



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 7, July 2023

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

Survey of Crime Activity Detection System

Prof. Prateek Meshram¹, Richa Kushwaha², Sapana Pande³, Shamali Deshmukh⁴, Priyanka Tatu⁵

Associate Professor, Department of Computer Engineering, DYPIEMR, Maharashtra, India¹

U.G. Student, Department of Computer Engineering, DYPIEMR, Maharashtra, India²

U.G. Student, Department of Computer Engineering, DYPIEMR, Maharashtra, India³

U.G. Student, Department of Computer Engineering, DYPIEMR, Maharashtra, India⁴

U.G. Student, Department of Computer Engineering, DYPIEMR, Maharashtra, India⁵

ABSTRACT: Video Surveillance plays a vital role in today's world. The technologies are been advanced too much when artificial intelligence, machine learning and deep learning used into the system. Using all such above combinations, different systems are placed which helps to differentiate various crime behaviours from the live tracking of footages. The most unpredictable one from among all is human behaviour and it is very difficult to find whether it is detected crime or not. Deep learning technology approach is used to detect crime or normal activity in an academic environment, and which sends an alert message to the corresponding authority, in case of predicting a crime activity. Monitoring is often performed through consecutive frames which are extracted from the video.

KEYWORDS: Deep Learning, Crime detection, CNN, image processing, etc

I. INTRODUCTION

In most nations, it is the job of the police to find criminal activity [1], while specialized law enforcement organizations may be tasked with finding certain kinds of criminal activity. The recognition of a crime has been committed, the identification of a suspect, and the gathering of adequate evidence to bring the suspect before a court are the three distinct stages of crime detection. Other than the police, many crimes are detected and reported. Certain crimes, especially those that require a subject's consent, like the sale of illegal drugs or prostitution, or those in which there may be no clearly defined victim, like obscenity, are frequently not identified until the police take any proactive measures to ascertain whether they have been committed. To detect such crimes, therefore, controversial methods are sometimes required (e.g., electronic eavesdropping, surveillance, interception of communications, and infiltration of gangs) [4].

With an increase in crime rates, it becomes a problem if they are not identified in time and necessary precautionary actions taken. Most urban and metropolitan areas have surveillance systems installed which constantly collects a large amount of data. With the vast accumulation of surveillance [1] data present, there are higher chances of suspicious activities that occur. But these tasks also require human supervision to detect such activities as they are too much complicated for artificial intelligence [3] to handle and require high resources. Breaking down complicated tasks and detecting sub tasks which leads to potential crimes is one of the way to simplify an activity to be automated [6]. We focus on two main potential leads to crimes which we attempt to detect through our models.

Humans have the capability to detect and distinguish objects and actions into innumerable categories. There certain thresholds beyond which even the human eye cannot detect differences in images. Therefore, object detection [14] using machines or computers is also an important factor. It is a tough task because the computer cannot find out the difference in images machines or computers is important. It is a most challenging task because the computer cannot find out the difference in images just by looking at them. Various algorithms can be initialising to do this. The machine splits the video into number of frames and then these frames are analysed to find out the motion in the video.

The system we created can identify crimes that are taking place in the real world. This technique aids in avoiding and protecting the public. Crimes can be discovered using face detection, object detection (such as a pistol or Knife) [14], and fire detection [15] Scope of project is the model can be improving by identifying the crime individual from the criminal activity.

A deep learning algorithm [3] which takes an input image, assign importance (learnable weights and biases) to various aspects/objects in the image, and be able to differentiate them one from the other. The pre-processing required

in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are handengineered, with enough training, ConvNets [6] have the ability to learn these filters/characteristics.

II. MOTIVATION

1. Knowing the importance of suspicious human activities. Recognizing it from video surveillances is to prevent the crime cases.
2. To prevent the accident in advance and if accident occurs then there will be a prevention system to notify emergency services before time.

III. OBJECTIVES

1. To design an intelligent algorithm for automatic classification of videos of single person to multi people and hence detect strange and crime activities.
2. To extend the approach by developing algorithms for detecting real-time changes in the scene and to trigger an alarm in real time, if the identified activity is crime case.
3. In this paper, we have compared different types of algorithms, for determining which algorithm suits best for crime activity detection.

IV. LITERATURE WORK

[1.] The detection of violent activities is done using Optical Flow Algorithms which has been discussed [1]. The main task in violence detection is to detect moving objects and classify their behavior using means of intensity as violence or non-violence. In this, an approach is implemented to detect any violent behaviour in videos using optical flow algorithms. Many algorithms exist to analyze the behavior of the moving objects. In this paper, an approach to detect any violent behavior is done through videos using optical flow algorithms. The SDHA 2000 dataset is used to analyze foreground behavior. The main task in violence detection is to detect moving objects and classify their behavior using mean of intensity as violent or non-violent.

[2.] With an increase in crime activities in urban and suburban areas, it is necessary to detect them to reduce and minimize such events. In earlier days surveillance was done manually by humans and was a tiring task as suspicious activities were uncommonly compared to the usual activities. With the arrival of intelligent surveillance systems, various approaches were introduced in surveillance. We mostly focus on analyzing on two cases, if those are ignored then it could lead to high risk of human lives, which detect potential gun-based crimes and detecting abandoned luggage on frames of surveillance footage. Presented model [2] can detect handguns in images and a machine learning and computer vision pipeline that detects abandoned luggage so that we could identify potential gunbased crime and abandoned luggage situations in surveillance footage

[3.] This paper proposes Secure-Pose, a novel cross-modal forgery detection and localization system for live surveillance videos using the help of WiFi signals near the camera spot. This application solves the problem by eliminating the need to store the videos itself. Instead, a better solution is to store only the accurate captions of the events happening in the video along with the respective timestamp. Crimes are detected based on the captions [3.] This paper proposes Secure-Pose, a novel cross-modal forgery detection and localization system for live surveillance videos using the help of WiFi signals near the camera spot. This application solves the problem by eliminating the need to store the videos itself. Instead, a better solution is to store only the accurate captions of the events happening in the video along with the respective timestamp. Crimes are detected based on the captions

[4.] It is quite challenging to detect the anomaly in crowded scenes for quite a long time. In this paper, a self-supervised framework, abnormal event detection network (AED-Net) [13], is composed of PCAnet and kernel principal component analysis (kPCA), is proposed to address this problem. Using surveillance video sequences of different scenes as raw data, PCAnet is trained to extract high-level semantics of crowd's situation. Next, kPCA, a one-class classifier, is trained to determine anomaly of the scene. In contrast to some prevailing deep learning methods, the framework is completely self-supervised because it utilizes only video sequences in a normal situation.

[5.] There is a huge growth in the use of surveillance cameras to maintain a log of events that would help in the identifying criminal activities. However, it is necessary to continuously monitor the acquired footage which contributes to increased labor costs but more importantly, it causes the violation of privacy. Therefore, we need decentralized surveillance systems that function autonomously in real-time to reduce crime rates even further. In our work, we discuss an efficient method of crime detection using Deep Learning [14], that can be used for on device crime monitoring. By making the inferences on-device, we can reduce the latency, the cost overhead for the collection of data into a centralized unit and counteract the lack of privacy.

[6.] Crowd analysis has become an extremely famous research point in the territory of computer vision. Computerized examination of group exercises utilizing reconnaissance recordings is a significant issue for public security since it permits the identification of hazardous groups and where they're going. We all can see how many problems are faced because of the crowd. This causes a planning for an terrorist attack. They place a bomb in a such an area which causes injuries. Thieves are mostly getting an advantage of the crowd. In that situation, crowd analysis is very important. This paper shows the design of the deep learning architecture that provides control over the crowd behavior which will help to avoid violence or any other act which occurs because of the crowd which causes harmful effects to the society. So, we are proposing a system that detects abnormal behavior of crowds using deep learning techniques.

V. ALGORITHM

A.) Optimal Flow Algorithm [1]:

- 1) Get input as video from database.
- 2) Extract frames from video.
- 3) Convert frames from RGB to HSV.
- 4) Select the values of intensity (V).
- 5) Calculate the mean of mean of magnitude.
- 6) Apply Gunnar Farneback Algorithm.
- 7) If mean of mean of magnitude is greater than threshold go to step 8 otherwise go to step 9.
- 8) Display activity as violent.
- 9) Display activity as normal.
- 10) End.

B.) Faster R-CNN for Suspicious Activity Detection:

Faster R-CNN [2] is used as an object detection algorithm which is similar to r-cnn. This algorithm utilises the region proposal network (rpn) that shares the full-image convolutional features with the detection network in a costeffective manner than r-cnn and fast r-cnn. A region proposal network (rpn) is basically a fully convolutional network that simultaneously predicts the object that bounds as well as objectness scores at each position of the object and is trained end-to-end to generate high-quality region proposals, which are then used by fast r-cnn for detection of objects [2].

C.) Feature Extraction

The following are the steps through which this project has been implemented with the help of known as system Real Time Crime Detection by Captioning Video Surveillance Using Deep Learning [3].

1. Firstly, perform Compression of video files.
2. Then extracting frames from videos and storing them as images.
3. Captioning of images to identify the threats like weapons, fire, violence etc.



4. Search for the features like timeframe, keywords etc.
5. Lastly the data encryption of the generated caption file.

D.] K-Means:

K-Means clustering algorithm [12] is an independent algorithm which is utilized in the section for intrigue zone from the foundation. It groups, or segments the given information into K-clusterd or parts dependent on the Kcentroids. This K-means algorithm is used when unlabelled data present in it (i.e., information without characterized classes or gatherings). The objective is to discover certain gatherings dependent on a closeness in the information with the quantity of gatherings group to by K.

A K-Means clustering algorithm is mostly used for motion detection of the objects which are detected by CNN. We created the cluster of objects pixels. Then we calculate the difference between the position of that object and then calculated the Euclidean distance on the bases of that we set the threshold value for detecting abnormal behavior of the crowd.

Knn is used to calculate the points distance between two image frames in sequence to track the motion of moving objects.

E.] AED-Net:

AED-Net [13] Intuitively, the anomaly detection task in our proposed AED-Net is to assign a score indicating abnormality to each frame of video. And during a training phase, the largest reconstruction error should be set as threshold for the anomaly detection. Thus, in testing phase, the abnormality of test frames can be determined by comparing the score of test frames to the threshold. To fulfill this task, we incorporate PCAnet and kPCA together and build AED-Net.

Alg.1 AED-Net

Input: optical flow maps $J = \{J11, J22, \dots, JN\}$

Output: Threshold of max normality score *threshold*, $M1$, $M22$, a

1. for $i = 1, 2, \dots, N$ do
2. Sample Jii by patches and get $XXii$ by vectorizing and concatenating all patches
3. end for
4. $XX \diamond = XX -$
 1
 $(\sum XX)$
 $u ii N ii$
5. $S1 = [X1, X$
 $2, \dots, XN$
 $]$
6. $M1 = \operatorname{argmin} \| S1 - M1$
 $(M1$
 $)$
 $TS1$
 $\| s. t. M1$
 $(M1$
 $)$
 $T = JL$
 $M1$
 1
7. $D = J * M1$
 $, D = \diamond D$
 $1 \diamond ii = 1, 2, \dots, N. J = 1, 2, \dots, L1$
 i
8. for $i = 1, 2, \dots, N, jj = 1, 2, \dots, L1$ do



9. Sample D_{jj} by patches and get Y_{jj} by vectorizing and concatenating all patches

ii

10. end for

11. \bar{Y}

y

$= Y_{jj} - \text{mean}(Y_{jj})$

u

12. $S = \bar{Y}^T$

, ..., \bar{Y}

$L1$

, \bar{Y}^T

, ..., \bar{Y}

$L1$

, ..., \bar{Y}^T

, ..., \bar{Y}

$L1$ $\diamond 2 \ 1 \ 1 \ 2 \ 2 \ N \ N$

13. $M2 = \text{argmin} \| S2 - M2$

($M2$

)

$TS2$

$\| s. t. M2$

($M2$

)

$T = IL$

$K2$

2

14. $Ol, = Dl * M2$

, $l = 1, 2, \dots, L, m = 1, 2, \dots, L, ii =$

$1, 2, \dots, N$

$ii \ m \ 1 \ 2$

15. $Tl = \sum L2$

$2m-1H \diamond Ol, \diamond ii \ m=1 \ ii$

16. $\mathcal{F} = Bhiist \diamond T$

$1 \diamond, \dots, Bhiist \diamond TL1 \diamond$

T

$ii \ ii \ ii \ \diamond$

17. for $i = 1, 2, \dots, N$, do

18. for $jj = 1, 2, \dots, N$, do

19. $VI, j = \kappa \diamond \mathcal{F}i$

, $\mathcal{F}jj \ \diamond$

20. end for

21. end for

22. $\bar{V} = V -$

1

$\sum N \ V -$

1

$\sum N \ V +$

1 $\sum N \ V$

$ijj \ ijj \ N \ a=1 \ iia \ N \ a=1 \ ajj \ N2 \ a,=1 \ ab$

23. $\alpha = \text{argmin} \| \bar{V} - \alpha \alpha T \bar{V} \| \ s. t. \ \alpha \alpha T = Iq$

α

24. Initialize threshold=0.

25. for $i = 1, 2, \dots, N$ do



```

2
26. (F) = 1 -
2
Σ N M +
1 Σ N M - Σ q
◆ P
(F)◆, and
ii N a=1 za N2 a,=1 ab jj=1
jj ii
27. P (F) = Σ N ajj ◆ M -
1
Σ N M -
1
Σ N M +
1
Σ N M ◆
jj ii ii=1 ii zii N a=1 iia N a=1 zb
N2 a,=1 ab
28. if R(Fii) > threshold, do
29. threshold = (Fii)
30. end if
31. end for
    
```

F.] Convolution Neural Network Algorithm

- Convolutional Neural Networks [16] is specially used for applications in image & video recognition. CNN is mainly used in image analysis tasks like Image recognition, Object detection & Segmentation [21].
- There are Four types of layers in Convolutional Neural Networks:
 - 1) Convolutional Layer: - In a typical neural network each input neuron is connected to the next hidden layer. In CNN, only a small region of the input layer neurons connects to the neuron hidden layer.
 - 2) Pooling Layer: - The pooling layer is used in reducing the dimensionality of the feature map. There will be multiple activation & pooling layers inside the hidden layer of the CNN.
 - 3) Flatten: - Flattening is used in converting the data into a 1-dimensional array for inputting it to the next layer. We flatten the output of the convolutional layers to create a single long feature vector.
 - 4) Fully-Connected layer: - Fully Connected Layers form the last few layers in the network. The input to the fully connected layer is the output from the final Pooling or Convolutional Layer, which is flattened and then fed into the fully connected layer.

CNN implementation steps: -

- Step1(a): Convolution Operation
(Filter image)
- Step 1(b): ReLU Layer
- Step 2: Pooling (used max pooling function)
- Step 3: Flattening (Covert Matrix into 1D Array)
- Step 4(a): Full Connection.
- Step 4(b): Dense()
- Step 4(c): Optimizer()
- Step 4(d): compile()

In this paper, we have compared many such algorithms and now we have come across the conclusion, that for crime activity detection CNN algorithm suits best.

VI. CONCLUSION

In this paper we have represented the working principle of CNN and how it is better from all other algorithms. In today's world, almost all the people know the awareness and importance of CCTV footages, but mostly in these cases the footages are being used for the investigation purposes after a crime/incident which have already been happened. The proposed model has the benefit of stopping the crime before it happens. The real time CCTV footages are being tracked and analysed at the same time. The result of the analysis is a command to the respective authority to take an action if in case the result indicates an untoward incident is going to happen or is hapenning. Hence this can be stopped. The advantage of paper using the Faster R-CNN is it can specifically detect gun-based crime and abandoned luggage but disadvantage is detecting abandoned luggage does not address issues like identification of objects in sudden changes of illumination.

REFERENCES

- [1] Prasad. D. Garje , Dr. M. S. Nagmode , Kiran. C. Davakhar , “Optical Flow Based Violence Detection in Video Surveillance ” , 2018 International Conference On Advances in Communication and Computing Technology (ICACCT) Amrutvahini College of Engineering, Sangamner, Ahmednagar, India. Feb 8-9, 2018.
- [2] Sathyajit Loganathan, Gayashan Kariyawasam, Prasanna Sumathipala , “Suspicious Activity Detection in Surveillance Footage”, 2019 International Conference on Electrical and Computing Technologies and Applications (ICECTA).
- [3] Nagesh Nayak, Shlesha Odhekar, Sapna Patwa, Sukanya Roychowdhury, “Real Time Crime Detection By Captioning Video Surveillance Using Deep Learning” , © 2022 IJCRT , Volume 10, Issue 7 July 2022,ISSN: 2320-2882.
- [4] Alexandrie, G. (2017). Surveillance cameras and crime:a review of randomized and natural experiments. *Journal of Scandinavian Studies in Criminology and Crime Prevention*, 18(2):210–222.
- [5] Tay, N. C., Connie, T., Ong, T. S., Goh, K. O. M., and Teh, P. S. (2019). A robust abnormal behavior detection method using convolutional neural network. In Al-fred, R., Lim, Y., Ibrahim, A. A. A., and Anthony, P., editors, *Computational Science and Technology*, pages 37–47, Singapore. Springer Singapore.
- [6] Tran, D., Bourdev, L., Fergus, R., Torresani, L., and Paluri, M. (2015). Learning spatiotemporal features with 3d convolutional networks. In *Proceedings of the IEEE international conference on computer vision*, pages 4489–4497.
- [7] Zhu, Y. and Newsam, S. (2019). Motion-aware feature for improved video anomaly detection.
- [8] P.Bhagya Divya, S.Shalini, R.Deepa, Baddeli Sravya Reddy, “Inspection of sus- picious human activity in the crowdsourced areas captured in surveillance cameras”,*International Research Journal of Engineering and Technology (IRJET)*, De- cember 2017.
- [9] Jitendra Musale,Akshata Gavhane, Liyakat Shaikh, Pournima Hagwane, Sneha- lata Tadge, “Suspicious Movement Detection and Tracking of Human Behavior and Object with Fire Detection using A Closed Circuit TV (CCTV) cameras ”, *Interna- tional Journal for Research in Applied Science Engineering Technology (IJRASET)* Volume 5 Issue XII December 2017.
- [10] U.M.Kamthe,C.G.Patil “Suspicious Activity Recognition in Video Surveillance System”, Fourth International Conference on Computing Communication Control and Automation (ICCCUBEA), 2018.
- [11] Zahraa Kain, Abir Youness, Ismail El Sayad, Samih Abdul-Nabi, Hussein Kassem, “ Detecting Abnormal Events in University Areas ”, *International conference on Computer and Application*,2018.
- [12] Riddhi Sonkar, Sadhana Rathod, Renuka Jadhav, Deepali Patil, ” CROWD ABNORMAL BEHAVIOUR DETECTION USING DEEP LEARNING”, *ITM Web of Conferences* 32, 03040 (2020) <https://doi.org/10.1051/itmconf/20203203040> ICACC-2020.
- [13] Tian Wang a, Zichen Miao a, Yuxin Chena, Yi Zhou, Guangcun Shan*c, Hichem Snoussid, ” AED-Net: An Abnormal Event Detection Network ” , <https://arxiv.org/ftp/arxiv/papers/1903/1903.11891> School of Instrumentation Science and Opto-electronic Engineering & International Research Institute for Multidisciplinary Science, Beihang University, Beijing 100191, China 2022.
- [14] Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun Faster r-cnn: Towards real- time object detection with region proposal networks *Advances in neural information processing systems*, 2015.
- [15] F. Yuan, “An integrated fire detection and suppression system based on widely available video surveillance,” *Mach. Vis. and Apps.*, vol. 21, pp. 941–948, 2010.
- [16] Y. Luo, L. Zhao, P. Liu, and D. Huang, “Fire smoke detection algorithm based on motion characteristic and convolutional neural networks,” *Multimedia Tools and Applications*, pp. 1–18, 2017.



- [17] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," in arXiv, vol. abs/1512.00567, 2015.
- [18] S. Kwak, G. Bae and H. Byun, "Abandoned luggage detection using a finite state," Optical Engineering, vol. 49, 2010.
- [19] S. Guler and M. K. Farrow, "Abandoned Object Detection in Crowded Places," 9th IEEE International Workshop on PETS, 2006.
- [20] L. Li, R. Luo, R. Ma, W. Huang and K. Leman, "Evaluation of An IVS System for Abandoned Object Detection," 9th IEEE International Workshop on PETS, 2006.
- [21] L. A. Lim and H. Y. Keles, "Foreground Segmentation Using a Triplet Convolutional Neural Network for Multiscale Feature Encoding".
- [22] Kumar, A., Goyal, S., and Varma, M. (2017). Resource- efficient machine learning in 2 kb ram for the internet of things. In Proceedings of the 34th International Conference on Machine Learning Volume 70, pages 1935–1944. JMLR. Org
- [23]] Kusupati, A., Singh, M., Bhatia, K., Kumar, A., Jain, P., and Varma, M. (2018). Fastgrnn: A fast, accurate, stable and tiny kilobyte sized gated recurrent neural network. In Advances in Neural Information Processin Systems, pages 9017–9028.



INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.423



ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com



www.ijirset.com

Scan to save the contact details