IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Artificial Intelligence Based Enhancing Automated Video Surveillance For Violence Detection

[1]Ramesh Kumar M, [2]Dilip T U, [3]Dhanabalan K, [4]Keshavaraj Sharmaa R

[1]Professor, [2]Student, [3]Student, [4]Student

Department of Computer Science and Engineering,

Paavai Engineering College,Namakkal, Tamilnadu, India

Abstract: one of the main objectives of the study and application of video surveillance is the detection of anomalous events. Surveillance cameras are being used more often in public areas such as roadways, crossroads, banks, and shopping centers in an effort to increase public safety. One of the most important tasks in video surveillance is recognizing anomalous occurrences, such as crimes, traffic accidents, or illegal behavior. Abnormal events usually occur significantly less frequently than normal activities. The goal of a workable anomaly detection system is to identify the window of time when the anomaly is occurring and to rapidly notify users of behavior that deviates from expected patterns. Consequently, anomaly identification can be considered a fundamental method of comprehending videos by distinguishing abnormalities from typical patterns. After an anomaly has been discovered, it can be categorized into one of the specialist activities by applying classification techniques. An overview of anomaly detection is provided in this article, with a focus on applications related to banking operations. A wide range of stakeholders, including employees, clients, debtors, and outside parties, participate in or are impacted by the routine, recurrent, and ad hoc activities and transactions involved in banking operations. Although things could take time to work out, early detection can significantly reduce any potential bad effects and, in certain cases, even completely avoid them. Time series based anomaly detection is used to locate persons in undesired times. This paper identifies both common and uncommon events using an anomaly detection technique based on machine learning. The biometric identity of the face is captured and cross-referenced with faces of well-known criminals. If a match is found, it is easy to identify and capture the offenders.

Index Terms - Abnormal Event Detection, Movement Detection, Gaussian Mixture Model, Face Detection, Haar Cascade Algorithm, Criminal Detection, Alert Sending, Image Sharing through Email.

I. INTRODUCTION

Video surveillance is one of the current study topics in image processing. When video surveillance first started, analogue CCTV systems were used to record data and monitor people, places, and activities. Current digital video surveillance systems rely entirely on human operators to detect dangers; they only provide the technology required to collect, store, and disseminate video. Manually reviewing security video is a laborious task. Finding several activities in real-time video requires a lot of work when done by hand. As a result, an intelligent video surveillance system was created. For security purposes, the analytics programme automatically recognizes events and objects (people, cars, and equipment) of interest by analyzing video flow images. The activity of monitoring or evaluating a specific area for business and security purposes is known as video surveillance. Surveillance camera installations are motivated by worries about safety and preventing crimes. Video security cameras are used in public areas, retail stores, banks, ATMs, and banking institutions. Research on network surveillance is growing all the time nowadays. The cause is the unstable times that are

occurring everywhere in the world. For intelligent monitoring, a smart surveillance system is therefore necessary. Real-time data collection, transmission, processing, and comprehension of information related to the monitoring targets are all requirements for this system. After a crime is committed, video evidence might be subjected to forensic examination. The affordability of video cameras has resulted in a rise in the utilization of video surveillance systems. Video surveillance systems have several uses, such as detecting human activity and monitoring traffic. Here's an example of how real-time activity analysis for video surveillance systems can be used to create content-based searches and alerts in real-time as events happen in the monitored area. In practice, the term "multi-view face recognition" refers exclusively to situations where multiple cameras are recording a subject (or scene) simultaneously, and an algorithm uses the collected images or videos in concert. Still, the expression has been used a lot to recognize faces in different positions. This ambiguity has no effect on (still) picture recognition; a set of photos taken simultaneously with separate cameras and a set of photos taken with the same camera but at different view angles are comparable in terms of posture variations. However, when it comes to video data, the two cases are different. A multi-camera system can always guarantee the acquisition of multiple viewpoints, but the possibility of achieving this with a single camera is not always assured. These kinds of distinctions become critical in non-cooperative recognition applications such as surveillance. Still, single-view recordings are used by most multi-view video facial recognition algorithms now in use. After being given two face photos to compare, they search through the collection to "align" the appearance of a facial component in one image with the same attitude and lighting in the other. Estimating the postures and lighting conditions will also be required for the two face pictures in this approach. Using the "generic reference set" approach, the holistic matching algorithm was also developed. It relies its matching measure on the ranking of look-up results. Additionally, some works calculate the pose implicitly, without explicitly accounting for stance changes.

1.2 VIDEO PROCESSING:

A video signal is essentially any set of images that change over time. A still image has a spatial intensity distribution that remains constant over time, however a time-varying image's spatial intensity distribution varies with time. Video signals are thought of as a set of images, or frames. It is possible to create the illusion of continuous video by rapidly varying the frame rate, commonly referred to as the frames. Digital video is becoming more and more necessary in sectors including education, multimedia authoring systems, video teleconferencing, and video-on-demand systems.

1.2.1 FRAME TYPES:

Video frames come in three main varieties: I-frame, P-frame, and B-frame. The symbols 'I' denote intracoded frame, 'P' indicates predictive frame, and 'B' indicates bidirectional predictive frame. Anticipated 'I' frames that are encoded without any motion correction are projected 'P' and 'B' type frames in the future. 'I' frames, on the other hand, take a lot of bits to encode. 'P' frames are decoded from a reference frame (which may be either a 'I' or a 'P' frame) using motion-compensated prediction. 'P' frames require more bits than 'B' frames, although being more sparing with the bits than 'I' frames. 'B' frames are computationally complex, but require less bits than 'I' and 'P' frames. The frames that follow the initial "I" frame and the next two "I" frames are referred to as a group of photographs (GOP). Figure 3 shows the Republican Party. The illustration consists of one "I" frame, two "P" frames, and six "B" frames. Usually, there are a lot of "B" frames positioned in between "P" frame pairings or between "I" and "P" frames. Features like random access, rapid forward, and regular and fast reverse playback are made easier to implement with GOPs. Video processing technology has revolutionized the multimedia industry with gadgets like the Digital Versatile Disc (DVD), Digital Satellite System (DSS), high definition television (HDTV), and digital still and video cameras. Among the several methods for processing videos are (i) segmentation (iv), indexing (iii), compression (v) and tracking (v).

1.3 ABNORMAL ACTIVITY DETECTION

Recently, anomalous detection-based machine learning has emerged as an intriguing field in video analysis. This makes sense in real-time video surveillance (RVS), where various motion patterns that occur periodically or often in a congested traffic scene are displayed. Analyzing the motion pattern and deriving a high-level understanding is the goal. If an incident seems improbable or unanticipated, it is deemed abnormal. An "anomaly" in statistics is an outlier data point in a data space or odd behavior in a distribution. Naturally, in RVS, behavior understanding is a subdomain of anomalous detection. After training on a normal dataset, anomaly detection algorithms interpret an outlier as aberrant behavior. Typically, these systems work together to identify and pinpoint unusual behavior. The location of an anomalous event can be found by identifying the anomalous pixels in each frame. An efficient data representation technique is needed for AED to respond

quickly and with great accuracy. In AED, data space is spatiotemporal, involving both movement and appearance. To find unusual regions in a video file, researchers employ a variety of techniques. The most often used method is called "gridding," which involves imprinting a predetermined grid on a sequence of frames to split them into smaller, fixed-size 3D patches. Any deviation from the typical behavior is referred to as an anomaly. Training is used to teach abnormal detection skills from a normal pattern. Essentially, it is achieved by applying a few specific learning strategies to the extracted attributes. A raw video (series of frames) is the primary input used for abnormal detection in video surveillance, from which relevant features are extracted. The proposed work is based on pattern-learning techniques and pixel-level features. Numerous issues, including occlusion, shadowing, and object overlap, arise while analyzing crowd scenes. To address this issue, several algorithms have been put forth; however, each has pros and cons of its own. There are multiple ways to interpret the scene parts' movements. These techniques can accurately simulate the direction and velocity of every single object. These techniques, however, typically take a long period. The perspective distortion of urban surveillance film, which produces different scale and movement patterns based on item placements and camera position, further adds to the complexity of the problem. Because different lighting circumstances and subtle distinctions between normal and abnormal cases exist, an appropriate discriminative model is needed to detect abnormal patches and frames. A significant amount of labelled data is needed for many machine learning models, including deep neural networks. However, collecting a large amount of tagged data for AED is a laborious and time-consuming operation because it is an unsupervised learning problem. Foreground modeling, adaptive learning, classification, and classifier are some of the key approaches for anomaly detection in RVS. These approaches have drawbacks, including an effective clustering center, imbalanced data (normal and abnormal occurrences), and a large number of objects in the traffic scene.

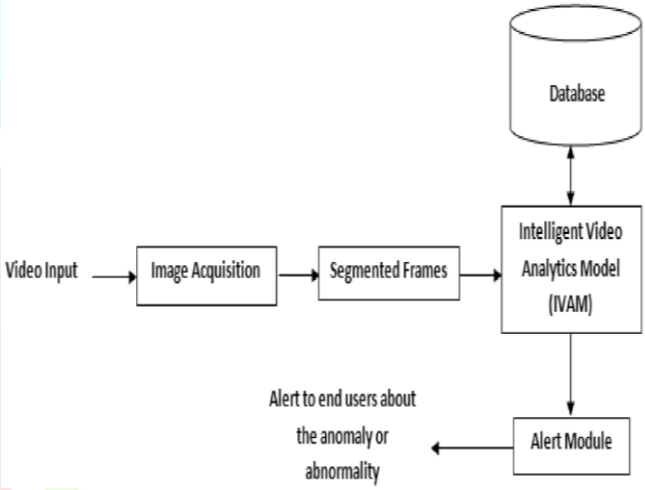


Fig.1 Video Anomaly Detection

II. RELATED WORKS

Ahmed abbasi, Abdul rehman javed, et.al,…[1] function as an essential tool for enhancing individual safety and safeguarding both private and public property. In order to classify odd sound events and detect anomalous audio, a bespoke dataset was created by fusing 15 background audio samples from the TUT Acoustic Scenes 2016 dataset with rare occurrences. In order to identify anomalies in the audio data, this study explored with a number of audio data features. This study uses the PCA feature selection technique to select the fewest number of best-performing features for optimal performance after extracting several features for feature engineering from the audio stream. Several machine learning techniques are applied to identify seven distinct anomalous events in fifteen different backdrop situations based on the selected feature set. Experiments revealed that the suggested strategy outperformed state-of-the-art research in every instance for the identification of anomalies in audio data. With the advent of new digital technologies, a notable rise in the amount of multimedia data generated by different smart devices has been seen. To obtain useful information from multimedia data, several challenges have emerged for data analysis. Finding anomalies in multimedia

data quickly and accurately is one of these challenges. This paper proposes an efficient way to identify anomalies and categorize infrequent events in audio data.

Lin Wang, et.al,…[2] ConvLSTM and VAE are combined in the DF-ConvLSTM-VAE model, which is used to learn the training data distribution for video anomaly detection. The ConvLSTM-VAE (Asymmetric) model is also suggested. ConvLSTMVAE (Asymmetric) model is created by weakening the decoder. In terms of training time and difficulty, the ConvLSTM-VAE (Asymmetric) model has several advantages over the ConvLSTM-VAE (Symmetric) model. Experiments show that the DF-ConvLSTM-VAE model outperforms the ConvLSTM-VAE (Asymmetric) model. Experiments conducted on many publicly accessible benchmark data sets verify the validity and competitiveness of the suggested DF-ConvlSTM-VAE in relation to other traditional methods. The security sector has recently shown a great deal of interest in video anomaly detection because of the market's rapid expansion in video surveillance sites. There is currently an unequal distribution of normal and aberrant data in unlabeled video footage. Variational autoencoder (VAE) is a common deep generative model that is gaining popularity in unsupervised anomaly identification. However, this paradigm has trouble processing time-series data, especially video data. Furthermore, a lot of autoencoder-based systems fail to identify anomalies because of their strong generalization ability, which over-reconstructs abnormal behavior.

Keval Doshi, et.al,…[3] offered a fresh, large dataset and a unique framework for further education in the field of video anomaly identification. It is anticipated that future VAD research will focus on practical and repeatable solutions as a result of the modified problem statement (Sec. 3) and updated dataset (Sec. 4). This also showcased a state-of-the-art video anomaly detector that is capable of learning continually through experience playback and incremental learning. Here, extensive testing on the suggested NOLA dataset and available benchmark datasets show that the suggested method outperforms two of the state-of-the-art methods in continual learning as well as in terms of the common frame-level AUC metric. The anomaly in the first case was a person carrying a snake while strolling down a crowded street. In the third, a couple is seen arguing with the restaurant's owners. To find such anomalies, a VAD algorithm needs a far deeper understanding of the intricate relationships between each detected object and how it affects its surroundings. But this also shows how comprehensive the proposed NOLA dataset is and how it could be applied to improve further VAD algorithms.

Mohana, et.al,…[4] created a smart surveillance system for security applications that allows for real-time object tracking and detection. The majority of domains currently use a manually operated mechanism that is too laborious to execute effectively in real-time. To increase its efficiency, it is now necessary to design an automated system that runs in real time and without the assistance of a human. The VHDL programming language and the Xilinx ISE software are used to develop the updated background subtraction technique. The Zynq XC7Z020 FPGA board is utilized, and an OV7670 camera is used to record the footage. Because FPGA systems offer both temporal and spatial parallelism and allow for module design flexibility according to system requirements, they are five to six times more efficient than DSP boards. Since FPGA contains parallel processing of algorithms, researchers are currently focusing on detecting and tracking anomalous behavior in persons or objects concurrently. Increasing the effectiveness of object classification using contour data. Following appropriate classification, it will be simple to identify anomalous behavior in large populations, which greatly strengthens the case for efficient surveillance. After that, the system will be integrated with artificial intelligence so that monitoring of any part of the surveillance process won't need human intervention. The idea of partial face detection with several key point descriptors can also be incorporated into the system to improve the surveillance system's overall effectiveness.

Vinaya Keskar, et.al,…[5] have determined the credit card disparity, wherein the erroneous information number was chosen on the basis of this component, and have demonstrated the methods for elimination. Many financial industries are currently using big data analysis to help them improve the services they offer to their customers, both internal and external, and to help them create both dynamic and passive security measures. While it is certainly more profitable and successful to undertake input validation with an eye toward the customer, employees should be able to accomplish this as well. Applications that operate in real-time must quickly analyze transient data sources in order to make decisions or take action in a timely manner. If decisions are made after the deadline has passed, they are no longer relevant. Because of the development of big data, examiners can now quickly request their results from massive datasets.

III. Existing Methodologies

In the field of pattern recognition and computer vision, the identification of anomalous events has become more and more important in recent years. The main difficulty is coming up with a range of scenarios that show anomalous happenings. It is difficult to define an interface that includes the whole spectrum of possible anomalous events. It is common practice to categorize anomalous events as low likelihood occurrences relative to normal occurrences in order to statistically interpret them. Disordered thinking events are behaviors that deviate from norms and are inconsistent with samples. The detection process used to find anomalous events can be broadly divided into two stages: event encoding and anomaly detection model. These steps correlate to the most widely accepted ideas in the machine vision and pattern recognition domains. The relevant elements from the video must be selected in order to depict an event. Because of the ambiguity in the event description, the event may be specified by features at the pixel or object level. Anomaly identification in videos remains a challenging task because of the imprecise definition of an anomaly and the complexity of visual conditions in real video data. Our research investigates a novel permutation auto-encoder architectural design that can detach the spatio-temporal portrayal to independently encapsulate the remote sensing data and the time data, since abnormal events often differ from normality in appearance and/or motion behavior. This contrasts with earlier research that attempted to grasp the temporal regularity by either reconstructing the data or utilizing predictions as a support assignment. Specifically, the temporal auto encoder uses the input of the first four consecutive frames and the output of the RGB difference to effectively replicate the movement of optical flow, while the spatial auto encoder learns to recreate the insight of the first frame (FIF) to model the data is normally distributed only on appearance feature space.

IV. ABNORMAL ACTIVITY DETECTION IN BANKING

Anomaly detection is the process of identifying patterns in data that do not correspond to expected behavior. These non-conforming patterns are sometimes referred to in various application fields as anomalies, outliers, discordant finds, exceptions, aberrations, surprises, oddities, or contaminants. Anomalies and outliers are the two terms most commonly used in the context of anomaly identification, while they are sometimes used synonymously. Applications for anomaly detection are numerous and include defect identification in safety-critical systems, intrusion detection for cyber-security, fraud detection for credit cards, insurance, or healthcare, and military surveillance for adversarial activity. Surveillance camera installations are motivated by worries about safety and preventing crimes. Businesses, public spaces, banks, ATMs, and banking institutions all use video security cameras. This includes shopping malls. Research on network surveillance is growing all the time nowadays. The cause is the global instances of instability that are occurring. Therefore, the deployment of a smart surveillance system is required for intelligent monitoring. Real-time data collection, transmission, processing, and comprehension of information related to the monitoring targets are all requirements for this system. After a crime is committed, video evidence might be subjected to forensic examination. Because of this, these systems offer strong security in public spaces, which is usually a very challenging issue. The affordability of video cameras has resulted in a rise in the utilization of video surveillance systems. Video surveillance systems have several uses, such as detecting human activity and monitoring traffic. Here's an example of how real-time activity analysis for video surveillance systems can be used to create content-based searches and alerts in real-time based on activities that happen in the monitored area. A facial recognition system is a computer program that can identify or verify an individual from a digital picture or a video frame from a video source. Comparing specific facial features in the image with a database of faces is one method to achieve this. Facial recognition systems based on face prints can accurately and quickly identify target individuals when the conditions are met. Numerous studies on face recognition are being developed for smartphone apps.

METHODOLOGY

Gaussian Mixture Model

Let us denote with $x = [x_1, x_2, \ldots . x_L]^T$ the generic image pixel, L being the number of sensor channels. Anomaly detection is viewed as a statistical binary decision problem, where by observing x one must decide if it is a background pixel (H_0 hypothesis) or a target pixel (H_1 hypothesis). Further, it is assumed that background consists of N different clusters C_i (i = 1,2,..., N) corresponding to different ground cover types. The generic cluster $C_i$is modeled as Gaussian distributed and its p.d.f. f_{c_i} is:

$$f_{c_i}(x) = f_G(x; \mu_i, \Gamma_i)$$

$f_G(x; \mu_i, \Gamma_i)$ denotes the multivariate Gaussian p.d.f. with mean vector μ_i and covariance matrixΓ_i.

HAAR Cascade algorithm:

The high-frequency components in the facial data that can be added to a low-resolution input image to produce a super-resolved image are first learned in this novel approach to super resolution. The suggested technique is different from traditional methods in that it estimates the high-frequency components that are not used in other ways to rebuild a higher-resolution image, rather than analyzing the direct relationship between the high- and low-resolution images. The codes for the HR and LR photos are X and Y, correspondingly. Here, bold uppercase indicates the transformation matrices (sometimes called projection matrices). To be more precise, the dual-tree complex wavelet transform in matrix form is represented by Ψ. Vectors are shown by bold lowercase letters. Regular matrices are represented by simple uppercase letters; the matrix representation of L, for instance, using a down sampling technique. Scalars are simply lowercase letters. The training is indicated by the superscript t. Integral pictures function as lookup tables and are two-dimensional matrices the same size as the original image. Each element of the integral image contains the total of all the pixels in the upper-left corner of the original image (relative to the element's position). This enables the sum of rectangular regions in the image to be calculated at any scale or position using only four lookups: Sum = I(C) + I (A) –I(B)- I(D) where points A,B,C,D belong to the integral image. Super-resolution image reconstruction is the process of combining low-resolution (LR) images into one high-resolution image. These aliased, low-resolution pictures are connected to one another by sub-pixel shifts and serve as different snapshots of the same scene with further data. The relationship between the ideal high-resolution (HR) image and the observed LR images can be described by the following observation model, yk = DBkMkx + nk,

where yk denotes the k = 1 ...p LR images, D is a subsampling matrix, Bk is the blur matrix, Mk is the warp matrix, x is the ideal HR image of the scene which is being recovered, and nk is the additive noise that corrupts the image. Mk can be represented by anything from a straightforward parametric transformation to motion flow fields, while D and Bk replicate the averaging procedure carried out by the camera's CCD sensor. In essence, an inversion procedure can be used to retrieve x given multiple yk. However, because there are many pixel values to be estimated from a limited number of known pixels, the problem is typically ill-posed.

Haar Cascade algorithm is a popular machine learning-based approach for object detection in images or video streams. Here are the main steps involved in the Haar Cascade algorithm:

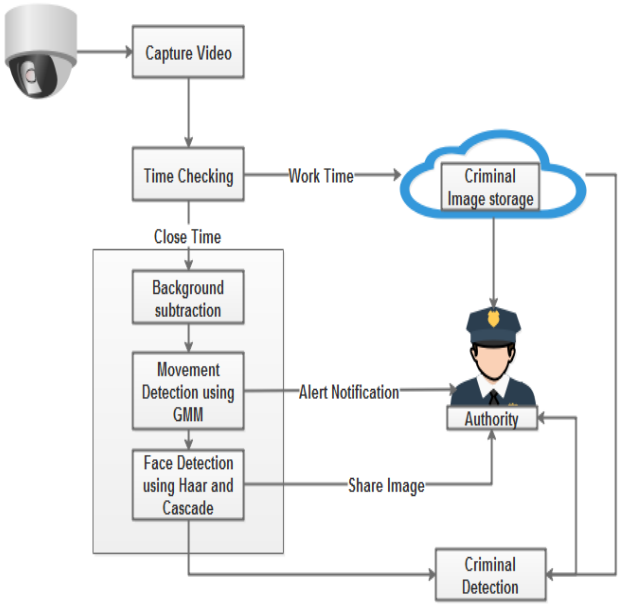
Feature extraction: Haar-like features are used to represent the object being detected. These features are computed at different scales and positions on the image.

Training the classifier: The next step involves training a classifier that can distinguish between the positive samples (images containing the object of interest) and negative samples (images without the object of interest). The most commonly used classifier is the Viola-Jones algorithm, which uses the AdaBoost algorithm to select the best features that can classify the images accurately.

Creating the Cascade Classifier: Here the classifier is divided into multiple stages and each stage consists of a set of weak classifiers. Each weak classifier is a simple classifier that uses a single Haar-like feature to make a decision. The cascade classifier is designed to reject non-object regions of the image quickly, reducing the amount of computation needed.

Detection: Once the cascade classifier is trained, it can be used to detect the object in the new image or video stream. The image is scanned with a sliding window at different scales, and the Haar-like features are computed at each location. The cascade classifier is then applied to each window, and if the object is detected, a bounding box is drawn around it

PROPOSED SYSTEM ARCHITECTURE



PROCEDURE

Video Capturing Framework

This module suggests using surveillance cameras to detect theft and track down the perpetrators. Here, image processing is used instead of sensors to identify theft and the movement of thieves in surveillance camera footage. The focus of this system is object detection. Security personnel may be notified in real-time when someone is suspected of breaking into a facility by using surveillance camera footage to analyze human activity. This allows the officers to intervene and prevent the crime from happening.

Set Time based Storage

Before recording the activities with a camera, the administrator should establish a time for predicting anomalous behavior based on an undesirable time period. This module receives data from the surveillance camera used for human detection. The surveillance system examines the timer to see how much time has passed when a user arrives before it does. The system notifies the administrator through email when the predetermined window of time for human detection is reached.

Movement Detection

The motion patterns of people are observed in front of the surveillance equipment. In the event that the system detects movement in the image, it automatically snaps a photo of the detected object and sounds the alert in accordance with user settings. Obtaining footage from security cameras is the initial stage. Those recorded video frames will be used in the motion detection process. Should motion be detected, the time stamp and the photos containing the motion detection will be saved. To determine if the activity is normal or aberrant, the gathered time value should be compared with the database. The motion value and the time threshold will be compared.

Face Identification

Real-time video is recorded as the input. Still photos are divided from video images. In the process, face detection is carried out. facial feature matching with the Haar Cascade algorithm in a database. The temporal information in video sequences enables the analysis of changes in facial dynamic characteristics and their application as a biometric identifier for person recognition. The feature vector that is used to train the classifier can be formed in a variety of ways. Some of them even perform categorization, which requires a lot of processing, using the entire image as a feature vector. In order to create a 40 value feature vector for each image, the key values from each filter—energy, mean, and standard deviation—are combined to create this feature vector. In this case, make use of the fact that people naturally make at least small movements with their mouths, eyes, and facial muscles. Since we are working with a video series, we can easily collect this information and obtain the full sequence of the object's movements. Considering that point can cut down on simulation time and eliminate errors caused by falsely detecting a human face.

Criminal Detection

The method of identifying fraudulent users and criminals is known as face categorization. A real-time camera captures the user's face image during face verification. Following that, the retrieved facial traits are compared to the database. Predicting criminals is another application of face categorization. The faces of criminals have previously been gathered and are kept in a database. If the taken picture matches the criminal database, it will be easy to analyze and predict the criminal population.

Send Alert Intimation

In a surveillance setting, automatically spotting unusual activity can be used to alert the appropriate authorities to potentially harmful or illegal behavior, like automatically reporting an individual. Unknown event alerts about certain officers are sent to the pre-defined contact numbers in the proposed system. Utilize image sharing here as well to make it simple to identify offenders.

V.CONCLUSION

The suggested remedy relies on creating a smart camera-based anomaly detection system that can identify any strange behavior by monitoring activities in banks. After that, the offenders would be found by applying motion detection algorithms and facial recognition technology based on undesirable time intervals. The security department will receive an automatic alert from the smart camera if any suspicious activity of this kind is noticed at an inconvenient moment. In order for the security to arrive with the necessary preparedness, the message not only specifies what kind of alert is issued, but it also includes the thief's face image and the time the detection was made, along with a web link to the live image.

REFERENCES

[1] Abbasi, Ahmed, et al. "A Large-Scale Benchmark Dataset for Anomaly Detection and Rare Event Classification for Audio Forensics." IEEE Access 10 (2022): 38885-38894.
[2] Wang, Lin, et al. "Unsupervised Anomaly Video Detection via a Double-Flow ConvLSTM Variational Autoencoder." IEEE Access 10 (2022): 44278-44289.
[3] Doshi, Keval, and Yasin Yilmaz. "Rethinking video anomaly detection-a continual learning approach." Proceedings of the IEEE/CVF winter conference on applications of computer vision. 2022.
[4] Aradhya, HV Ravish. "Elegant and Efficient Algorithms-Real Time Implementation of Object Detection, Classification, Tracking and Counting using FPGA Zynq XC7Z020 for Automated Video Surveillance and its Applications."
[5] Keskar, Vinaya, Jyoti Yadav, and Ajay Kumar. "Perspective of anomaly detection in big data for data quality improvement." Materials Today: Proceedings 51 (2022): 532-537.
[6] Feng, Jiangfan, Yukun Liang, and Lin Li. "Anomaly detection in videos using two-stream autoencoder with post hoc interpretability." Computational Intelligence and Neuroscience 2021 (2021).
[7] Sisodia, Shefali Shahane1 Shreya Dajgude2 Shreya, and Aditi Darade4 Mr Rahul Chakre. "Anomalous Motion Identification for Bank Surveillance." 2021
[8] Prarthana Sanas, “Anomaly Detection at ATM Center using Machine Learning Algorithm” International Journal of Advanced Research in Science, Communication and Technology (IJARSCT), 2021
[9] Ganji, Venkata Ratnam, and Aparna Chaparala. "An Efficient Pre and Post filter-based anomaly detection technique for credit card fraud detection." NeuroQuantology 20.8 (2022): 1987-2021.
[10] Chang, Yunpeng, et al. "Video anomaly detection with spatio-temporal dissociation." Pattern Recognition 122 (2022): 108213.