HUMAN SUSPICIOUS ACTIVITY DETECTION SYSTEM USING CNN MODEL FOR VIDEO SURVEILLANCE

Tejashri Subhash Bora[1], Monika Dhananjay Rokade[2]

[1] PG Student, Department of Computer Engg., SPCOE, Maharashtra, India
[2] Assistant Prof., Department of Computer Engg., SPCOE, Maharashtra, India

ABSTRACT

This paper brings forward one amongst the foremost significant applications of human suspicious activity recognition that is termed as anomaly detection. A key concern of any society today is providing safety to an individual. The main reason behind this concern is due to the constantly increasing activities causing threats, starting from deliberate ferocity to an injury caused through an accident. Simple installation of a traditional closed circuit television(CCTV) is not sufficient as it requires a person to continuously stay alert and monitor the cameras, which is quite inefficient. This call for the requirement to develop an security system which is fully automated system that recognizes anomalous activities in real time and brings instant help to the victims. Hence we proposed a system which will examine and detect the suspicious human action from real-time CCTV footage with help of machine learning techniques and generates the alert if the abnormal activity is occurred. The method is implemented on the dataset containing both normal and anomaly activity and experiment has shown better results.

Keyword : - *Video Surveillance, Anomaly detection, Machine learning, Convolutional neural networks, Image processing.*

1. INTRODUCTION

Human activity recognition can be useful to a variety of scenarios, and anomaly detection in security systems is one of among them. Seen the increasing demand for security, surveillance cameras have been widely set up as the infrastructure for video analysis. One of the major challenges faced by surveillance video analysis is detecting abnormal activity which requires exhausting human efforts. Fortunately, such a labor-intensive task can be recast as an anomaly detection problem which aims to detect unexpected actions or patterns. Anomaly detection varies from the traditional classification problem in the following aspects: 1) It is very difficult to list all possible negative (anomaly) illustrations. 2) It is a daunting job to collect adequate negative samples due to the rarity.

An activity recognition system is projected to identify the basic day to day activities performed by a human being. It is challenging to achieve high rate accuracy for recognition of these activities due to the complexity and diversity in human activities. Activity models required for identification and classification of human activities are constructed based on different approaches specific to the application. The activities of a human being can be generally categorized into normal activities or anomalous activities. A human being's deviation from normal behavior to abnormal causing harm to the surrounding or to himself is classified as an anomalous activity. To achieve anomaly detection, one of the most widespread method is using the videos of normal events as training data to learn a model and then detecting the suspicious events which would do not fit in the learned model. For example, human pose guesstimate is used in applications including video surveillance, animal tracing and actions understanding, sign language recognition, advanced human-computer interaction, as well as marker less motion capturing.

Low cost depth sensors consist of limitations like limited to indoor use, and their low resolution and noisy depth information make it difficult to estimate human poses from depth images. Hence, we are to using neural networks to overcome these problems. Anomalous human activity recognition from surveillance video is an active exploration part of image processing and computer visualization.

2. HISTORY AND BACKGROUND

According to [1] Sparse coding has constructed anomaly detection which shown better performance, even contain the theories are feature learning, sparse representation, and dictionary learning. In this paper, a innovative neural network is proposed for anomaly detection which is also labeled as AnomalyNet by deeply accomplishing feature learning, sparse representation as well as dictionary learning in three joint neural processing blocks. Specifically, to learn improved features, the authors design a motion fusion block accompanied by a feature transfer block to relish the benefits of eliminating background noisy, capturing motion and improving data insufficiency.

According to [2] An suspicious activity is any observation of action that could state a person may be involved in a crime or is about to commit a certain criminality. Anomaly detection is the process detecting suspicious activity. Surveillance cameras are one of the best solution to the issue of security in various places. Present-day system needs man power for monitoring the system as detecting and identifying criminal and abnormal activity is so challenging. So this paper carry out a survey on anomaly detection for video surveillance using different concepts like deep learning, RNN etc.

Then Research paper [3] automates the detection of anomalous actions within long video series is challenging due to the uncertainty of how such events are defined. The authors tactic the problem by learning generative models that can discover anomalies in videos using restricted supervision. Projected end-to-end trainable complex Convolutional Long Short-Term Memory (Conv-LSTM) networks that are able to predict the development of a video sequence from a minor number of input frames.

According to the paper [4] , authors inspired by the capability of sparse coding based suspicious detection, projected a Temporally-coherent Sparse Coding (TSC) where they implement similar neighboring frames be encoded with alike reconstruction coefficients. Then mapped the TSC with a distinct type of stacked Recurrent Neural Network (sRNN). The contributions of the paper are- i) proposed a TSC, which can be recorded to a sRNN which facilitates the parameter optimization and speed up the doubtful prediction. ii) Build a very huge dataset that is even larger than the summation of all existing dataset for finding anomalous activity.

The research paper from Springer [5] presented an efficient technique for identifying anomalies in videos. Recently applications of convolutional neural networks have shown possibilities of convolutional layers for object detection and recognition, specifically in images. Though, convolutional neural networks are supervised and have need of labels as learning signals. Authors as well as proposed a spatiotemporal architecture for suspicious detection in videos with crowded scenes.

The paper [6] proposed end-to-end trainable complex Convolutional Long Short-Term Memory (Conv-LSTM) networks. These Conv-LSTM networks are capable to predict the evolution of a video sequence from a minor number of input frames. Consistency scores are derived from the reconstruction errors of a set of estimates with irregular video sequences yielding lower regularity scores as they separate further from the actual sequence over time. The models employ a composite structure and observe the special effects of conditioning in learning more meaningful representations.

According to [7], the approach for this problem by learning a generative model for consistent motion patterns using multiple resources with very restricted supervision. Specifically, paper contains two methods that are built upon the autoencoders for their capacity to work with little to no supervision. The first method is to leverage the conventional handcrafted spatio-temporal local features and then study the fully connected autoencoder. Secondly, construct a fully convolutional feed-forward autoencoder to learn together the local features and the classifiers as an end-to-end learning structure. The proposed model is able to capture the regularities from numerous datasets.

The paper [8], authors has proposed the technique for actual time anomaly detection and localization in crowded scenes. Each video is well-defined as a set of non-overlapping cubic spots, and is explained using two local and global descriptors. The descriptors used here capture the video assets from different phases. By integrating simple and cost-effective Gaussian classifiers, we can distinguish normal events and anomalies in videos.

Then Research paper [9] is basically on inherent redundancy of video structures, authors propose an effective sparse combination learning framework. It accomplishes decent performance in the detection phase deprived of compromising result quality. The short running time is fail-safe because the new method efficiently turns the original complex problem to one where only a few less in cost small-scale least square optimization steps are considered. The process scopes high detection rates on benchmark datasets when figuring on an usual desktop PC by using MATLAB.

By the reference of the paper [10] in which a fully unverified dynamic sparse coding methodology for spotting unusual events in videos based on online sparse reconstructibility of query signals is proposed. Based on an perception that usual events in a video are mostly liked to be reconstructible from an event dictionary, whereas infrequent events are not, the algorithm works on a principled convex optimization formulation that permits both a sparse reconstruction code, and an online dictionary to be mutually inferred and modernized.

2.1 Research Gap

- Existing system stores the data in form of records only and needed continuous monitoring which is labor intensive task.
- Different detection techniques are implemented to find suspicious activity.
- Some researches are carried where the features are typically learned from scratch without considering the well-established pre-train model.
- There are large number of computations which result in high computation cost.
- Classification accuracy is low.

3. PROPOSED SYSTEM AND DESIGN

In our proposed system, for detecting anomalous behavior, the CNN i.e. convolution neural network have been used. For effectively classification of anomalous activities, it is essential to recognize the temporal data in the video. Recently, CNN is mostly used for extracting key features from each frame of the video. CNN is only the algorithm best suited for this purpose. For classifying the given input successful, it is necessary that the features get extracted from CNN, therefore CNN should be capable of knowing and extracting the needed features from the frame of videos.

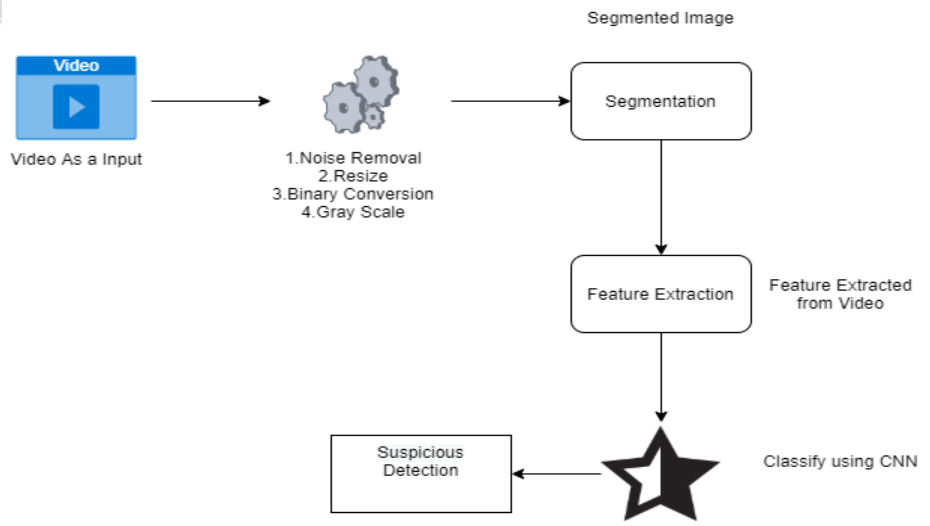

Fig -1: System architecture

Proposed work of model:

1. Data Collection: First of all, the information for different Websites and Social Media applications based on certain parameters is extracted data.

2. Preprocessing: Then we will apply various pre-processing steps such as Noise removal, resizing, binary conversion and gray scaling in order to make our dataset proper.

3. Noise removal: Noise is removed from the input video. In image processing, the key process for denoising is filtering. Generally average filters, median filters, Wiener filters and Kalman filters are utilized to reduce noise.

4. Resizing: Image resizing is necessary when we need to increase or decrease the total number of pixels, whereas remapping can be done when we are adjusting for lens distortion or rotating an image.

5. Binary conversion: A binary image is one that holds the pixels that can have any one of precisely two colors, classically black and white. Binary images are also well known as bi-level or as two-level. This means that each and every single pixel is put in storage as a solitary bit – i.e. in value of 0and 1.

6. Gray scaling: Gray-scaling is the method of transforming a continuous-tone image to an image that a computer can manipulate effortlessly.

7. Segmentation: Image segmentation is the significant process in which isolation of a digital image into multiple segments is carried out i.e. (sets of pixels, also recognized as image objects).6. Data Training: We compile artificial as well as real time using online news data and provide training with any machine learning classifier.

8. Feature extraction: Feature extraction is a part of the dimensionality decrease procedure, in which, an initial set of the raw data is separated and compact to more controllable groups.

9. Classification: Classification is the method of sorting and labeling groups of pixels or vectors with in an image based on definite rules and instruction

10. Data Training: We gathered artificial as well as real time using social media data and provide training with any machine learning classifier.

11. Testing with machine learning: We give testing dataset to system and apply machine learning algorithm to detect the activity accordingly.

12. Analysis: We determine the accuracy of proposed system and estimate with other existing systems.

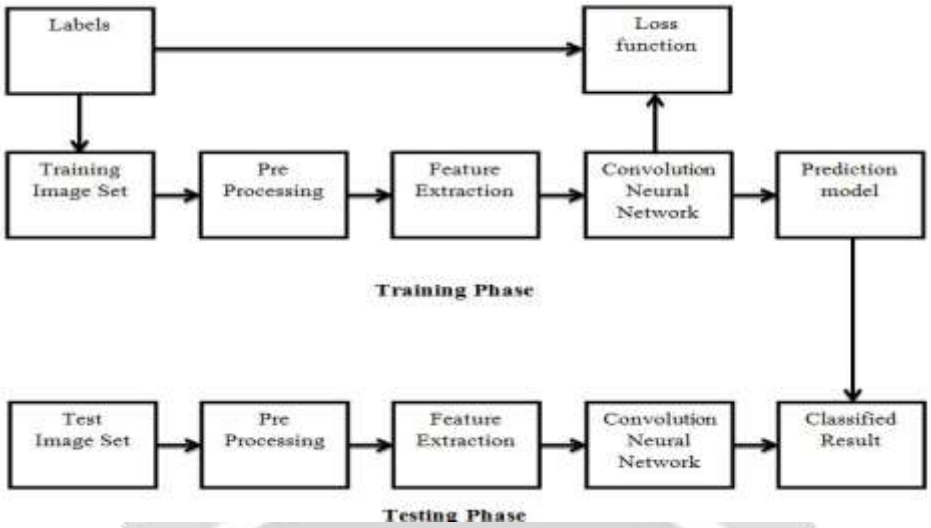

Fig -2: Block diagram of proposed model using CNN.

3.1 Algorithm Design

3.1.1 Algorithm : Convolution Neural Network(CNN)

Step 1: Input is given as image / video.
Step 2: Then many different filters are applied to the input to create a feature map.
Step 3: Next a ReLU function is applied to increase non-linearity.
Step 4: Then applies a pooling layer to each and every feature map.
Step 5: The algorithm compresses the pooled images into one long vector.
Step 6: In next step, inputs the vector to the algorithm into a fully connected artificial neural network.
Step 7: Processes the features via the network. At the end fully connected layer delivers the "voting" of the classes.
Step 8: In this last step trains through forward propagation and back propagation for numerous epochs. This repetition occurs until we have a well-defined neural network with trained weights and feature detectors.

4. RESULT AND ANALYSIS

The dataset of the proposed model includes videos of anomalous behavior which are Arson, Burglary, Fighting as well as it also contains videos of normal behavior. Since the surveillance videos are very sensitive and hard to get, the videos for the normal behavior had to be created manually with the help of standard camera and for the anomalous behavior the videos are taken from the work dataset and through social media.

We extract various existing data as well as currently posted information through various websites. The dataset includes a total of 15 videos for training purposes and 7 videos for the testing purpose. . We downloaded about 4-5 GB data as a video samples to use machine learning algorithms for testing the proposed method.

Table -1: Dataset description used in proposed model

Total Size	**4-5 GB**
Training Samples	**15 Videos**
Testing Samples	**7 Videos**

Table 2: System performance evaluation with proposed vs existing.

Technique	**Dataset**	**Accuracy**
Joey T. Z.	UFC Crime Dataset	76%
Proposed System	UFC Crime Dataset(Two classes anomalous and non-anomalous)	85%

The proposed system aims to detect the anomalous behavior happening in the video and the system is achieving the accuracy of 85% on created data set. Following are the images of result of the proposed model.

Fig -3:- Normal Activity Detection

Fig-4:- Anomaly Activity Detection

5. CONCLUSIONS

The proposed system is a machine approach to detect real-world criminal Activity identification in surveillance videos. The necessity to develop such a security system is increasing with the increasing number of crimes that are happening everyday. The result of the proposed system will be able to detect whether any anomaly action is taking place or not. And most of the previous researches had lower accuracy in determining the abnormal behavior. Therefore, in this a new approach CNN is used for better results.

6. ACKNOLEDGEMENT

Authors want to acknowledge Principal, Head of Computer department Prof. S.S. Khatal and the guide Prof. M.D. Rokade for all the support and help rendered. To express profound feeling of gratitude to the regarded guardians for giving the motivation needed for finishing of paper.

7. REFERENCES

[1] Joey Tianyi Zhou, Jiawei Du, Hongyuan Zhu, Xi Peng, Rick Siow Mong Goh, ”AnomalyNet: An Anomaly Detection Network for Video Surveillance, 2019.

[2] Monika D. Rokade and Tejashri S. Bora, "Survey On Anomaly Detection for Video Surveillance" 2021 International Research Journal of Engineering and Technology(IRJET).

[3] Jefferson Ryan Medel, Andreas Savakis, "Anomaly Detection in Video Using Predictive Convolutional Long Short-Term Memory Networks" under review.

[4] W. Luo, W. Liu, and S. Gao, "A revisit of sparse coding based anomaly detection in stacked rnn framework," in The IEEE International Conference on Computer Vision (ICCV), Oct 2017

[5] Y. S. Chong and Y. H. Tay, "Abnormal event detection in videos using spatiotemporal autoencoder," in International Symposium on Neural Networks. Springer, 2017, pp. 189–196.

[6] J. R. Medel and A. Savakis, "Anomaly detection in video using predictive convolutional long short-term memory networks," arXiv preprint arXiv:1612.00390, 2016.

[7] M. Hasan, J. Choi, J. Neumann, A. K. Roy-Chowdhury, and L. S. Davis, "Learning temporal regularity in video sequences," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 733–742.

[8] M. Sabokrou, M. Fathy, M. Hoseini, and R. Klette, "Real-time anomaly detection and localization in crowded scenes," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, June 2015.

[9] C. Lu, J. Shi, and J. Jia, "Abnormal event detection at 150 fps in matlab ," in Proceedings of the IEEE international conference on computer vision, 2013, pp. 2720–2727.

[10] H. Mousavi, M. Nabi, H. K. Galoogahi, A. Perina, and V. Murino, "Abnormality detection with improved histogram of oriented tracklets," in International Conference on Image Analysis and Processing. Springer, 2015, pp. 722–732.

[11] Monika D.Rokade, Dr. Yogesh Kumar Sharma, "MLIDS: A Machine Learning Approach for Intrusion Detection for Real Time Network Dataset", International Conference on Emerging Smart Computing and Informatics (ESCI), IEEE 2021.

[12] Monika D.Rokade, Dr. Yogesh Kumar Sharma, "Identification of Malicious Activity for Network Packet using Deep Learning ", in 2020.

[13] Monika D.Rokade, Dr. Yogesh Kumar Sharma, "Deep and Machine Learning Approaches for Anomaly-Based Intrusion Detection of Imbalanced Network Traffic", IOSR Journal of Engineering, 2019.

[14] Y. K Sharma, S Khatal Sunil, " Health Care Patient Monitoring using IoT and Machine Learning", IOSR Journal of Engineering, 2019.

[15] S Khatal Sunil, Y. K Sharma, "Analyzing the role of heart disease prediction system using IOT and machine learning", International Journal of Advanced Science and Technology, 2020.

[16] Z. Zhu, J. Wang, and N. Yu, "Anomaly detection via 3d-hof and fast double sparse representation," in Image Processing (ICIP), 2016 IEEE International Conference on. IEEE, 2016, pp. 286–290.

[17] T. Xiao, C. Zhang, H. Zha, and F. Wei, "Anomaly detection via local coordinate factorization and spatio-temporal pyramid," in Asian Conference on Computer Vision. Springer, 2014, pp. 66–82.

[18] M. D. Zeiler, "Adadelta: an adaptive learning rate method," arXiv preprintarXiv:1212.5701, 2012.