International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 Impact Factor: 8.176 ISSN: 2582-3930

ANOMALY DETECTION IN SURVEILLANCE VIDEOS USING DEEP LEARNING

Mr.M.Thirunavukkarasu
Assistant professor
Department of Computer Science and Engineering
SCSVMV
login2tiru@gmail.com

Dasari Manikanta
Department of Computer Science and Engineering
SCSVMV
11199A055@kanchiuniv.ac.in

Ala Jagan Mohan
Department of Computer Science and Engineering
SCSVMV
11199A008@kanchiuniv.ac.in

Abstract- Deep learning has established itself as a ground breaking computing strategy in the field of computer vision. The identification of abnormalities in surveillance footage is one of the challenging cognitive tasks for which it has been frequently employed. In this instance, anomaly detection refers to the recognition of anomalous occurrences in surveillance films that may constitute security incidents or threats. For anomaly detection, deep learning approaches have outperformed more conventional approaches. Traditional closed-circuit television (CCTV) systems are inefficient and expensive because they require constant human monitoring. Thus, a system that can efficiently detect human activities in real time is required. Due to the size of a surveillance video and the time required to analyze it for activity, the video must be compressed using adaptive compression techniques. With adaptive video compression, only the areas of the video with the least emphasis are compressed; the remainder is left uncompressed. The goal of the discussion is to be able to use surveillance video to construct an automated anomalous human activity identification system that can identify unusual behaviour and alert the real-time user. The research is divided into two parts that discuss methods for adaptive video compression of surveillance films and using that compressed video as an input to identify unusual human activities.

Keywords: **Anomaly Detection, CCTV, Convolutional Neural Network, Deep learning**

I. INTRODUCTION

Human identification largely relies on behavioural patterns and the human face. For these identifications, visual information is a crucial source. Surveillance videos offer this kind of visual data that may be seen as live movies or that can be replayed for later use. Even in the area of video analytics, the current "automation" movement is having an effect. Video analytics may be applied to a wide range of tasks, including motion detection, human activity prediction, person identification, anomalous activity detection, vehicle counting, crowd counting, etc. The two elements that are employed in this field to identify people are called, strictly speaking, face recognition and gait recognition, respectively. Face recognition is the more flexible method of these two for automatically identifying people from surveillance footage. The orientation of a person's head may be predicted using face recognition, and this will assist anticipate that person's behaviour. Motion recognition combined with face recognition is particularly helpful in many applications, including person verification, identification, and determining the presence or absence of a person at a certain location and time. A system that can successfully identify and classify suspicious conduct among students in an exam hall is also developed using human interactions such as delicate eye contact between two people, head motion detection, hand gesture identification, and estimating. This study presents a face-based technique for detecting suspicious human activities. Lastly, the system recognizes student contact and thereby inhibits the dissemination of potentially damaging information among students. Our study has helped develop a system that can analyze real-time footage of test rooms filled with students and categorize their behavior as suspicious or not. This study suggests an intelligent algorithm that can watch and analyze students' behaviour in a testing

environment and notify the management of the educational institution of any malpractice or suspicious activity.

II. LITERATURE SURVEY

"Learning Spatiotemporal Features for Pedestrian Detection in Surveillance Video Analysis" by Xiaoyu Zhang [1].This paper proposes a deep learning-based method for pedestrian detection and anomaly detection in surveillance videos. The proposed method uses a combination of convolutional neural networks (CNNs) and long short-term memory (LSTM) to capture spatiotemporal features in the video data.

"Anomaly Detection in Time Series Sensor Data Using Variational Autoencoder with Reconstruction Probability" by Rui Zhao [2] et al. (2020).This paper proposes a deep learning-based approach to anomaly detection in time series sensor data using variational autoencoder (VAE) networks. The approach works by training the VAE network to learn the normal patterns in the sensor data and then identifying anomalies as data points with low reconstruction probability. The proposed method achieved high accuracy on several benchmark datasets and demonstrated the effectiveness of the approach for real-world industrial monitoring applications.

"Anomaly Detection in ECG Signals Using Deep Autoencoder Neural Networks" by Kaveh Hassani [3] et al. (2018).This paper proposes a deep learning-based approach to anomaly detection in electrocardiogram (ECG) signals using autoencoder neural networks. The approach works by training the autoencoder network to learn the normal patterns in the ECG data and then identifying anomalies as data points that deviate significantly from the learned patterns. The proposed method achieved high accuracy on several benchmark datasets and demonstrated the effectiveness of the approach for real-world medical applications.

"Deep Anomaly Detection in Complex Industrial Processes Using LSTM Networks" by Fabrizio Riguzzi [4] et al. (2019).This paper proposes a deep learning-based approach to anomaly detection in complex industrial processes using long short-term memory (LSTM) networks. The approach works by training the LSTM network to learn the normal patterns in the industrial process data and then identifying anomalies as data points that deviate significantly from the learned patterns. The proposed method achieved high accuracy on several benchmark datasets and demonstrated the effectiveness of the approach for real-world industrial applications.

"Unsupervised Anomaly Detection in Surveillance Videos Using GANs and Spatial Temporal Autoencoder" by Shaoqing Ren [5] et al. (2020).This paper proposes an unsupervised approach to anomaly detection in surveillance videos using generative adversarial networks (GANs) and spatial temporal autoencoders (STAEs). The approach works by training a GAN to learn the distribution of normal video frames and then using a STAE to reconstruct the video data. Anomalies are identified as frames that deviate significantly from the reconstructed data. The proposed method achieved high accuracy on several benchmark datasets and demonstrated the effectiveness of unsupervised anomaly detection in surveillance videos.

"Deep Learning for Anomaly Detection: A Survey" by Zhou [6] et al. (2019).This paper provides a comprehensive survey of deep learning techniques for anomaly detection, including autoencoders, GANs, and variational autoencoders (VAEs). The paper discusses various applications of these techniques, including computer network intrusion detection, medical diagnosis, and industrial process monitoring. The authors also discuss some of the challenges and open research questions in this area, such as handling class imbalance and incorporating domain knowledge into deep learning models.

III. PROBLEM STATEMENT

The problem at hand is best summed up as the identification of anomalies in surveillance footage. In most surveillance recordings, a fixed camera is used to look out at a designated area where activities are taking place. occurrences are seen on a security video. We are dealing with enormous amounts of video data, which may easily wear people out and cause errors from manual intervention. The effectiveness of the system is significantly impacted. Video surveillance automation has provided a solution for this. Currently, it is difficult to manually watch every incident captured on a CCTV (Closed Circuit Television) camera. Even if the incident has already occurred, manually looking for it in the recorded video is a time-consuming process. This study makes advantage of both typical and unusual occurrences recorded by the surveillance CCTV cameras to overcome the limitations of previously utilized datasets. It is a brand-new, extensively studied huge dataset. The list of anomalies, consists of a variety of lengthy, unedited real-world surveillance recordings taken at various locations. So, the level of public safety may be gauged using these anomalies.

IV. PROPOSED METHODS

The approach suggested in this study makes use of deep learning algorithms to detect abnormalities and their activity patterns in video surveillance. It is employed in a deep learning-based video surveillance system for anomaly identification. Because human behaviour detection is an automated method of shrewdly identifying any suspicious conduct in video surveillance systems. In public places like airports, train stations, banks, workplaces, and exam rooms, we can employ CNN algorithms to automatically recognize human behaviour. Artificial intelligence, machine learning, and deep learning are increasingly being used in the field of video surveillance. Deep learning, which extracts features from raw pixel data to represent abstract semantic notions,

has a tremendous capacity for learning. Some feature extraction techniques use single frames while others rely on patch frames to cut costs and training time. The detection score is obtained by comparing each frame to the previous and subsequent frames, and the first attribute, appearance, is connected to object detection in each frame. The final score is determined using frame comparison and average speed, and the second attribute is density, which is the number of objects in each frame. The third characteristic, motion, creates optical flow and a video sequence that is subsequently utilised for another anomaly score. It is based on the movement of objects between patch frames. The last feature is called the scene, which builds a scene from a trained model using patch frames. These characteristics are combined to generate and identify scores as well.

Video-to-Frame Conversion: This Approach's initial step is to extract frames from the recorded CCTV footage. After a predetermined and brief period, the process extracts the frame (say 1 sec). The retrieved Frame is then scaled to the Inception V3 standard input dimensions of 1200 × 1200 pixels. The pre-process input function's goal is to convert the downsized picture to the format required by the model.

InceptionV3: The ImageNet dataset is used to train InceptionV3. A sizable dataset for the Visual Recognition competition was made public. The model attempts to categorize the dataset into frames, as is customary in computer vision. Throughout the first half, the model focused on generic traits from input images. In the second half, it categorizes those photos using the traits that were extracted. The conventional inception model's layers are fully described.

Convolution Neural Network: The transfer Learning is utilized to train our CNN using the InceptionV3 model, which has previously been trained. The feature extraction component of the new model is used in transfer learning, while the classification component is retrained using the original dataset. The feature extraction portion of the model, which is quite complicated, does not need to be learned, hence learning as a whole takes less time and computer resources. The CNN, which isn't the final classification model, receives the output of the Inception model as input. Instead, the output of the last pooling layer a vector with 2,048 features is retrieved and used as the RNN's input. A high-level feature map is a term used to describe the vector.

Grouping of feature maps into a single pattern: Several prepossessed frames are taken into consideration to provide the framework a feeling of the sequence. The final categorization is then based on this portion. A group of these frames may be used to categorize a period in the movie and to create the illusion of motion. For this, a few feature maps that the inception model (CNN), created during that predetermined period of the movie, anticipated are preserved. To create a high-level feature map, low-level features have been taken into consideration. These tools are used to locate items and forms in computer graphics.

Architecture

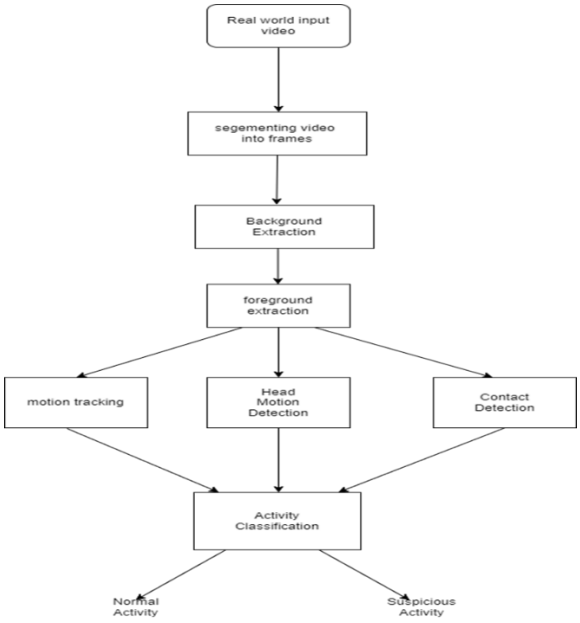

Fig 2. Anomaly activity Detection Architecture

From the flow chart Surveillance videos are used for detecting anomaly activity in the public place to reduce dangerous activity. The video These can be videos can be captured by the surveillance camera of taking the given dataset from the source and these videos get segment videos into frames of the given dataset from this dataset. And that can be removed the background extract and fore ground of the given video. From the extracting of all the videos tracking the human motion tracking by head motion detection and contact detection by this process we can identify which activity is happening is the captured is normal activity or anomaly activity. Through must consider several important elements, including time and computational complexity. The system that usages one algorithm including a relatively lesser period complexity, consuming a reduced amount of hardware properties then whichever produces respectable results resolve standard extra useful designed for period critical requests corresponding bank theft detection, noticing also writing doubtful actions by the public place, etc. Once the information collected after the respectively layer has been integrated, the image or visual interpreted or classed. Similar to this, CNN uses a variety of filters that each collect data after the image, on the point of edges and numerous forms (vertical, horizontal, and round), which are combined to identify every image.

Algorithm

The steps of CNN are listed below

Step 1: Input is given as video.

Step 2: The uploaded video is converted into multiple frames and saved in the frames file.

Step 3: Next frames go under the convolution layer.

Step 4: Then apply a pooling layer to every feature map.

Step 5: The algorithm compresses the pooled images into one long vector.

Step 6: In the next step, input the vector to the algorithm into a fully connected artificial neural network.

Step 7: Processes the features via the network. In the end, a fully connected layer delivers the Anomaly and the percentage of anomaly activity that has taken place.

V.EXPERIMENTAL SETUP

Python 3.7 must be installed on a 64-bit laptop in order to put the concept into practice. The source code for this application is important to check the box labelled "Add path to system variable" while installing the software. This choice will show on either first or second installation screen. After installing the software, run the commands below. In the whole episode computer has to be online. These are nine commands that should be executed for installation.

1)pip install TensorFlow

2) pip install NumPy

3) pip install SciPy

4)pip install OpenCV-python s

5)pip install pillow

6) pip install matplotlib

7) pip install h5py

8) pip install keras

9) pip installs

The first step in automating the process is to build a training model using a large collection of photos (all images that could potentially identify aspects of suspicious activity) and then use the Python library TENSOR FLOW to build a convolution neural network. Then, following the upload of any video, the app will use the video's frames to train a model that will identify the type of the video, such as "suspect or normal."

The order of all the execution steps is listed below. Two types of videos are fed, one with thermal images and other with the images whose faces are fully covered with masks. People with completely covered faces are assumed to be untrustworthy.

VI. RESULTS

After completion of implementation according to the Experimental setup. Go to the location of the code and we have set the path for both videos and frames. Now select the path for the command prompt to run the python file name AnomalyDetection.py in the cmd. After running the python file, it creates a user interface for uploading CCCTV videos. Take a video from our data set that we selected video file. We test every video to detect normal and also train our code that can precisely increase its capacity to find abnormal/anomaly events. Every video creates its frames as in Fig1, the frames are saved by continuously testing every video it helps to find the video whether s are normal as in Fig2 and abnormal as Fig 3.

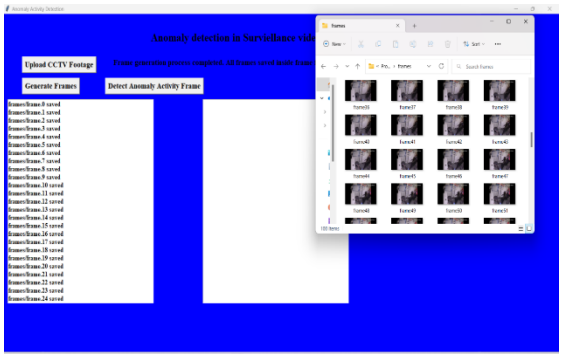

Fig1. Frames and Mining and saving

Fig2. Detected as an Anomaly video

Fig3.Detected as Normal video

VII. CONCLUSION

Nearly everyone in the modern world is aware of the value of CCTV video, yet in many situations, it is only used for investigative purposes after a crime or incident has occurred. The benefit of the suggested paradigm is that it prevents crime before it occurs. The live CCTV video is being monitored and examined. If the analysis's conclusion predicts an undesirable incident will occur, the appropriate authority is instructed to take action. Thus, it is possible to stop this. Although the suggested approach is only applicable to academic settings, it may be used to anticipate more shady activities in both public and private settings. Every scenario where training has to be supplied with suspicious activity appropriate for that scenario can utilize the model. By distinguishing the suspect person from the suspicious action, the model may be enhanced.

VIII. REFERENCES

[1] Xiaoyu Zhang, et al. (2019“Learning Spatiotemporal Features for Pedestrian Detection in Surveillance Video Analysis’’.

[2] Rui Zhao et al. (2020)."Anomaly Detection in Time Series Sensor Data Using Variational Autoencoder with Reconstruction Probability".

[3]Kaveh Hassani et al. (2018)."Anomaly Detection in ECG Signals Using Deep Autoencoder Neural Networks".

[4] Fabrizio Riguzzi et al. (2019)."Deep Anomaly Detection in Complex Industrial Processes Using LSTM Networks".

[5] Shaoqing Ren et al. (2020) "Unsupervised Anomaly Detection in Surveillance Videos Using GANs and Spatial Temporal Autoencoder".

[6] Zhou et al. (2019). "Deep Learning for Anomaly Detection: A Survey"

[7] Dinesh Jackson Samuel R, Fenil E, Gunasekaran Manogaran, Vivekananda G.N, Thanjaivadivel T, Jeeva S, Ahilan A, (2019) “Real-time violence detection framework for football stadium comprising of big data analysis and deep learning through bidirectional LSTM”, The International Journal of Computer and Telecommunications Networking.

[8]Kwang-Eun Ko, Kwee-Bo Sim(2018)“Deep convolutional framework for abnormal behavior detection in a smart surveillance system.”Engineering Applications of Artificial Intelligence,67.

[9]Dinesh Kumar Saini, Dikshika Ahir and Amit Ganatra, (2019) “Techniques and Challenges in Building Intelligent Systems: Anomaly Detection in Camera Surveillance”, Proceedings of First International Conference on Information and Communication Technology for Intelligent Systems, Springer International Publishing Switzerland.

[10]W. Li, V. Mahadevan, and N. Vasconcelos. (2020) “Anomaly detection and localization in crowded scenes”. TPAMI.

19. Shean Chong, Yong Haur Tay, Yong, (2020) “Modeling Representation of Videos for Anomaly Detection using Deep Learning: A Review”, arXiv:1505.00523v1.

[11]Yuke Li (December 2018) “A Deep Spatiotemporal Perspective for Understanding Crowd Behavior”, IEEE Transactions on Multimedia, Vol. 20, NO. 12.

[12]Javier Abellan-Abenza, Alberto Garcia-Garcia, Sergiu Oprea, David Ivorra- Piqueres, Jose Garcia-Rodriguez (December 2017) "Classifying Behaviours in Videos with Recurrent Neural Networks", International Journal of Computer Vision and Image Processing.

[13]Zahraa Kain, Abir Youness, Ismail El Sayad, Samih Abdul-Nabi, Hussein Kassem, (2018) " Detecting Abnormal Events in University Areas ", International conference on Computer and Application.

[14]Tian Wanga, Meina Qia, Yingjun Deng, Yi Zhouc, Huan Wangd, Qi Lyua, Hichem Snoussie,(January-2018) "Abnormal event detection based on analysis of movement information of video sequence" ,Article-Optik,vol-152

[15]Joey Tianyi Zhou, Jiawei Du, Hongyuan Zhu, Xi Peng, Rick Siow Mong Goh, (2019) "Anomaly Net: An Anomaly Detection Network for Video Surveillance", IEEE Transactions on Information Forensics and Security, 1(1), pp. 99-105