity by their voice in multilingual environment. The speech samples are recorded in three Indian languages Hindi, Marathi and Rajasthani for multilingual environment. Pitch, formant frequencies, MFCC and GFCC feature are extracted from the speech signals. For training and testing, neural network using radial basis functions are used. In [7], Images of the banknotes will be acquired using a scanner or camera and the acquired image which will be in RGB will be converted to grayscale. The banknote will be denoised after which edge detection operation will be performed. Image segmentation operation will be performed through the PCM which will identity the features of the study banknote that are to be compared with those of the original pre-stored picture in the system. Matches in the banknotes features determine its genuinity.

III. PROPOSED METHODOLOGY

The proposed approach in this paper is a complete solution that ensure secure ATM transactions .Video content analysis, IoT sensors and biometric finger scan techniques are used in this proposed ATM surveillance system for an effective result oriented outcome. Such a system have the three major phases. 1) Identify the activity as suspicious 2) confirm the activity as suspicious 3) initiate the proactive measures to block the offence.

In this proposed model, Biometric finger scanner is integrated with ATM machine to initiate the ATM transaction. This finger scanner enable to fetch the demographic information from aadhar server when suspicious action confirmer approves there is a suspicious activity in progress. This automatically disable the facility to open the door of ATM cabin. The tower location of mobile number linked with aadhar number used to identify the location of suspect. Since, the aadhar number is linked with bank accounts, all those bank transactions can be freeze. The facial information of the suspect can be sent to various key places like railway station, police stations, bus stations and airports to block his further conveyance if he/she escaped from ATM cabin.

Various IoT sensors are also used for recognizing the suspicious activity. Pressure sensor is used for detecting the physical attack on cash storage chamber, Proximity sensor is attached along with video camera, used for detecting any trial to block the vision of video camera, Gas sensor is used to monitor changes in air quality and to detect the presence of various gases to detect any variant of gas cutters used inside the ATM cabin.

In this proposed innovative methodology, video content analysis and IoT sensors are used for identifying the activity as suspicious. In figure 1, detailed work flow of proposed work is described. Video camera used in ATM cabin generate the video output which contains video frames as well as audio information. Video summariser summarises the

video frames based on a predefined image frame distance to reduce the computational complexity by reducing the number of frames need to be processed. Following phase is used for recognising the human action whether it is suspicious using the Skeleton-based action recognition by part-aware graph convolutional networks. In [8], Skeleton-based action recognition by part-aware graph convolutional proposed, which is inspired by splitting skeleton into various parts to given as input to deep networks. The part-aware convolutions is considered to substitute common convolutions which is implemented on all the neighbouring joints. Skeleton-based action recognition by part-aware graph convolutional networks is capable of mining both temporal and spatial features from the input frames given as output of video summariser. The performance of this model is improved by exploiting the entire usage of spatial features and structures on various datasets. The major Skelton based dataset used as bench mark is NTU RGB-D.

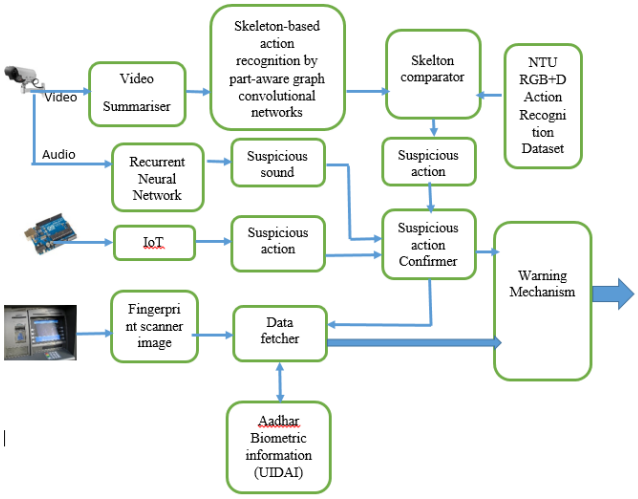

Figure 1. Architecture- automated ATM surveillance system

Figure 2 shows an example for skeleton consisting of 20 joints and 19 body-parts. The bounding boxes show regions of interest corresponding to different body-parts NTU RGB+D action recognition dataset consists of 56,880 action samples containing RGB videos, depth map sequences, 3D skeletal data, and infrared videos for each sample. This dataset is captured by 3 Microsoft Kinect v.2 cameras concurrently. The resolution of RGB videos are 1920×1080, depth maps and IR videos are all in 512×424, and 3D skeletal data contains the three dimensional locations of 25 major body joints, at each frame. The dataset contains 60 different action classes in three broad categories: daily actions, mutual actions, and medical conditions.

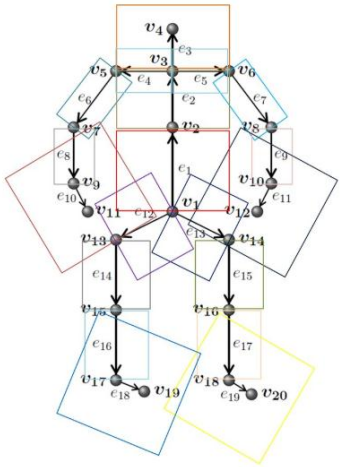

Figure 2. An example skeleton consisting of 20 joints and 19 body-parts. The bounding boxes show regions of interest corresponding to different body-parts.

Figure 3 shows the output of Skeleton-based action recognition model by part-aware graph convolutional networks. Each and every human action is represented in the form of skeletons. Suspicious activities inside the ATM kiosk are identified by the predefined dataset of Skelton models. In a normal ATM activity, the customer should be in standing position. If any other position if he/she take, it won't be take long time. The face of the suspect is not important and considered. But, the activities are closely monitored and analysed whether they are usual or unusual. Also, how much time, the suspect continue, how many suspect present in the ATM kiosk are also considered for recognising the unusual activity.

Any criminal activity normally start inside the ATM kiosk with blocking the vision of video camera. The blank output of video initiates first sign of suspicious. Similarly, any ATM transaction in this proposed approach start with authentication using biometric Finger print scanner and aadhar information. It also fails in the case of suspicious activity.

Suspicious action detection in the video frames is carried out using three major techniques. 1) Identify the action of suspect in the ATM cabin using Skeleton-based action recognition by part-aware graph convolutional networks and NTU RGB+D data set, whether he/she is focused or attacking the ATM storage chamber. 2) How long time, the suspect spare time in ATM cabin using the Euclidean distance between two frames by Video summarizer. 3) Whether the suspect used any Helmet or any variant of mask which hide the actual face, detect using the technique in [9] which proposes Faster RCNN framework by combining a number of strategies, including feature concatenation, hard negative mining, multi-scale training, model pre-training, and proper calibration of key parameters.

Audio generate during the Suspicious action is detected by another RNN and compare with the pre-defined suspicious sound models.

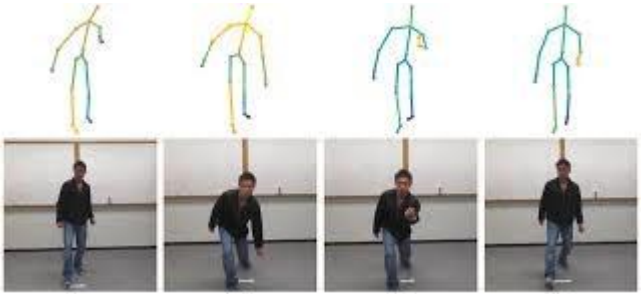

Figure 3. An example output of skeleton-based action recognition neural networks

IV. RESULTS AND DISCUSSION

The methodology presented in this work implemented and checked stage wise. The result of each segment wise implementation is awesome. Python IDE frame work with tensor flow is used for implementing skeleton-based action recognition neural networks by part-aware graph convolutional Networks for detecting suspicious human action using the NTU RGB-D data set. Recurrent Neural Network-RNN is used for audio detection to check the suspicious sound. Video content analysis using Python IDE is used for checking the presence of multiple suspect and masked face. Blank video surveillance camera output and how much time suspect is present in ATM kiosk is also checked using the Video content analysis.

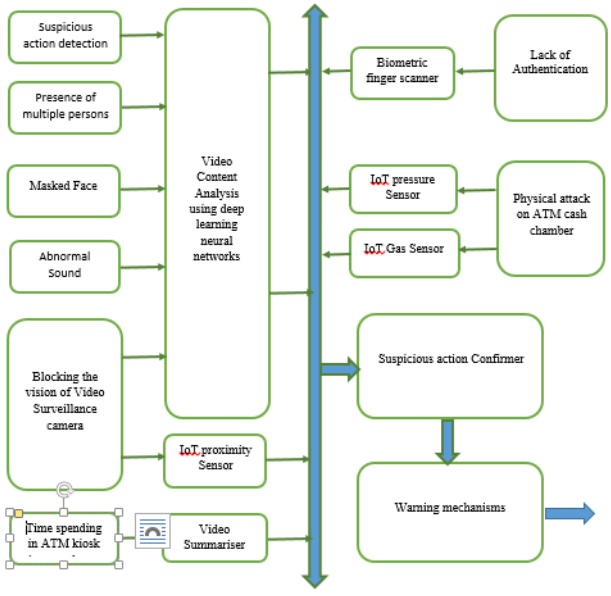

Figure 4. Summary of various suspecious activities considered and how it detected.

V. CONCLUSION AND FUTURE SCOPE

The proposed methodology in this paper is a complete solution to prevent all probable crimes in ATM cabin. Figure 4 describes various probable suspicious activities considered and how it detected using the Video content analysis and IoT sensors.

The techniques applied in this work are state of art technologies to ensure maximum accuracy. The computational complexity and time taken for confirming the activity as suspicious is reasonably high. Hence, future works can be focused on reducing the computational complexity to produce accurate results very quickly.

ACKNOWLEDGMENT

The success and final outcome of this project required a lot of guidance and assistance from many people and I am extremely privileged to have got this all along the completion of my work. All that I have done is only due to such supervision and assistance and I would not forget to thank them. I am thankful to and fortunate enough to get constant encouragement, support and guidance from all my research guides which helped me in successfully completing my research work. Also, I would like to extend our sincere esteems to all my friends and colleagues for their timely support.

REFERENCES

[1] S.Shriram, S. B.Shetty, V.P. Hegde , KCR Nisha, Dharmambal.V, "Smart ATM Surveillance System", 2016 International Conference on Circuit, Power and Computing Technologies [ICCPCT].

[2] M. Baranitharan, R. Nagarajan, G. ChandraPraba, *"Automatic Human Detection in Surveillance Camera to Avoid Theft Activities in ATM Centre using Artificial Intelligence*", International Journal of Engineering Research & Technology (IJERT)-NCICCT – 2018. Volume 6, Issue 03.

[3] K. Archana, P. B. Reddy , A. Govardhan, "To Enhance the Security for ATM with the help of Sensor and Controllers", International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS-2017).

[4] S. Aakur, D.Sawyer,S.Sarkar, *"Fine-grained Action Detection in Untrimmed Surveillance Videos*", IEEE winter applications of computer vision workshops Computer Science and Engineering, University of South Florida, Tampa, **2019**, pp-**38-40.**

[5] K. Soomro, H. Idrees, M. Shah, "Online Localization and Prediction of Actions and Interactions", IEEE Transactions on pattern analysis and machine intelligence.

[6] O.Ayankemi ONI , "*A Framework for Verifying the Authenticity of Banknote on the Automated Teller Machine (ATM) Using Possibilistic C-Means Algorithm*", International Journal of Scientific Research in Computer Science and Engineering, Vol.6, Issue.2, pp.57-63, 2018.

[7] V.K. Jain, N. Tripathi, "*Speech Features Analysis and Biometric Person Identification in Multilingual Environment*", International Journal of Scientific Research in Network Security and Communication, Vol.6, Issue.1, pp.7-11, 2018
[8] Y. Qin, L. Mo1, C. Li1, J. Luo1, *"Skeleton-based action recognition by part-aware graph convolutional networks"*, Springer-Verlag GmbH Germany, part of Springer Nature 2019.
[9] X. Suna P. WuaSteven, C.H.Hoi, *"Face detection using deep learning: An improved faster RCNN approach"*, Neurocomputing (2018).

Authors Profile

First Author- Jaimon Jacob, attained the degrees B.Tech in Computer Science and Engineering from University of Calicut in 2003, M.Tech in Digital Image processing from Anna University, Chennai in 2010, MBA in Information Technology from Sikkim Manipal University in 2012, M.Tech in Computer and Information Science from Cochin University of Science and Technology in 2014. Currently working as Asst. professor in Computer Science and Engineering, Department of Computer Science, Govt. Model Engineering College. Thrikkakara, Ernakulam, Kerala. Four International Conference papers and Two National Conference research papers published. Author passionate in research area "video processing". Associate with professional bodies ISTE,IETE and IE.

Second Author-Prof.(Dr.) Sudeep Ilayidom attained the degrees B.Tech, M.Tech, PhD. Currently Working as Professor, Division of Computer Engineering ,School of Engineering, Cochin university of Science and Technology. Ernakulam, Kerala. Published a Text book on "Data mining and warehousing" by Cengage Fifty Five research papers published in the related area Data mining. A well known musician in Malayalam Film Industry. Passionate in research area Data Mining, Big Data and related areas.

Prof.(Dr.) V.P.Devassia attained the degrees B.Sc. Engineering from MA College of Engineering, Kothamangalam, in 1983, M.Tech in Industrial Electronics from Cochin University of Science and Technology, Ph.D in Signal Processing from Cochin University of Science and Technology in 2001. Worked as Graduate Engineer(T) in Hindustan Paper Corporation Ltd, Design Engineer, HMT Limited, Principal, Govt. Model Engineering College, Ernakulam. Author passionate in research area Signal Processing. associate with professional bodies ISTE,IETE and IE.

