# Unusual Event Detection Using Machine Learning Techniques

**Ganesh Hathwar B S**

*Department of Computer science and Engineering,*
SMVITM, Bantakal.
*Udupi, Karnataka, India*

**Sandesh Prabhu**

*Department of Computer science and Engineering,*
SMVITM, Bantakal.
Udupi, Karnataka, India

**Vivek Shenoy Y**

*Department of Computer science and Engineering,*
*Udupi, Karnataka, Indi*

***Abstract*---- At the context, situation, due to variation of definitions between normality and abnormality unusual event detection for video sequences is a tough challenge. Though, it may generally be taken into consideration that unusual event or an activity, by way of unusual events happens much to a smaller extent compared to normal events. In order to get rid of unusual activities in surveillance videos, numerous kinds of modelling approaches are proposed inside the literature, inclusive of trajectory-based models, feature-based models and sparse reconstruction-based version, the main aim of these proposed models was to address the ambiguity in event detection and to examine and extract the hand-crafted-models or deep-learning-based models. Currently, the existing unusual-event detection models using deep learning techniques focus on data represented using the feature form, which gives a slight importance to the impact of the internal structure characteristics of feature vector. Additionally, it is bit difficult to ensure the classification accuracy using single classifier. In order to overcome the above issues, we proposed an unusual event detection system using machine learning techniques using CNN algorithm. First, in order to extract the spatiotemporal features of video frame We use convolutional neural network (CNN) and long short-term memory models. This helps us to process the constructed model at the faster rate, and then in order to capture the internal sequential and topological relational characteristics of structured feature, the feature expectation for each key frame of every video is done. Finally, the constructed model is tested along with the trained model experiments by providing the video dataset as input. Comparison with some developed models the performance of this proposed method is better than several the state-of-the-art approaches.**

***Keywords*** **--- *CNN*, surveillance, algorithm, accuracy, deep learning, training, testing.**

## I INTRODUCTION

Due to urbanization the places like airports, railway stations, public parks and shopping centers malls will be crowded with people. Due to this there is an increase in crime activities in these places. To reduce these situations, the big crowd should be continuously observed, and detect these incidents as fast as possible. To analyze crowd, security and safety of the people are main objectives. The continuous monitoring of these events by the security personnel is a very challenging task and it also requires a lot of time, resources and patience. In order to overcome these drawbacks, it inspires researchers to build a model which automatically detects these unusual event activities in the crowds and facilitating the appropriate operations.

In order to detect unusual events in the crowd the many research has been conducted in the domain of computer vision and image processing methods. Recently, many research works has been made using machine learning and deep learning models to fulfill the objectives of the problems. In spite of all the completed research work and success in this domain, there are drawbacks like ground truth data, type of the unusual event and many more.

In order to develop well organized unusual event detection systems, the computer vision community still faces a variety of challenges, which includes unavailability of adverse climate conditions, cameras, night vision problems etc. All these challenges often lead to inaccuracy in outputs and capturing the particular area of interest.

The main objective of the system is to identify unusual events in the populated area by simultaneously processing the visual data asynchronously acquired by the CCTV.

## II. LITERATURE REVIEW

[1] This paper deals with survey on various features-based models, and later review was done on deep learning-based models

for unusual event detection. In order to filter unexpected characteristic representation from the variations in the video frames, they have constructed a feature expectation sub-graph. [2] This paper deals with the methods to find vehicles and their uses in crime, The methodology used here is deep learning techniques which include SSD algorithm that is coupled with K-nearest neighbors' algorithm and CNN. In this paper, author used vehicle datasets and system developed gives vehicle with registration number and color information. The drawback here is the system was developed to identify vehicles with only few colors. This system fails to identify the vehicles with multiple colors. [3] In this paper, author used the distributed deep learning approach for abnormal behavior detection using big data analytics and multi-layer ensemble learning machine. The developed system shows high performance or provides good accuracy in detection of abnormal behavior in distributed way. The Drawback of this system is no privacy mechanism for this project. [4] The author published a paper which deals with a model that studies combinations based on sparse basis. In order-to obtain required data, each frame is brought down into various scales and each layer sets of nonoverlapping patches were portioned uniformly. This model reaches high accuracy rates on standard datasets at a speed of 138~152 FPS while evaluating on a standard desktop personal computer using MATLAB. The drawback in this project is low accuracy. [5] In this paper author conducted a survey on unusual event detection where training orders were not available and anomalies are labelled independently of temporal ordering. The algorithm used here were based on the classical density estimation approach of learning high-dimensional models and finding events of less probability. The drawback is, no training data available for initial process. [6] The author presented a probabilistic multi-task learning approach for visual definition estimation present in the video. The algorithm used here learns various fusion strategies, which are used to integrate the stimulus-driven and task-related components that are used to obtain the visual proper definition. The result says that the approach performs eight state-of-the-art approaches remarkably. The drawback is here there is no efficiency in this system [7] The author proposed more unique way in order to determine annotation outliers using the interestingness prediction task as unified robust learning to rank problem that detects both the outlier detection and interestingness prediction tasks. Applications of this on dataset containing images and videos denote that this approach can excel state-of the-art alternatives significantly. The drawback of this system is low performance of the system. [8] In this paper the proposed model is based on video event that are studied at each and every pixel with no need of supervising, using densely constructed spatio-temporal video volumes. Here spatio-temporal contextual data is decomposed into distinctive spatial and temporal contexts. The proposed structure studies the models of the provided spatial and temporal occurrence. Hence, this model is qualified of modeling simultaneously high-level behaviors as well as low-level spatial, temporal and spatio-temporal pixel level changes. Drawback of this system is it does not consider trajectories, and hence it does not learn long-term behaviors. [9] This paper that deals with methods to learn anomalies by using both normal and anomalous videos. Here first surveillance videos are divided into fixed number of segments during training phase. These segments make instances in a bag. Using both positive and negative bags instance the anomaly detection model is trained using the deep Multiple Instance Learning ranking loss. [10] The author published a paper that is used for detecting unusual event such as fighting and overcrowding using low-resolution video. This model can be particularly used in the ATM. Understanding and identifying the object that is moving, another important part of surveillance system is entering the scene.

## III. PROPOSED MODEL

Now a days the surveillance for activities is been done through cameras which are station outside the buildings and corner of the roads, but still fail to give efficient results on the crime activity or the abnormal activity around the surveillance areas still these systems need human interaction to differentiate between the activities. Some of the drawbacks of existing system are low accuracy, less efficient, long-term behaviours are not learned and it requires supervisor. The aim of the proposed system is to develop an intelligent supervision on video footage which will replace the traditional passive video surveillance footage. So that crime activities performed by human being or any unusual incidents can be captured. This system also helps to differentiate between unusual activities and normal activities. We used a Deep learning approach to detect the unusual event classification. In this proposed system, the algorithm used is CNN. We used this algorithm to increase the accuracy of the system and to find the Name of the material correctly

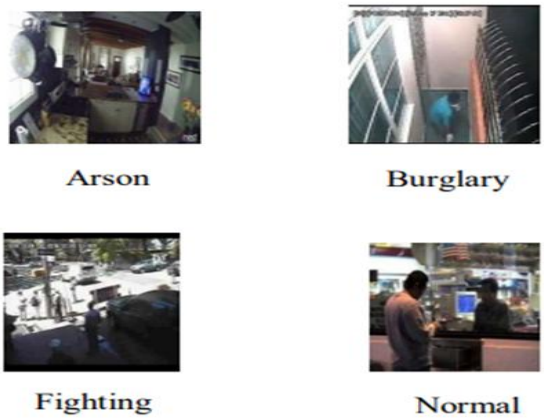


Fig 3.1 – some normal and abnormal activities

The above figure 3.1 shows come of the categories of unusual event. The datasets used in this model is video datasets. The business intelligent model has been developed, to classify different unusual events, based on a specific business structure with Anomaly classification using deep learning approach. The model is then evaluated by a scientific approach using confusion matrix to measure accuracy. Some of the advantages of this system are fault tolerant, more effective compared to existing model, provides good accuracy and scales well.

**IV METHODOLOGY**

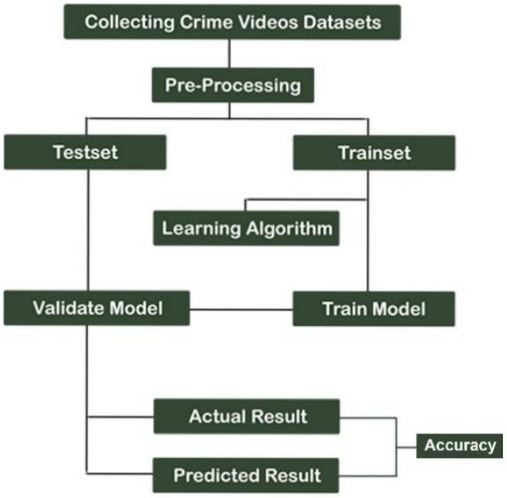


Fig 4.1 - System Architecture

There is total five steps. Data Collection, Image Acquisition and Pre-processing, Data Preparation and Model construction, Model training, Model testing and evaluation.

**Data Collection** – The first is to collect data from various sources such as YouTube, Google. We will collect the various types of footages like Arson, Burglary, Fighting and Chain Snatching etc. The data that is being collected in this step is videos. Now, this video datasets are processed further in next step.

**Image Acquisition and Pre-processing** – The video datasets are processed in this step. Using various python built in libraries such as open cv etc. We will obtain various frames of the video. Video frame convertor is used to obtain all the frames in one particular format. We will then use geometric image transformation which is used to resize the image or scale the image. Image Interpolation Algorithm is used here to modify the image size based on algorithm. To improve or reduce the target resolution of the image the image scaling process is being carried out this will reduce further noise that is present in the image and the image data size is reduced or adjusted.

Computers cannot interpret/translate images in the way humans do, they are can perform computations only on numbers. We will convert the images into numbers. The images are converted to greyscale. Based on the pixel how dark it is value will be assigned to each pixel by the computer. Then all the numbers are inserted into the array and the computer performs the computations on that array. In the next step this resulting array is used.

**Data Preparation and Model construction** – The next step is splitting the data into training dataset and testing dataset. The testing dataset is kept aside, which is used to test the model once the model is constructed and trained successfully. Randomly we choose X percent of the training dataset to be the actual training set, which is used to train the model and the remaining (100-X) percentage used as test set or validation set to test the model, where the X is a fixed value (say 70%), we then train the model iteratively and validate it using testing datasets. During this phase we successfully get training dataset and testing dataset. We then build model using Convolution Neural Network (CNN). CNN is a kind of Artificial Neural Network (ANN). Now all the pre-processing steps are done and we will start actual implementation of neural network. There are three convolution layers along with 2 x 2 max-pooling. The three layers are input layer which accepts inputs, the next layer is hidden layer and last layer is output layer that gives output or generate predictions.

**Max-pooling**- It is a method which reduces the proportion of an image by considering the maximum pixel value of the grid. This method helps to reduce over fitting and makes the model more general. After that, adding of two fully connected layers takes place. Since, the input of fully connected layers is 2D, and the output of convolution layer is 4D, it requires a flattening layer in-between the input and output layer. Fully connected later is SoftMax layer.

**Model training** - After model construction is done, we then train the model. We build artificial CNN that will identify images. After the model is trained, testing of the model will be done using the predefined train and test dataset.

**Model testing and evaluation** - Once the model has been trained the model is then tested and evaluated using confusion matrix to find the accuracy of the system. During this phase testing dataset is given as input to the model. This dataset has not been seen by the model. This gives the right accuracy of the model. Finally, the model is saved and used in the real world. The model can be used to evaluate the new data so this phase is called model evaluation.

## V IMPLEMENTATION

### Data Input and Data Processing Model

Step 1: Collecting video datasets from various data sources
Step 2: Importing Required libraries
Step 3: Classify the videos based on categories
Step 4: Saving the path of each video file in the text document for further references
Step 5: Fetching the video files, using path stored in text document.
Step 6: Frame Conversion
Step 7: Extracting frames of each video using python models
Step 8: Storing the extracted frame in specific folder
Step 9: Convert image instance into array of computable values
Step 10: Initialize "CNN Model"
Step 11: Compile and configure the model for training.
Step 12: Predicting the output of the testing data
Step 13: comparing testing output with predicted output
Step 14: Checking the overall accuracy of the CNN Model, using Test-set

**Output Model**

Step 14: Storing the video file, that is to be predicted/ Classified by the model in a specific folder
Step 15: Capturing the video from the given path
Step 16: Converting all the frames into array (as previously done during Training model)
Step 17: Predicting the output using the Train model
Step 18: Classify the video based on the index value

## VI RESULT

The below Figure shows the User Interface of the proposed model.

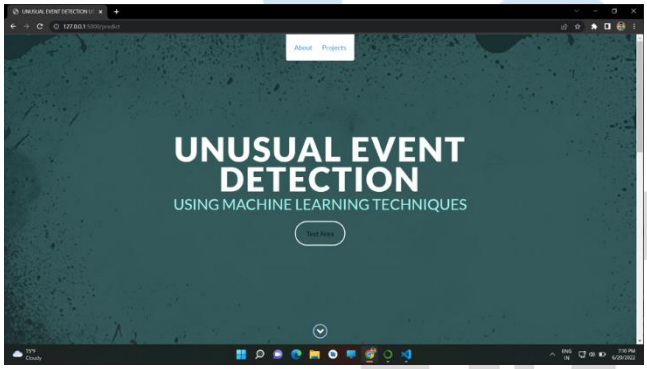


Fig 6.1 -Here, the field test area specifies the area where we can actually test the videos for unusual events.

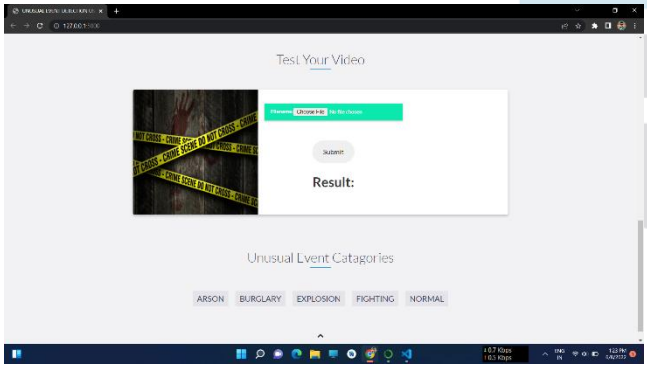


Fig 6.2 - Next step is to upload the video, here we can upload only video files.

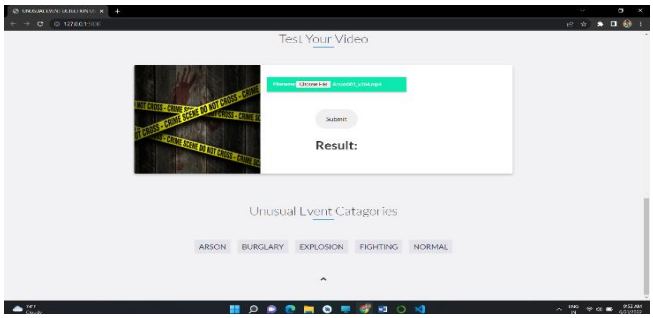

Fig 6.3 - Now, if we submit the file, it runs the model and gives the output.

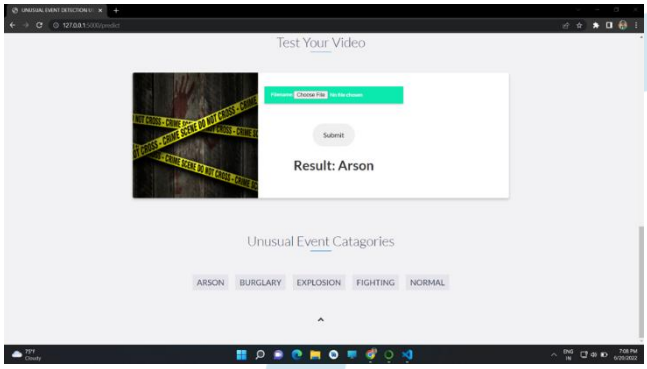


Fig 6.4 – result is shown as Arson in the above figure since we have uploaded the footage of Arson incident.

## VII CONCLUSION

The proposed system will provide the safety measures for the city corporate. In this system, using the CNN algorithm and the DEEP LEARNING if any abnormal event is occurred it checks the appropriate category and gives the result. This developed model shows around 97 % accuracy.

## VIII FUTURE ENHANCEMENT

We can improve this project by developing much faster and reliable system that will recognize any unusual events which helps in biometric features to incorporate variation into Human activity Recognition System. Developing the system that consists of a login system in order to enhance security. Develop system which will automatically alert concerned department about crime. This developed system can also be improved to process the live footage of CCTV and alert the concerned department if it founds any unusual events so that criminal activities or incidents can be prevented.

## IX REFERENCES

[1] O. Ye, J. Deng, Z. Yu, T. Liu and L. Dong, "Abnormal Event Detection via Feature Expectation Subgraph Calibrating Classification in Video Surveillance Scenes," in IEEE Access, vol., pp.9756497575, 2020, doi:10.1109/ACCESS.2020.2997357.

[2] Praveen Saini, Karambir Bidham, Sona Malhotra, " A Detection System for stolen vehicle using vehicle attributes with deep learning", 5th International Conference on Signal Processing (Computing and Control (ISPCC 2k19), Oct 10-12, 2019, JUIT, Solan, India

[3] Naila Marir, Huiqiang, Wang, Guangheng Fenc, Bingyang Li, and Neijuan Jia. "Distributed Abnormal Behavior Detection Approach Based on Deep Belief Network and Ensemble SVM Using Spark". In IEEE Access,(2018), doi : 10.1109/ACCESS.2018.2875045

[4] Cewu Lu, Jianping Shi, Jiaya Jia. "Abnormal Event Detection at 150 FPS in MATLAB", 2013 IEEE International Conference on Computer Vision

[5] A. D. Giorno, J. A. Bagnell, and M. Hebert, ``A discriminative framework for anomalydetection in large videos, " in Proc. 14th Eur. Conf. Comput.Vis., Amsterdam, The Netherlands, Oct. 2016, pp. 334_349.

[6] J. Li, Y. Tian, T. Huang, and W. Gao, "Probabilistic multi-task learning for visual saliency estimation in video," Int. J. Comput. Vis., vol. 90, no. 2, pp. 150_165, Nov. 2010.

[7] Y. Fu, T. M. Hospedales, T. Xiang, S. Gong, and Y. Yao, “Interestingness prediction by robust learning to rank," in Proc. 13th Eur. Conf. Comput. Vis., Zurich, Switzerland, Sep. 2014, pp. 488_503.

[8] M. J. Roshtkhari and M. D. Levine, “Online dominant and anomalous behavior detection in videos," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Portland, OR, USA, Jun. 2013, pp. 2611_2618.

[9] Waqas Sultani, Chen Chen2, Mubarak Shah, “Real-world Anomaly Detection in Surveillance Videos”. Department of Computer Science, Information Technology University, Pakistan Center for Research in Computer Vision (CRCV), University of Central Florida

[10] Sudhir Goswami, Jyoti Goswami, Nagresh Kumar,” Unusual Event Detection in Low Resolution Video for enhancing ATM security,” Computer Science Department MIET, Meerut, Electronics Department NITTTR, Chandigarh