

# ABNORMAL EVENT DETECTION BY MACHINE VISION USING DEEP LEARNING

Akshara Alex  
Department of CSE  
Acropolis Institute of Technology & Research  
Indore, M.P., India

Avni Tanwar  
Department of CSE  
Acropolis Institute of Technology & Research  
Indore, M.P., India

Kavita Namdev  
Department of CSE  
Acropolis Institute of Technology & Research  
Indore, M.P., India

Ashi Sahu  
Department of CSE  
Acropolis Institute of Technology & Research  
Indore, M.P., India

Nisha Rathi  
Department of CSE  
Acropolis Institute of Technology & Research  
Indore, M.P., India

**Abstract**— This paper proposes an abnormal event detection system through surveillance camera using machine vision, in corporation with deep learning, which analyzes footages of crowded scenes and detects abnormal events. In the proposed solution, the concept of basic deep neural network model has been widely adopted. DNN is a mathematical model that categorizes video input frames and separates the anomaly present in the video events. If any one of the anomaly namely arrest, assault and abuse is occurred in a video, that anomaly is detected from specified frames of video. The proposed model mainly depicts how deep learning elevates the quality of machine vision in abnormal event.

**Keywords**— Abnormal Event detection, Convolution Neural Network, Deep Learning, Framing, Machine Vision Modeling

## I. INTRODUCTION

The surveillance cameras are a form of Cyber Physical System, which are deployed broadly to supply visual surveillance monitoring of activities. Surveillance cameras are widely being employed publically to increase public safety and to remain an eye fixed on everything. The goal of the abnormal event detection system is to signal an activity that deviates from normal patterns, and then it identifies the time window of the occurring anomaly. Once an anomaly or the abnormal event is detected, it can further be sorted into the told activities using classification techniques. It takes sufficient computing power on investigation and to detect an anomaly happening in real time, because if the event cannot be identified in real time, the aim of CCTV model isn't served. Many approaches are undertaken in analyzing and classifying

video events, as an example, object tracking, pedestrian detection, crowd counting and action detection. During the previous couple of decades, surveillance cameras installation has increased abundantly in numerous locations.

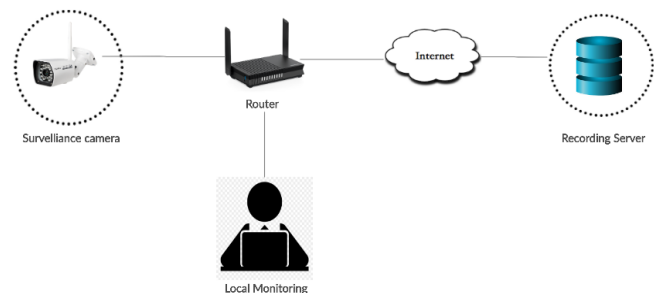


Figure1:Process flow diagram of existing system

However, hiring people to look at all the surveillance video is expensive. Figure1 shows the process flow diagram of local monitoring system for anomaly detection. The aim of our project is to implement an algorithm which can detect abnormal event like assault, abuse or arrest. Nowadays, several AI techniques are used to detect peculiarity, amongst them convolution neural networks using deep learning techniques improved the discernment accuracy remarkably. Figure2 describes the process flow of proposed system in which deep learning is applied for anomaly detection.

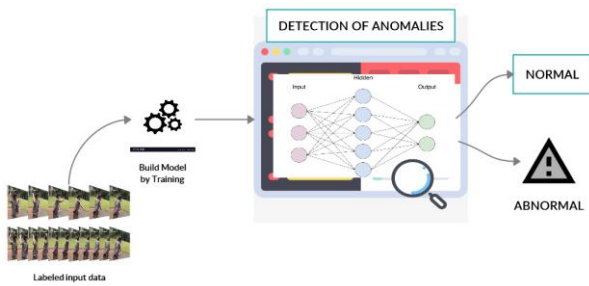


Figure2: Process flow diagram of proposed system

## II. RELATED WORK

Despite the substantial advances made by deep learning methods in many machine learning problems, there's a relative scarcity of deep learning approaches for anomaly detection. A survey by Chalapathy et al. (2019) elaborates and reviews various structured and comprehensive deep anomaly detection (DAD) methods. Anomaly detection is nearly related to outliers detection and novelty detection, therefore there exists various researches and reviews on this issue [3,4,5,6,7]. In the system [8] abnormal event at ATM systems are detected using image processing based on IOT technologies. The system recognizes human activity through CCTV cameras and analyzes them using various signal properties. The system faces low accuracy due non usage of feature extraction methods like SVM, etc.

Also trajectories are used for anomaly detection. In the system [9] anomalies in videos are detected by learning regularity in skeleton trajectories. Tran et al. (2019) propose to leverage 2D human skeleton trajectories for detecting abnormal events related to human behaviour in surveillance videos.

Recently, various methods primarily based on deep learning are applied for anomaly detection [10,11,12]. Zhou et al. (2016) proposed descriptor based on a novel simplified Histogram of Oriented Tracklets (sHOT), which provides promising results in detecting abnormalities. The author used 3D-CNN for feature extraction. Since, the algorithm applied was based on supervised learning hence not suitable for real world. Hasan et al. (2016) used an autoencoder to learn regularity in video sequences. The method identifies the deviations from regularity, and hence detects abnormal deviations. Ravanbakhsh et al. (2017) proposed the use of Generative Adversarial Nets (GANs) which are trained by normal frames and their representations are generated by optical flow.

We have proposed a system in which anomalies are detected by machine vision and is enhanced by deep learning methodologies like CNN.

## III. MACHINE VISION

In the work [14] importance of machine vision is surveyed along with its industrial applications and algorithmic applications. As stated in [14] Machine vision is the technology

to replace or complement manual inspections and measurements with digital cameras and image processing. Methods such as deep-learning technologies and convolutional neural networks (CNNs) from the field of artificial intelligence (AI) are evolving with machine vision to make image-processing systems learn and distinguish between defects and hence making identification processes even more precise.

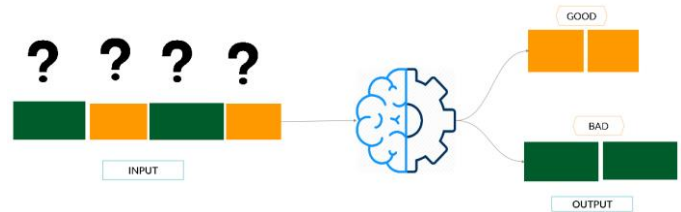


Figure3: Analyzing data with the help of deep learning and machine vision

Machine vision is a technology which gives machines high-level understanding from digital images and videos. One of the prominent application of this is anomaly detection in surveillance cameras. Basically, video is a bunch of images that are moving very fast we cannot distinguish between two images taken at a same time but a machine can due to the fact that images are nothing but collection of pixels.

Hence, anomaly detection can also possible with this process by detecting any change in intensity, opacity, density, etc in image or video.

## IV. DEEP LEARNING

Deep learning is a fast-growing field of machine learning and subfield of AI focused on using neural networks for complex practical problems. Deep neural networks have wide application and are used for visual perception and image analysis and classification, for various modules of self-driving cars which contains object detection and many more. Deep learning is a fast-growing field of machine learning and subfield of AI focused on using neural networks for complex practical problems.

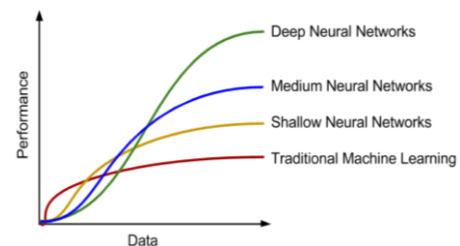


Figure4: Performance comparison between Deep learning based and traditional algorithms [2]

As a language may be a tool to specify something. A national language may be a tool which you'll express a minimum of to any or all the humans. And similarly deep learning, in turn, may be a language that permits you to hint your machine learning model about what we want this model to specify and



perform. Hint is about what quite features you would like it to possess, and what quite expert knowledge may be applied for on this dataset. It is a practical field.

Computer vision is one among the sphere that's been advancing rapidly thanks to deep learning. Also deep learning is even enabling new kinds of possibilities to be created by advancing different algorithms. So, the 2 reasons one should work in deep learning for computer vision. First, rapid advances in computer vision are enabling novel applications to look at, though they solely were impossible some years ago. Second, because the machine vision research community has been so creative then inventive in designing new neural network architectures and algorithms, is what truly inspire that makes lots cross-fertilization into other areas additionally.

### V. CNN

Convolution Neural Network is a part of deep neural network to analyze and process any image. Hence basically it is a approach to reduce the features and parameters of input image and assign different weights and bases to different objects to reduce the scale of enormous dataset. It has various building blocks including convolution using edge detection, padding, strided convolution, pooling and convolution on multiple layers.

The convolution operation is one of the fundamental building blocks of a convolutional neural network. So let's see for a computer to determine what are the objects in this picture, the primary thing you might do is probably detect vertical edges in the image. In order to make deep neural networks one modification to the basic convolutional operation that you can make use of padding. And the math of this seems to be that if you have a n by n image and to involved that with an f by f filter, then the dimension of the output will be; n -f+1 by n -f+1. Padding helps to capture the boundary features efficiently. Then comes strides, Strided convolutions is another piece of the basic building block of convolutions as employed in Convolutional Neural Networks.

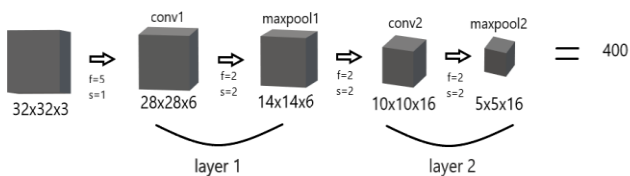


Figure5: Structure of Convolutional Neural Network

Here, rather than stepping over by one step, we are visiting step over by two steps. So, we are supposed to make it hop over two steps like and so on. Also pooling is a function is to progressively interrupt the dimensions of the image to reduce the amount of parameters and computation within the network by selecting largest number from a given frame.

Model used is C3d model: C3D are deep 3-dimensional CNN with a regular architecture containing 3 x 3 x 3 conv net and 2 x 2 x 2 pooling at each layer.C3D is obtained by training a

deep 3D convolutional network on an outsized annotated video dataset. In this project we can use a binary SVM classifier as a baseline method. Specifically, we treat all anomalous videos as a single class and normal videos as different class.C3D features are computed for every video, and a binary classifier is trained with linear kernel. For testing, this classifier provides the probability of every video clip to be anomalous.

### VI. ANOMALY DETECTION

Anomaly detection comes under the most challenging and long standing problems in computer vision. For video surveillance and continuous monitoring applications, there are several attempts to detect violence or aggression in videos. Beyond violent and non-violent patterns discrimination, we proposed to use the traditional motion of individual and detect deviation from that standard motion as an anomaly. And these can be achieved using different methods, we use a new method to model the normal patterns of human movements in surveillance video for anomaly detection using dynamic skeleton features. We decompose the skeletal movements into two sub-components: global body movement and native body movement.

Also we can obtain irregular events from human interaction using coupled features by applying encoder and decoder recurrent network. This coupled features interact with one another in spacial temporal model which helps in identifying the patterns. Compared to primary models, our model can achieve superior outlier detection performance.

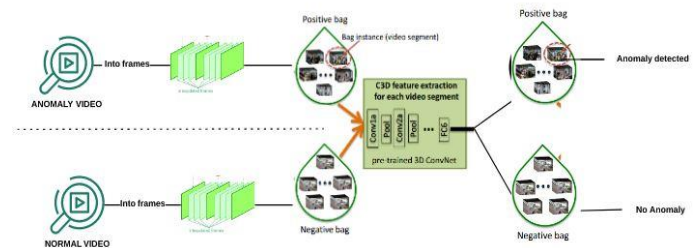


Figure6: Process flow of anomaly detection method by training of normal and anomaly videos

### VII. EVALUATION OF EXPERIMENT AND IMPLEMENTATION

#### A. Implementation details-

We extract visual perception from the fully connected (FC6) layer of the C3D network. Before computing features, we resize each video frame. We compute C3D feature followed by normalization. We input these features (4096D) to a 3-layer FC neural network [13]. We use tanh activation and Relu activation for the primary and tertiary FC layers respectively. We divide the video images into overlapping segments and calculated the gradient. Each video passes through the network and we get the score for every temporal segments. Then we compute loss and back-propagate the loss for the entire batch.



### B. Evaluation-

Sultani et al. (2018) mentioned a proposed dictionary based approach to acknowledge the conventional behaviors and used reconstruction errors to detect anomalies. Following their code, we extract some cuboids from each of the conventional training video and compute gradient based features in each volume [13]. After reducing the feature dimension using PCA, we learn the dictionary using sparse representation. Also Sultani et al. (2018) mentioned a completely convolutional feedforward deep autoencoder based approach to learn local features and classifier. Using their implementation, we train the network on normal videos using the temporal window of 40 frames. Similarly reconstruction error is employed detect anomaly. We also use a binary SVM classifier. Specifically, we treat all anomalous videos collectively and normal videos as another class. C3D features are computed for each video, and a binary classifier is trained with linear kernel [13]. For testing, this classifier provides the probability of each video clip to be anomalous.

### C. Training and Testing-

The proposed model is loaded with lot of positive and negative videos with labels, the network can automatically learn to predict the location of the anomaly in the video. To overcome this, the network should learn to produce high scores for anomalous video segments during training iteration [13] and accordingly we set the number of epochs less than 1000 to prevent overfitting. To prevent the model from false alarm rate, maximum of our dataset contains normal videos.

## VIII. CONCLUSION

The Proposed multi-instance learning model is a way to deal with discerning inconsistencies and abnormalities in surveillance videos, using deep learning approach coupled with machine vision. The abnormalities detected mainly are arrest, assault and abuse. The proposed detection method was enhanced via clustering through the competitive learning process. In future works, the system could be considered for implementing more different types of anomalies and also an alarm system to alert the concerned authorities.

## IX. ACKNOWLEDGMENT

We would also like to show our gratitude to everyone in Department of Computer Science and Engineering, Acropolis Institute of technology and research, Indore for sharing their pearls of wisdom with us during the course of this research. We thank Dr. Sanjay Bansal, HOD, CSE department for assistance and for the comments that greatly improved the manuscript.

## X. REFERENCE

- [1] Chalapathy Raghavendra, Chawla Sanjay (2019). "Deep learning for anomaly detection:A survey", A Preprint, *ResearchGate*, available online: <https://www.researchgate.net/publication/330357393>
- [2] Deep Learning Vs Traditional Learning. In <https://blog.easysol.net/wp-content/uploads/2017/06/image1.png>
- [3] Chandola V., Banerjee A., and Kumar V.(2009). "Anomaly detection: A survey," *ACM Computing Surveys*, vol. 41, no. 3.
- [4] Omar S.,Ngadi A. and Jebur H. (2013). "Machine learning techniques for anomaly detection: An overview," in *International Journal of Computer Applications*, vol. 79/2, pp. 33–41.
- [5] M. A. Pimentel, D. A. Clifton, L. Clifton, and L. Tarassenko (2014) "A review of novelty detection," *Signal Processing*, vol. 99, pp. 215–249.
- [6] Agrawal S. and Agrawal J. (2015). "Survey on anomaly detection using data mining techniques," *Procedia Computer Science*, vol. 60, pp. 708–713.
- [7] D. Kwon, H. Kim, J. Kim, S. C. Suh, I. Kim, and K. J. Kim (2017). "A survey of deep learning-based network anomaly detection," *Cluster Computing*, vol.9,no.5,p.205.
- [8] Kande Archana and P Bhaskara Reddy (2018) "To detect abnormal event at ATM system by using image processing based on IOT technologies," in *International Journal of Engineering & Technology*, vol. 7, no 3.
- [9] Tran Truyen, Morais Romero and Venkatesh Svetha (2019). "Learning Regularity in Skeleton Trajectories for Anomaly Detection in Videos", *Researchgate*, available online: <https://www.researchgate.net/publication/33851180>
- [10] Zhou S., Shen W., Zeng D., Fang M., Wei Y., Zhang Z., (2016) "Spatial-temporal convolutional neural networks for anomaly detection and localization in crowded scenes", *Signal Process. Image Commun.*, 47, pp 358–368.
- [11] Hasan M., Choi J., Neumann J., Roy Chowdhury, A.K. Davis, (2016) "Learning temporal regularity in video sequences". In *IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016*; pp. 733–742.
- [12] Ravanbakhsh M., Nabi M., Sangineto E., Marcenaro L., Regazzoni C., Sebe N. (2017) "Abnormal event detection in videos using generative adversarial nets". In *IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017*; pp. 1577–1581.
- [13] Sultani W., Chen Chen, Shah M. (2018), "Real-world Anomaly Detection in Surveillance Videos," *arXiv:1801.04264v1 [cs.CV] 12 Jan 2018*.
- [14] Patel Kirtan B, Zalte M. B. and Panchal (2013) "A Review: Machine vision and its Applications," *IOSR Journal of Electronics and Communication Engineering*



(IOSR-JECE), e-ISSN: 2278-2834, p-ISSN: 2278-8735.  
Volume 7, Issue 5 (Sep. - Oct. 2013), pp.72-77.