Detecting Human Suspicious Behavior Using CNN for Video Surveillance

1Shyam Sunder Reddy Dodla, 2Devi Priya Ramini

1Computer Science and Engineering, 2Information Technology
BV Raju Institute of Technology, City, Narsapur, India

Abstract- Due to a rise in antisocial activity, security has recently been given primary importance. Several organizations have installed CCTVs to continuously watch over people and their interactions. In a contemporary nation of 64 million people, every individual is photographed 30 times every day. Large amounts of video data are generated and stored for a predetermined period of time. An image with a resolution of 704x576 and a frame rate of 25 will generate around 20GB each day. Since it requires a workforce and constant attention, determining whether the events are abnormal is nearly impossible for people to do while continuously monitoring data. This necessitates automating the same procedure. Additionally, it is important to demonstrate which frame and which regions of it exhibit aberrant activity. This will aid in determining the activity’s abnormality more quickly. To do this, a video is divided into frames, and the individuals and activities in each frame are examined using processing software. In a broad sense, deep learning and machine learning algorithms and techniques assist us in making.

Index Terms- antisocial, aberrant activity, abnormality, automation.

Introduction

Human face and human behavioural pattern play an important role in person identification. Visual information is a key source for such identifications. Surveillance videos provide such visual information which can be viewed as live videos, or it can be played back for future references. The recent trend of ‘automation’ has its impact even in the field of video analytics. Video analytics can be used for a wide variety of applications like motion detection, human activity prediction, person identification, abnormal activity recognition, vehicle counting, people counting at crowded places, etc. In this domain, the two factors which are used for person identification are technically termed as face recognition and gait recognition respectively. Among these two techniques, face recognition is more versatile for automated person identification through surveillance videos. Face recognition can be used to predict the orientation of a person’s head, which in turn will help to predict a person’s behaviour. Motion recognition with face recognition is very useful in many applications such as verification of a person, identification of a person and detecting presence or absence of a person at a specific place and time. In addition, human interactions such as subtle contact among two individuals, head motion detection, hand gesture recognition and estimation are used to devise a system that can identify and recognize suspicious behaviour among pupil in an examination hall successfully. This paper provides a methodology for suspicious human activity detection through face recognition.

Video processing is used in two main domains such as security and research. Such a technology uses intelligent algorithms to monitor live videos. Computational complexities and time complexities are some of the key factors while designing a real-time system. The system which uses an algorithm with a relatively lower time complexity, using less hardware resources and which produces good results will be more useful for time-critical applications like bank robbery detection, patient monitoring system, detecting and reporting suspicious activities at the railway station, etc

Manual monitoring of exam hall through invigilators and manual monitoring of exam hall through surveillance videos is performed throughout the world. Monitoring an examination hall is a very challenging task in terms of man power. Manual monitoring of examination halls may be prone to error during human supervision. Such a system when implemented as an ‘automatic suspicious activity detection system’ will not only help in detecting suspicious activities but also helps in minimizing such activities. Moreover, the probability of error will be much lesser. This system will serve as a useful surveillance system for educational institutions.

This paper describes a technology in which real time videos are analysed and are used for human activity analysis in an examination hall, thus helping to classify whether the particular person’s activity is suspicious or not. The system developed identifies abnormal head motions, thereby prohibiting copying. It also identifies a student moving out of his place or swapping his position with another student. Finally the system detects contact between students and hence prevents passing incriminating material among students. In our research, we have contributed upon a system that will intellectually process live video of examination halls with students and classify their activities as suspicious or not.
research proposes an intelligent algorithm that can monitor and analyse the activities of students in an examination hall and can alert the educational institute’s administration on account of any malpractices/suspicious activities.

The Suspicious Human Activity Detection system aims to identify the students who indulge in malpractices/suspicious activities during the course of an examination. The system automatically detects suspicious activities and alerts administration.

![Diagram of the Suspicious Human Activity Detection system]

**Fig 1: Example Process**

**LITERATURE SURVEY:**
Data monitoring following image sensor capture spans a wide range of topics. They have been used in practice by many scientists. various image processing and anomaly detection techniques. When it comes to video surveillance, there is a substantial evaluation of the literature with numerous authors. Below is a simple explanation of the same.

A sparse coding method was proposed by Joey and others [1] for analyzing video. Constructed anomaly was the name given to utilized methodology. a detecting method that has performed more effectively. It is suggested to use a unique neural network to detect anomalies. Which is also referred to as an anomalous network because it successfully completes feature learning, sparse representation, and dictionary, three interrelated neuronal learning building components. The authors, in particular, developed a motion fusion to learn enhanced features. To benefit from the advantages of eliminating background noise, catching motion, and lowering the shortage of data, block should be accompanied with a feature transfer block.

Suspicious behavior is any observation of conduct that would suggest a person is either committing or planning to commit a criminal crime in some circumstances. Anomaly detection is a technique for spotting unusual behaviour. The camera is one of the best surveillance gadgets, a fix for the various sites’ security issue. Human resources are needed for both detection and monitoring in modern systems. Moreover, it is very challenging to tell the difference between criminal and aberrant behaviour.

Jefferson and Andreas [3] worked on Predictive Convolutional Long Short-Term Memory Networks. A plan of action was suggested. It is challenging to automate the detection of anomalous actions within extensive video series due to the unpredictability of how such. Events are predetermined. Studying generative models that can spot anomalies in videos helped the authors identify a solution, little supervision. The end-to-end trainable Convolutional Long Short-Term Memory (Conv-LSTM) is a complicated system. the network that, given a short number of input frames, can predict how a video sequence would unfold.

Luo and co. [4] offered a useful method for finding anomalies in videos. Inspired by this ability, the authors wrote this. Temporally coherent Sparse Coding (TSC), which they implement similarly, is based on sparse coding and is projected for suspicious detection. Similar reconstruction coefficients are used to encrypt adjacent frames. then used a particular
kind of stacked to map the TSC, neural network with recurrence (sRNN). The paper's contributions are as follows: (i) it proposed a TSC that can be recorded to a sRNN which speeds up the shaky prediction and facilitates parameter optimization. Building a very sizable dataset that is even is option two, greater than the total size of all currently available datasets for detecting anomalous activity.

Using the videos, Chong and Tay [5] demonstrated an effective method for spotting anomalies. Applications in the recent past. Convolutional neural networks have demonstrated the potential of convolutional layers for object recognition and detection in particular. Picture form. However, convolutional neural networks require labels as learning signals because they are supervised. Writers as well. In videos with crowded scenes, a spatiotemporal architecture has been proposed for suspicious detection. End-to-end trainable complex Convolutional Long Short-Term Memory (Conv-LSTM) networks were proposed by Medel and A. Savakis [6]. These Conv-LSTM networks are able to predict a video sequence's evolution from a small number of input frames. Consistency scores are created by subtracting the renovation errors of a set of estimates from a set of irregular video sequences. Lower regularity scores over time as they drift further from the real sequence. The models use a composite structure and take note of the unique effects of conditioning in acquiring more accurate representations.

Hasan et al. [7] the method for solving the issue of anomaly detection by developing a generative model for dependable motion patterns. Using a variety of resources under very limited supervision. The paper specifically includes two methods that are based on the, because they can operate with little to no supervision, autoencoders. Utilizing traditional handcrafted methods is the first strategy, study the fully connected autoencoder after studying the spatiotemporal local features. The second step is to build a fully convolutional feedforward autoencoder to learn the classifiers and local features as one ua suggested unit. The regularities from many datasets can be captured by the model.

The method for real-time anomaly localization and detection in crowded scenes was proposed by Sabokrou et al. [8]. Every video uses two local and global descriptors to be explained and is well-defined as a collection of non-overlapping cubic spots. The descriptive terms captured the video assets from various stages and are used here. We can by incorporating straightforward and economical Gaussian classifiers. Recognize regular occurrences and anomalies in video.

Lu and co. [9] proposed the method based on the built-in redundancy of video structures. The authors offered a practical. Volume 13 Special Issue 1 of the 1237 Journal of Pharmaceutical Negative Results, 2022 on the framework for sparse combination learning. Without compromising, it achieves a respectable performance in the detection phase. Outcome caliber. The short running time is fail-safe because the new method efficiently turns the original complex problem to one where only a few low-cost small-scale least squares optimization steps are taken into account. High detection rates are possible through the process when calculating benchmark datasets using MATLAB programming on a typical desktop PC.

Mousavi et al. [10] suggested the idea of using a dynamic sparse coding methodology for identifying the unusual that is completely untested. On the basis of online sparse re-constructability of query signals, it is suggested to identify events in videos. Based on the impression that commonplace events, whereas infrequent events are not, the algorithm prefers that frequent events in a video be re-constructible from an event dictionary, and employs a principled convex optimization formulation that allows for both a sparse reconstruction code and an online dictionary to be updated and mutually inferred.

However, there are a number of ways to identify anomalies in security applications [11–14] that are included in our suggested model. We have created automated video screening, which finally distinguishes between normal and suspended activities. In the system any suspicious activity is detected, an alarm is sounded.

**METHODOLOGY**

Manual exam hall monitoring is carried out employing invigilators and manual test hall monitoring using surveillance footage anywhere in the world. Monitoring an exam room requires a lot of labor, which makes it a very challenging task. Manual monitoring in exam rooms conducted by humans may be inaccurate. Such a system can be utilized as a "automated suspicious activity detection system," which will help both identify and curtail suspicious activities. Additionally, there will be a lot less potential of error. As a surveillance technique, this strategy will be useful for educational institutions.
The technique used to analyze real-time videos and use it to categorize human activity in a testing environment to evaluate whether or not it is suspicious is discussed in this work. The intended technology stops copying by detecting abnormal head motions. Additionally, it shows when a student changes places with another student or vacates his position. The system also detects student contact, which stops the spread of negative information among pupils. In our research, we have made contributions to a system that can watch video of test rooms full of students in real time and determine whether or not their behavior is suspicious. This study presents an intelligent program that can monitor and assess students' performance throughout testing and may provide the administration of the educational institution of any malpractices or questionable behaviour.

The goal of the suspicious human activity detection system is to locate pupils who engage in unethical or suspicious behaviour during an exam. The technology automatically recognises unusual activity and notifies management.

![System Architecture](image)

Python, Django, MySQL, and Wamp Server are required software. CNN's procedures are as follows:

Step 1: Image or video input is provided.

Step 2: Many filters are then applied to the input to generate a feature map.

Step 3: To boost non-linearity, a ReLU (Rectified LinearUnit) function is used.

Step 4: Next, apply a pooling layer to each feature map. Step 5: The algorithm merges the merged pictures into a single long vector.

Step 6: The vector is then sent into the algorithm, which creates a fully linked artificial neural network.

Step 7: The features are processed through the network. Finally, the completely linked layer provides the "voting" of the classes.

Step 8: In this final stage, training is carried out over a number of epochs via forward and reverse propagation. This process is repeated until a well-defined neural network with training weights and feature detectors is obtained.
I. IMPLEMENTATION
Applications for identifying human behavior can be found in many real-world settings, including intelligent video surveillance and analysis of consumer behavior. Video surveillance has many applications, notably in indoor, outdoor, and public settings. Surveillance is an essential part of security. Security cameras are a necessity in the modern world for both safety and security. E-surveillance is one of the main objectives of the Indian government's development program, Digital India. In some way, video surveillance is still a part of it. The advantages of video surveillance include effective monitoring, the need for less employees, cost-efficient auditing capabilities, adoption of new security trends, etc. Tracking has traditionally been done by hand. Dealing with much visual information could rapidly weary people, and manual intervention could lead to errors. When it comes to the system is significantly impacted. Video surveillance automation has provided a solution for this.

Currently, every occurrence that a CCTV (Closed Circuit Television) camera records cannot be manually inspected. It takes time to manually search for the incident on the recorded video, even if it has already happened. The analysis of anomalous occurrences from video is a growing subject in the realm of automated video surveillance systems. Human behavior detection in video surveillance systems is an automatic means of determining any suspicious behavior. There are several efficient algorithms available for the automatic identification of human behavior in public contexts like airports, train stations, banks, businesses, exam rooms, etc.

Video surveillance is a new application for artificial intelligence, machine learning, and deep learning. Computers can think like people thanks to artificial intelligence. Machine learning relies heavily on learning from training data and creating predictions based on future data. Due to the availability of huge datasets and GPU (Graphics Processing Unit) processors, deep learning is now used. Computer vision and video surveillance will be used to provide public safety and security. Computer vision techniques comprise the following steps: environment modeling, motion detection, object classification, tracking, behavior interpretation, and description and combining of data from numerous cameras. This approach requires substantial pre-processing to extract features from different video sequences. Categorization that is both supervised and unsupervised There are techniques. Unlike supervised classification, which uses manually marked training data, unsupervised classification is totally computer-operated and does not require any human intervention.

II. EXPERIMENTAL RESULTS

Fig 3: Home Screen
FIG 4: UPLOAD CCTV FOOTAGE

Fig 5: Generate Frames

Fig 6: Frames
CONCLUSION
Almost everyone in the modern world is aware of the importance of CCTV footage, but most of the time, it is only used for investigations after a crime or incident has occurred. The suggested model has the advantage of stopping crime before it starts. CCTV footage is being watched and studied in real-time. If the analysis’s findings indicate that a regrettable incident may occur, the relevant authority is told to intervene. Consequently, this can be stopped. Despite being limited to the academic field, the suggested approach can be utilized to predict more suspicious behaviors in public or private situations. The model can be applied to any case where guidance should be given utilizing the dubious behavior that is appropriate for that circumstance. By distinguishing the suspect person from the suspicious action, the model can be enhanced.

REFERENCES: