IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

## IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 4, April 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.488

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 4, April 2021 ||

| DOI: 10.15680/IJIRCCE.2021. 0904021 |

# An Event-Driven Crime Data Reporting Service Using AI

**SOUNDARARAJAN, Mrs. V. VIJAYAKUMARI, M.E.,**

ME Final Year, Department of Computer Science, The Kavery Engineering College, M.Kalipatti, Tamil Nadu, India

Assistant Professor, Department of Computer Science, The Kavery Engineering College, M.Kalipatti, Tamil Nadu, India

**ABSTRACT:** The project designs and implements a distributed Internet of Things (IoT) framework called ATMIoT for an intelligent, resource-efficient and real-time security management system. The system, which consists of Edge Fog computing layers, helps in crime prevention and predicts crime events in ATM centers. The ATMIoT detects and confirms criminal events in real time using artificial intelligence (AI) and an event-driven approach to send crime data to protection services and police units, which enable immediate action while at the same time conserving resources such as energy and bandwidth (BW). and memory and central processing unit (CPU) usage. The main contributions of the proposed system are: (i) a resource efficient smart edge node implementation to detect human interference and initiate fog processing; (ii) a forgeable infrastructure for the detection and confirmation of a crime; and (iii) an event-driven crime reporting service to the police station to deal with an established crime. Therefore, the proposed framework integrates image processing, AI computer vision and network communication methods to detect crime incidents in real time to ensure resource efficiency and good distribution of the processing load in an IoT-based video surveillance system. Deep Neural Network (DNN) is a popular deep learning (DL) structure made up of layered input models. The DCNN architecture that we used to train and build the classifier model. The model can predict criminal objects. The DCNN model, which runs on a fog node, recognizes and labels the images with the name of the crime objects with the highest probability and stores those images. The Fog Node then collects crime data and sends it to the nearest crime help or police unit. It also sends a real-time warning to the security service. On the other hand, the cloud is responsible for generating updated DCNN crime data models (i.e. using transfer learning methods).

**KEYWORDS:** Artificial Intelligence, Bandwidth, Deep neural network, Deep learning, multiple- layered models, Deep Convolutional Natural Network.

## I. INTRODUCTION

Criminal activities in ATMs, such as robbery, have been occurring frequently as reported in newspapers worldwide. Not only developing and underdeveloped but developed countries as well are not safe from the threat of ATM robbery. As reported in recent news, robbers are not afraid of contemporary video recording surveillance techniques because the recorded videos are used for forensic investigation of the robbery's aftermath and do not help much to prevent ATM crime.

Therefore, an instant detection of ATM crime using surveillance camera has become a demanding issue to ensure secured ATM environment. With the flourishing advances in modern technologies, video surveillance cameras integrated with image processing techniques can be used to detect criminal activities in ATM and to give a warning alarm.

Although there are numerous image processing algorithms available for occluded or covered face detection, human abnormal behavior analysis and illegal object detection, not all the conventional state of the art algorithms will work for ATMs owing to different environment (illumination, camera view, captured scene, etc.),abnormal gestures and crime devices or accessories.

Therefore, extensive research has been carried out for application of ATM crime detection using surveillance camera. Despite being a potential research area a very few traces of critical reviews has been found for the available articles on ATM crime detection using surveillance camera.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 4, April 2021 ||

| DOI: 10.15680/IJIRCCE.2021. 0904021 |

### Digital Images and Video

Digital images and video refer to 2D or 3D still and moving (time-varying) visual information, respectively. A still image is a 2D/3D spatial distribution of intensity that is constant with respect to time. A video is a 3D/4D spatio-temporal intensity pattern, i.e., a spatial- intensity pattern that varies with time. Another term commonly used for video is image sequence, Since a video is represented by a time sequence of still images (pictures). The spatio- temporal intensity pattern of this time sequence of images is ordered into a 1D analog or digital video signal as a function of time only according to a progressive or interlaced scanning convention.

### Human Visual System and Color

Video is mainly consumed by the human eye. Hence, many imaging system design choices and parameters, including spatial and temporal resolution as well as color representation, have been inspired by or selected to imitate the properties of human vision. Furthermore, digital image/video-processing operations, including filtering and compression, are generally designed and optimized according to the specifications of the human eye. In most cases, details that cannot be perceived by the human eye are regarded as irrelevant and referred to as perceptual redundancy.

### Digital Video

We have experienced a digital media revolution in the last couple of decades. TV and cinema have gone all-digital and high-definition, and most movies and some TV broadcasts are now in 3D format. High-definition digital video has landed on laptops, tablets, and cellular phones with high-quality media streaming over the Internet. Apart from the more robust form of the digital signal, the main advantage of digital representation and transmission is that they make it easier to provide a diverse range of services over the same network.

Digital video brings broadcasting, cinema, computers, video surveillance and communications industries together in a truly revolutionary manner, where telephone, cable TV, and Internet service providers have become fierce competitors. A single device can serve as a personal computer, a high-definition TV, and a videophone. We can now capture live video on a mobile device, apply digital processing on a laptop or tablet, and/or print still frames at a local printer. Other applications of digital video include medical imaging, surveillance for military and law enforcement, and intelligent highway systems.

### Spatial Resolution and Frame Rate

Digital-video systems use component color representation. Digital color cameras provide individual RGB component outputs. Component color video avoids the artifacts that result from analog composite encoding. In digital video, there is no need for blanking or sync pulses, since it is clear where a new line starts given the number of pixels per line. The horizontal and vertical resolution of digital video is related to the pixel sampling density, i.e., the number of pixels per unit distance. The number of pixels per line and the number of lines per frame is used to classify video as standard, high, or ultra-high definition, as depicted in Figure 1.4.1. In low-resolution digital video, pixilation (aliasing) artifact arises due to lack of sufficient spatial resolution. It manifests itself as jagged edges resulting from individual pixels becoming visible.

## II. RELATED WORKS

### 1. Crime Detection and Avoidance in ATM

Sujith B was told that Now a day's research is going on in the field of crime detection and avoidance in the ATM. But till now there is no advanced technology came in the field of ATM. So the idea of designing and implementation of security for ATM project are born from the observation of our real life incidents happening around us. Over the past three decades consumers have come to depend on and trust the ATM to conveniently meet their banking needs. In recent years there has been a proliferation of ATM frauds across the globe. The suspicious action in ATM is many. The suspicious object's visual properties so that it can be accurately segmented from videos. After analyzing its subsequent motion features, different abnormal events like Crimes and robbery can be effectively detected through videos. This paper deals comparison existing technologies and propose a new frame work for ATM security in cooperation ATM software. The proposed method will uses multiple object detection method and event recognition techniques of computer vision.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 4, April 2021 ||

| DOI: 10.15680/IJIRCCE.2021. 0904021 |

**2. Surveillance Video Based Robust Detection and Notification of Real Time Suspicious Activities in Indoor Scenarios**

Nithya Shree R was said Over recent years, surveillance camera is attracting attention due to its wide range of applications in suspicious activity detection. Current surveillance system focuses on analyzing past incidents. This paper proposes an intelligent system for real-time monitoring with added functionality of anticipating the outcome through various Image processing techniques. As this is a sensitive matter, human decisions are given priority, still facilitating limited logical intervention of human resource. This framework detects risk in the area under surveillance. One such dangerous circumstance is implemented, like a person with a knife. Here the prediction is that in the firm places like ATM, Banks, Offices etc. a person possessing knife is unusual and likely to cause harmful activities like threatening, injuring and stabbing. The experiment demonstrates the effectiveness of the technique on training dataset collected from distinct environments. An interface is developed to notify concerned authority that boosts reliability and over all accuracy.

**3. Video-Based Tracking, Learning and Recognition Method for Multiple Moving Objects**

HidetomoSakaino was said Proposes a cost lessening strategy for the MCMC approach by taking moves, i.e., birth and demise, out of the emphasis circle of the Markov chain when diverse moving items collaborate. The GUI proposed thus offers an auto-portion module of images from pictures and a hand- drawing module for productive direction learning and for premium direction expansion.

**4. Comparative Analysis of Various Feature Descriptors for Efficient ATM Surveillance Framework**

Zhiqian Chen said that Show relative examination of various descriptors like Gradient based descriptor i.e. HOG(Histogram of Oriented Gradients) and shape based descriptors like Hu Moments and Zernike Moments for powerful anomalous movement discovery.HOG is in charge of pulling back shape data of question in picture utilizing force angles and edge headings. Zernike minutes are orthogonal minutes which are compelling in picture portrayal. These are revolution invariant. Zernike minute developed from Zernike polynomials which are orthogonally autonomous and subsequently the picture portrayal doesn’t experiences covering or repetition.

Hue minutes give seven qualities as a removed component from a given picture. These minutes are invariant to interpretation, scale, and turn of a picture. Out of seven invariants, six are supreme orthogonal invariants and the seventh one is skew orthogonal invariant. Hoard gives better exactness then different descriptors. As for shape descriptors Hue minutes give better outcomes in video when contrasted with Zernike minute.

**5. Realtime 3D Skeletonisation in Computer Vision-Based Human Pose Estimation Using GPGPU**

Rune HavnungBakken and Lars MolandEliassen was said the proposed a calculation yields predominant exactness and vigor when utilized as a part of a stance estimation setting. General purpose computing on graphics processing units (GPGPU) that is appropriate for use in human stance estimation, and accomplishes constant execution. A diminishing calculation iteratively expels limit voxels from a question create a topologically identical skeleton. The proposed calculation ought to be contrasted with other cutting edge skeletonisation calculations. Exhibited ongoing diagram based human posture estimation approach, utilizing skeletons to recoup the substance. The skeletons were created with a strategy called voxel scooping.

## III. PROPOSED SYSTEM

Proposed Bank/ATM IoT-guard. With the help of AI, each fog node can detect and identify a possible crime event and crime object by processing the motion-captured images sent by an edge node. DCNN model running at a fog node detects and labels the images with the name of the crime objects having the highest probability, and saves those images. The crime unit receives the crime image, alert message, and crime location information. The alert message consists of the crime data and location, and the labeled-image verifies and confirms the crime weapon. Finally, the location information tells the police where the crime might occur so they can take necessary steps to prevent it.

## IV. SYSTEM DESIGN

The system is far more efficient even without the video compression algorithm. Proved the superiority of our proposed system in terms of agility, scalability, energy, and CPU and memory usage. Efficient crime predictive system. Real-time crime event detection, ensuring resource efficiency and good distribution of the processing load in an IoT-based video surveillance system.

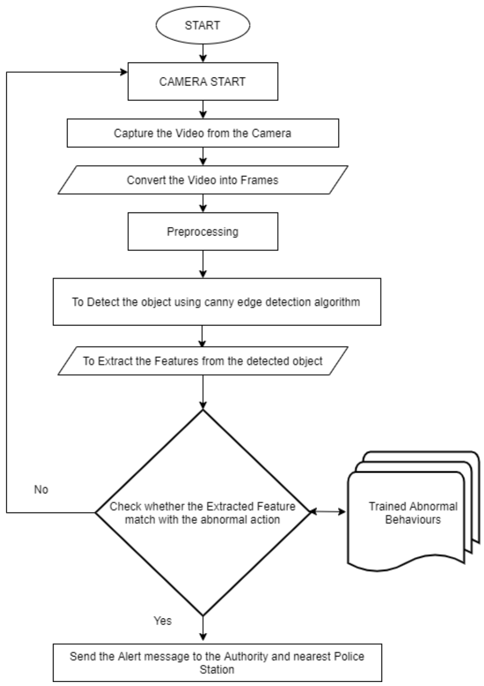


**Figure 4.1: System Architecture**

## V. METHODOLOGY

### 5.1 Video Surveillance Data Set Annotated

The videos provided to the algorithm presented have the minimum resolution of 320 x 240, which are recorded in the indoor and outdoor environment in ATM Center. Videos captured are analyzed under the four categories: (i) single: when single person is in the video and performing normal activities. (ii) Single abnormal: when single person is in the video and performs abnormal activity. (iii) Multiple: when multiple people are in the camera view and act normally. (iv)Multiple abnormal: when multiple people are in the view and abnormal activities are performed.

### 5.2 Testing Phase

Live Video Data Set Annotated

- Cameras should be deployed in critical areas to capture relevant video.
- Computer and camera are interfaced and here webcam is used.

### 5.3 Frame Extraction

Frames are extracted from the video input. The video has to be broken down into a sequence of images that are further processed. The speed at which a video needs to be broken down into images depends on the implementation of each person. We can say that mostly 20 to 30 frames per second are taken, which are sent to the next phases.

### 5.4 Preprocessing

By improving the various characteristics of an image, we obtain, for example, intensity, contrast and saturation for various types of image processing. Low pass filter a grayscale image that has been affected by constant power additive

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 4, April 2021 ||

| DOI: 10.15680/IJIRCCE.2021. 0904021 |

noise. It uses a pixel-by-pixel adaptive Wiener method based on statistics estimated from a local neighborhood of each pixel.

**5.5 Action Detection**

The background subtraction approach is mostly used when the background is static. The principle of this method is to use a model of the background and compare the current image with a reference. The foreground objects present in the scene are recognized. An attempt is made to recognize moving areas in an image by distinguishing pixel by pixel between the current image and a reference background image. We're going to use the static background for the image subtraction that gives us the people we need to chase. This step detects objects of interest as they move in the scene. The action detection process is applied independently to all static cameras present in the scene. For human recognition, feature extraction and representation applies, in which the important features of picture frames are extracted and systematically represented as features.

**5.6 Feature Engineering**

Deep neural network (DNN) is a popular deep learning (DL) structure that consists of multiple-layered models of inputs. The DCNN architecture that we used to train and build the classifier model. The model can predict crime objects. Hence, the DCNN model running at a fog node detects and labels the images with the name of the crime objects having the highest probability, and saves those images.

**5.7 IoT Integration**

IEEE 802.11(WLAN) as the physical layer protocol, IPv4 as the network layer protocol, and, finally, MQTT as the application layer protocol for the proposed system. The MQTT client publishes multimedia data (text, image, and video) through a specific topic to the broker/server, and, then, the broker forwards the data to the clients who have subscribed to that topic.

**5.8 Crime Prevention Unit**

The crime unit receives the crime image, alert message, and crime location information. The alert message consists of the crime data and location, and the labeled-image verifies and confirms the crime weapon. Finally, the location information tells the police where the crime might occur so they can take necessary steps to prevent it.

## VI. CONCLUSION

Human activity detection is a challenging problem in many applications in areas such as visual surveillance, human-computer interaction, autonomous driving, and entertainment. There are many possible approaches to motion estimation to solve this problem. In this study it is proposed to construct a hybrid depth model for the purpose of HAR. The proposed architecture is based on a dense optical flow approach and additional motion information in videos using deep learning methods. First, deep learning models, namely 3D convolutional neural network (3D-CNN), 3D-CNN with optical flow and long-term short-term memory network (LSTM), are combined to determine the motion vectors. The video classification task is then processed by the support vector machine algorithm. A wide range of comparative experiments will be performed on two newly generated chess data sets, namely the Magnetic Wall Chessboard Video Data Set (MCDS) and the Standard Chessboard Video Data Set (CDS) to demonstrate the contributions of the proposed study. Finally, the experimental results show that the proposed hybrid depth model shows considerable classification success compared to studies according to the prior art. In addition, to the best of our knowledge, this is the first study based on a novel combination of 3D-CNN, 3D-CNN with optical flow, and LSTM over video images to detect human activity. In summary, the experimental results show that the proposed architecture has significant advantages for recognizing and classifying human activity in videos. First, the proposed hybrid depth model is flexible, extensible and customizable in that it can determine many complex activities in different video data sets, including playing chess, playing checkers, playing solitaire and playing cards. Second, any number of features can easily be consolidated as support information for the proposed architecture. In addition to these advantages, the proposed hybrid deep model architecture allows other deep learning models to be combined with our proposed model as auxiliary features for purposes such as object recognition, hand tracking, etc.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 4, April 2021 ||

| DOI: 10.15680/IJIRCCE.2021. 0904021 |

## VII. FUTURE ENHANCEMENT

As a future work, we plan to expand the functions through the use of other deep learning techniques. In addition, heuristic optimization algorithms can be used to extend the functionality and improve the classification performance of HAR.

## REFERENCES

[1] Sujith B, “Crime detection and avoidance in ATM: a new framework,” International Journal of Computer Science and Information Technologies,2014.
[2] Nithya Shree R, RajeshwariSah and Shreyank N Gowda, “Surveillance video based robust detection and notification of real time suspicious activities in indoor scenarios,” The sixth international conference on computer science, engineering and information technology,2016.
[3] HidetomoSakaino, “Video-Based Tracking, Learning and Recognition Method for Multiple Moving Objects,” IEEE Transactions on Circuits and Systems for Video Technology, VOL. 23, NO. 10, OCTOBER2013.
[4] Zhiqian Chen, Vishal Sanserwal, VikasTripathi and Monika Pandey, “Comparative Analysis of Various Feature Descriptors for Efficient ATM Surveillance Framework,” IEEE International Conference on Computing, Communication and Automation (ICCCA)2017.
[5] Rune HavnungBakken and Lars MolandEliassen, “Realtime 3D Skeletonisation in Computer Vision-Based Human Pose Estimation Using GPGPU,” IEEE Conference on Image Processing Theory, Tools and Applications, 2012.
[6] HaithamHasan and S. Abdul Kareem, “Human Computer Interaction for Vision Based Hand Gesture Recognition: A Survey.” IEEE International Conference on Advanced Computer Science Applications and Technologies,2013.
[7] Rishabh Agrawal and Nikita Gupta, “Real Time Hand Gesture Recognition for Human Computer Interaction,” IEEE 6th International Conference on Advanced Computing,2016.
[8] Hong Cheng, Lu Yang and Zicheng Liu, “Survey on 3D Hand Gesture Recognition,” IEEE Transactions on Circuits and Systems for Video Technology, VOL. 26, NO. 9, SEPTEMBER2016.
[9] Miwa takai, “Detection of suspicious activity and estimate of risk from human behaviour shot by surveillance camera” Nature and Biologically Inspired Computing (NaBIC), 2010 Second World Congress, IEEE2011.
[10] Sambarta Ray, Souvik Das and Dr. AnindyaSen, “An Intelligent Vision System for monitoring Security and Surveillance of ATM” IndiaConference.