Original Article
ISSN (Online): 2582-7472

ShodhKosh: Journal of Visual and Performing Arts
May 2024 5(5), 1086–1093

THEFT DETECTION WITH CRIMINAL IDENTIFICATION USING MACHINE LEARNING

Dr. K. Gowsic[1], Thilagavathi R[2], Vairam Sountharya K[3], Varsha R[4]

[1] Associate Professor, Mahendra Engineering College, Namakkal
[2] UG Scholar CSE Department, Mahendra Engineering College, Namakkal

Corresponding Author
Dr. K. Gowsic, kgowsic@gmail.com

DOI
10.29121/shodhkosh.v5.i5.2024.2721

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

ABSTRACT

One of the main goals of video surveillance research and practical implementations is abnormal event detection. In order to improve public safety, the usage of surveillance cameras in public spaces—such as roadways, crosswalks, banks, retail centers, etc.—is expanding. One of the most important tasks in video surveillance is the detection of anomalous occurrences, such as criminal activity, traffic accidents, or crimes. In general, abnormal events are rare in comparison to normal activities. A useful anomaly detection system aims to pinpoint the anomaly's temporal range and instantly notify users when any behavior deviates from expected norms. Consequently, it is possible to think of anomaly identification as coarse-grained video knowledge that separates anomalies from regular patterns. Classification techniques can be used to further categories an anomaly into one of the specific activities once it has been recognized. An overview of anomaly detection is provided in this work, with a particular emphasis on applications in banking operations. Banking operations involve a wide range of daily, weekly, and monthly tasks and exchanges carried out by or impacting several parties, including staff members, clients, debtors, and outside organizations. Events could develop gradually, and early identification greatly reduces the likelihood of negative consequences and, in certain situations, even completely prevents them. Finding people at unfavorable periods is accomplished using anomaly detection based on time series. This research offers a machine learning based anomaly detection technique to discriminate between normal and abnormal occurrences. A comparison is made between the biometric identity of the captured face and the biometric identities of known criminals. If a match is found, we can identify and capture the culprit right away.

Keywords: Video Surveillance System, Movement Detection, Face Detection, Face Recognition, Criminal Identification, Alert System

1. INTRODUCTION

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. Present-day digital video surveillance systems merely offer the technology needed to gather, store, and distribute footage; they do not need human operators at all to identify threats. 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. The analytics tool automatically identifies events and objects (people, cars, and equipment) of interest from video flow photos for security reasons. The process of monitoring or analyzing 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 moments of instability that are occurring globally. Therefore, 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. The affordability of video cameras has resulted in a rise in the utilization of video surveillance systems. There are several uses for video surveillance systems, such as traffic monitoring and the identification of human activity. 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.

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. The aim is to analyze the movement pattern and obtain a high-level understanding. If an incident seems improbable or unanticipated, it is deemed abnormal. In statistics, 'anomaly' is defined as an unusual behaviour in a distribution or an outlier data point in a data space. Naturally, in RVS, behavior understanding is a subdomain of anomalous detection. An outlier is interpreted as abnormal behavior by anomaly detection algorithms after they have been trained on a typical dataset. These systems usually cooperate to recognize and localize anomalous activity. The location of an anomalous event can be found by identifying the anomalous pixels in each frame. For AED to react swiftly and precisely, an effective data representation method is required. Data space in AED is spatiotemporal, encompassing both appearance and movement. A range of methods are used by researchers to identify anomalous areas within a video file. 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. When examining crowd scenes, a number of problems come up, such as occlusion, shadowing, and object overlap. Many algorithms have been proposed to solve this problem, but each has advantages and disadvantages of its own. The movements of the scene's elements might be interpreted in a variety of ways. These methods are able to replicate the velocity and direction of each individual object with precision. However, these methods usually require a lot of time. The perspective distortion of urban surveillance film, which produces different scale and movement patterns based on item placements and camera position, further ads 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. Deep neural networks are among the many machine learning models that require a large amount of labeled data. 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 approaches used for anomaly detection in RVS. These approaches have drawbacks, including an ineffective clustering center, imbalanced data (normal and abnormal occurrences), and a large number of objects in the traffic scene.

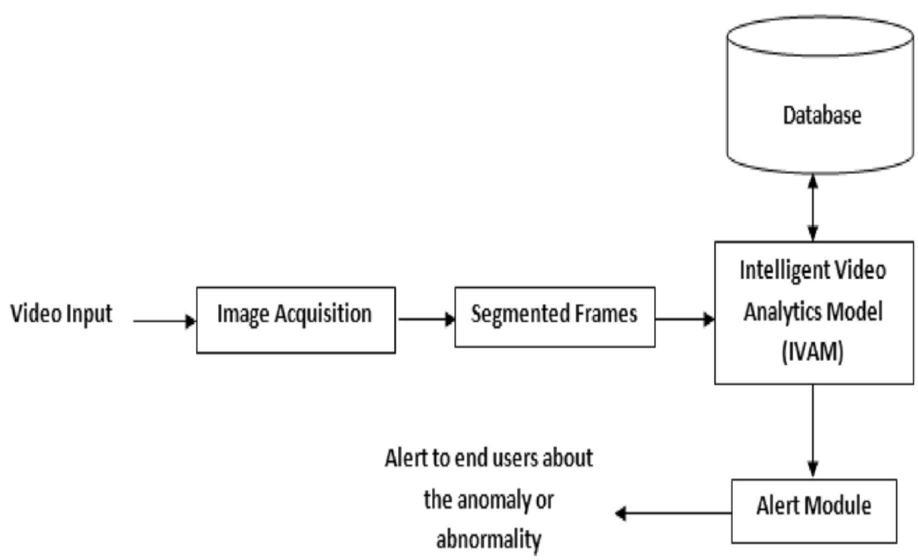

Fig 1. Video Anomaly Detection

2. LITERATURE SURVEY

2.1 A LARGE-SCALE BENCHMARK DATASET FOR ANOMALY DETECTION AND RARE EVENT CLASSIFICATION FOR AUDIO FORENSICS.

Ahmed abbasi, 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. In this work, many features are retrieved for feature engineering from the audio signal, and the fewest number of best-performing features is selected using the PCA feature selection technique for optimal performance. 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. Several challenges have emerged for data analysis in order to extract useful information from multimedia data. 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.

2.2 UNSUPERVISED ANOMALY VIDEO DETECTION VIA A DOUBLE-FLOW CONVLSTM VARIATIONAL AUTOENCODER

Lin Wang, et.al, [2] The ConvLSTM-VAE (Asymmetric) model, which combines ConvLSTM and VAE, is suggested to learn training data distribution for video anomaly detection. The DF-ConvLSTM-VAE model has been applied. 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.

2.3 RETHINKING VIDEO ANOMALY DETECTION-A CONTINUAL LEARNING APPROACH

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.

2.4 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

Mohana, et.al, [4] created a smart surveillance system for security applications that allows for real-time object tracking and detection. Most domains are now using a manually operated system that is too time-consuming to use efficiently 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 the OV7670 camera is used to record the footage. Five to six times more efficient than DSP boards, FPGA systems provide both temporal and spatial parallelism, allowing for flexible module design in accordance with system requirements. 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.

2.5 PERSPECTIVE OF ANOMALY DETECTION IN BIG DATA FOR DATA QUALITY IMPROVEMENT

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. Big data analysis is currently being supported by a number of financial industries, helping them to improve the services they offer to customers, both internal and external, and to build their passive and dynamic 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.

3. EXISTING SYSTEM

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 for anomalous events can be roughly divided into two phases: anomaly detection model and event encoding. These phases correlate to the most popular 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. After merging the output from the movement auto encoder and the spatially auto encoder, we combine the prediction with group distance at the pixel-level to evaluate the anomaly and produce the forecast of the last individual frame (LIF). 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. These 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.

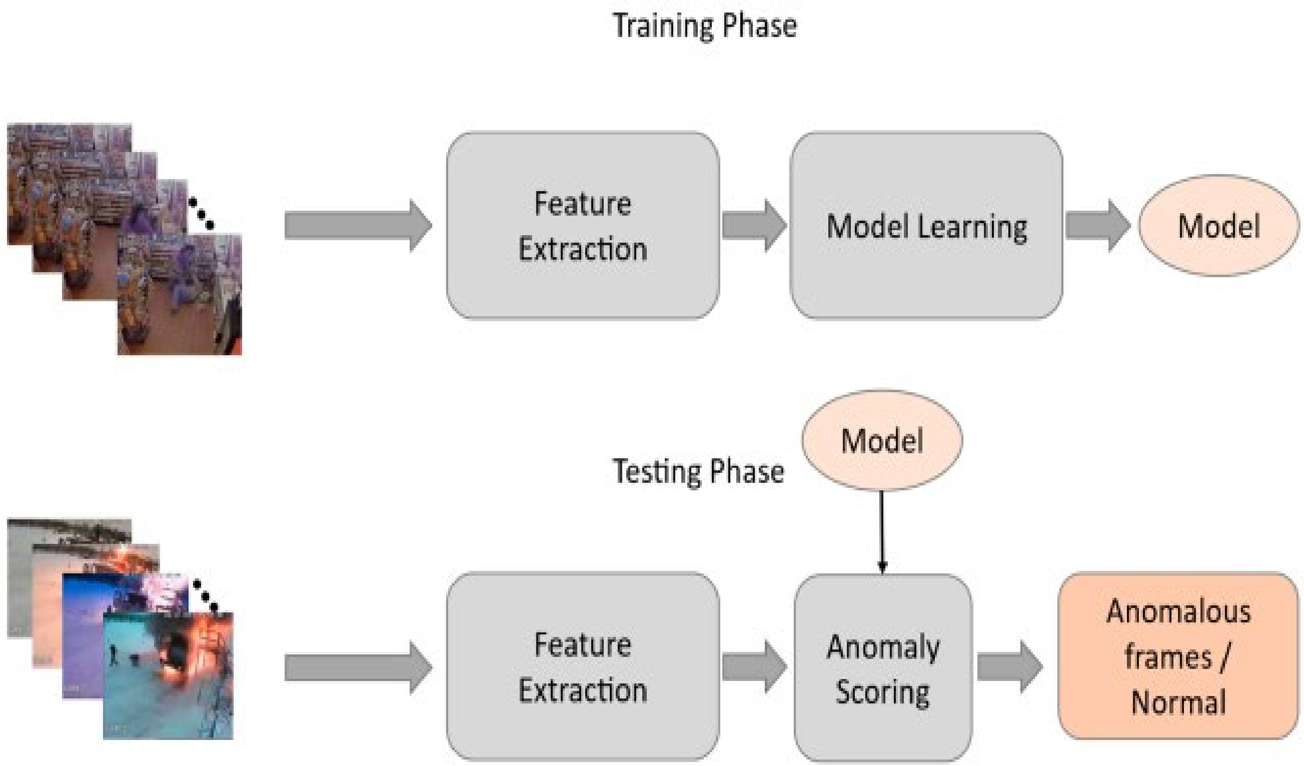

Fig 3.1: Existing System Model

4. PROPOSED SYSTEM

The installation of a smart camera, which monitors criminal behavior at banks and may spot any suspicious activity, is the primary element of the suggested system. The camera tracks thieves using motion detection and time-based facial recognition. The security department will receive an automatic notice from the smart camera if a suspicious face is spotted at an inconvenient time. In order for security to be ready, the message explains the type of warning that is sent out and provides a link to a website where the live image is saved when a face is spotted. The suggested research focuses primarily on developing the anomaly identification method rather than examining the frames; by learning from the output of previously trained models, this also reduces the complexity of the outlier detection model. The appearance, velocity, and position of unusual objects are just a few of the many possible causes of anomalous events. We use pre-trained models for backdrop segmentation, multi-object tracking, and classification methods to extract the associated properties. The three essential operations performed by the inversion layer are pooling layer, pooling, and activation. This is the basic formula. Filters are applied to the input matrix during the convolution process to form a feature matrix. Every feature matrix is put through an activation method in order to collect high-level properties throughout the activation operation. Combining an activation process's output can reduce its dimension. The next convolution layer is then fed the pooled output. In addition to producing more complex features like forms and specific objects, convolution also produces low-level features like lines, points, and edges. Convolution is the process of applying filters to incoming images in order to generate visual features that highlight and indicate the presence of distinct visual elements. Every filter point, which extends from top left to bottom right, generates a single weighted value, which is the unique result of multiplying the associated overlay image's pixel by the coefficients of each filter. 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.

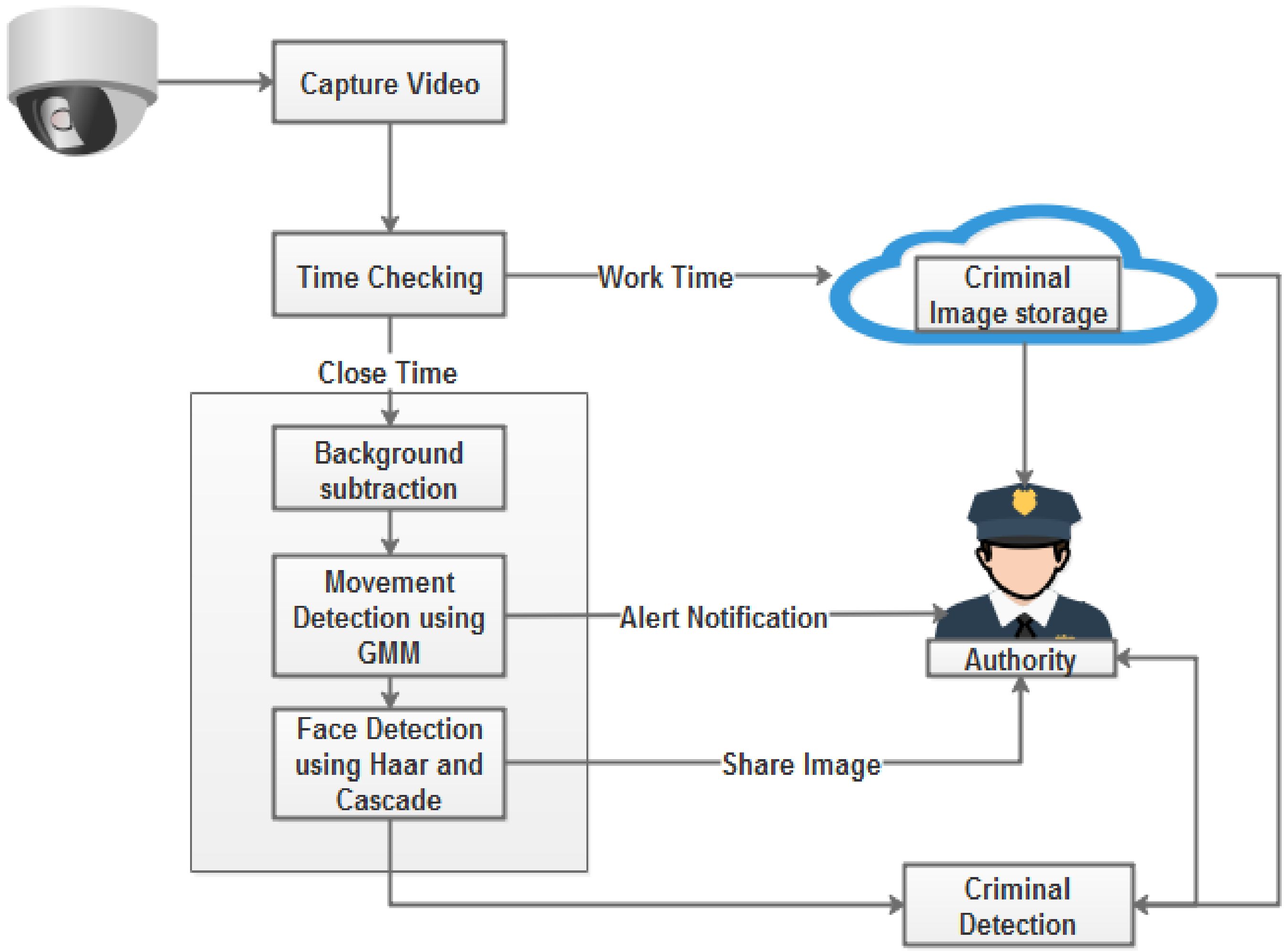

Fig 4.1: Architecture for Proposed Work

4.1 Gaussian Mixture Model

Let us denote with 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 (hypothesis) or a target pixel (hypothesis). Further, it is assumed that background consists of N different clusters (i = 1,2,..., N) corresponding to different ground cover types. The generic cluster is modeled as Gaussian distributed and its p.d.f. is: denotes the multivariate Gaussian p.d.f. with mean vector and covariance matrix

4.2 HAAR Cascade algorithm:

In this project propose a new method for super resolution by first learning the high-frequency components in the facial data that can be added to a low-resolution input image to create a super-resolved image. Our method is different from conventional methods as we estimate the high-frequency components that are not used in other methods, to reconstruct a higher-resolution image, rather than studying the direct relationship between the high- and low resolution images. The HR and LR images are denoted as X and Y, respectively. We use bold uppercase to denote the transformation matrices (or projection matrices). Specifically, we use Ψ to denote the dual-tree complex wavelet transform in a matrix form. Bold lowercase letters denote vectors. Plain uppercase letters denote regular matrices (i.e., L is used as a down sampling operation in matrix form). Plain lowercase letters are used as scalars. Superscript t is used to denote training. Integral images can be defined as two-dimensional lookup tables in the form of a matrix with the same size of the original image. Each element of the integral image contains the sum of all pixels located on the up-left region of the original image (in relation to the element's position). This allows to compute sum of rectangular areas in the image, at any position or scale, 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 low-resolution images are aliased and related to each other through sub-pixel shifts; essentially representing different snapshots of the same scene which carry complementary information. 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. D and Bk simulate the averaging process performed by the camera's CCD sensor while Mk can be modelled by anything from a simple parametric transformation to motion flow fields. Essentially, given multiple yk's, x can be recovered through an inversion process. The problem is usually ill-posed however, due to the large number of pixel values to be estimated from a small number of known pixels.

5. CONCLUSION

The proposed solution focuses on developing an anomaly detection system based on smart cameras that monitors activity in banks and can spot any unusual behaviour. 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. The message not only indicates what kind of warning is given, but it also includes the thief's face image, the moment the detection was made, and a web link to the live image so that the security may arrive prepared. By employing machine learning techniques, particularly in time series analysis, this research proposes a method for discriminating between normal and abnormal occurrences. The system utilizes movement detection, face detection, and face recognition technologies to identify potential threats or criminal activity. By comparing the biometric identity of captured faces with known criminal identities, the system can swiftly alert authorities to intervene and apprehend perpetrators. This approach not only enhances public safety but also contributes to the prevention and mitigation of negative consequences in banking operations and other public spaces. Early identification of anomalies can significantly reduce the likelihood of adverse events and facilitate timely intervention. Overall, the integration of advanced technologies in video surveillance systems holds immense potential for improving security measures and safeguarding communities.

CONFLICT OF INTERESTS

None

ACKNOWLEDGEMENTS

None

REFERENCES

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.

Wang, Lin, et al. "Unsupervised Anomaly Video Detection via a Double-Flow ConvLSTM Variational Autoencoder." IEEE Access 10 (2022): 44278-44289.

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.

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."

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.

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).

Sisodia, Shefali Shahane1 Shreya Dajgude2 Shreya, and Aditi Darade4 Mr Rahul Chakre. "Anomalous Motion Identification for Bank Surveillance." 2021

Prarthana Sanas, "Anomaly Detection at ATM Center using Machine Learning Algorithm" International Journal of Advanced Research in Science, Communication and Technology (IJARSCT), 2021

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.

Chang, Yunpeng, et al. "Video anomaly detection with spatio-temporal dissociation." Pattern Recognition 122 (2022): 108213.

Authors

Thilagavathi R
Department of Computer Science and Engineering, Mahendra Engineering College, Tamil Nadu, India
Thilagavathi R, a B.E. Student at Mahendra Engineering College. Her passion for machine learning drives her to delve into advanced disease detection which contributes to research in analysis of images.

Vairam Sountharya K
Department of Computer Science and Engineering, Mahendra Engineering College, Tamil Nadu, India
Vairam Sountharya K, a B.E. Student at Mahendra Engineering College, excels in transfer learning and image analysis, aiming to innovate in accuracy detections.

Varsha R
Department of Computer Science and Engineering, Mahendra Engineering College, Tamil Nadu, India
Varsha R, a B.E. Student at Mahendra Engineering College, specialized in machine learning. Her proficiency in application development underpins her project to create disease detection.