Efficient Detection of Human Misbehaviour in CCTV Real Time Footage

[1]MR. DR. S. MARUTHU PERMAL,
GUDIE, HOD
CSE
BIHER

TEAM MEMBERS
[2]M.KARTHIK REDDY, [3]N.HARINADH REDDY, [4]S. DINESH BABU, [5]SAYYAD SAMEER
CSE
BIHER

***Abstract*—In this paper, we propose real-time image-based recognition of human activities from series of images considering different human actions performed in an indoor environment. The proposed image-based human activity recognition (IHAR)system can be utilized for assisting the life of disabled persons, surveillance and human tracking, human computer interaction, and efficient resource utilization. The proposed IHAR system consists of closed-circuit television (CCTV) camera-based image acquisitioning, various filtering-based image enhancement, principle component analysis (PCA) based features extraction, and various machine learning algorithms for recognition accuracy performance comparison. We collected dataset of 10 different activities such as walking, sitting down and standing up consists of 35,530 images.**
The dataset is divided into (90%,10%), (80%,20%), and (70%,30%) training and testing respectively and evaluated three classifier K-nearest neighbors (KNN), Random Forest (RF), and Decision Tree (DT). The experimental results show the accuracy of 95%, 97%, and 90% by KNN, RF, and DT respectively.

***Index Terms*—Human Activity Recognition, CCTV, Principal Component Analysis, Random Forest, Decision Tree**

I. INTRODUCTION (*HEADING 1*)

The development in image processing technology makes it possible to intelligently monitor daily activity with cameras for a variety of applications, including emergency assistance for the elderly and disabled [1], safety and cognitive support [2, and driver driving status and drowsiness detection [3]. A series of observations on the activities of the individuals are made using an image, and the image-based human activity recognition (IHAR) method is utilized to identify distinct activities and gestures. The identification of human activities has been addressed in a number of ways. Several cameras and computer vision videos are used in some of the systems [2], while portable sensors are used in other systems [4], [5]. Users of wearable sensor-based solutions must remember to equip these sensors and other devices, such as smartphones. If the sensors or devices are removed or forgotten, identification is either canceled or has false recognition [6]. sensors

This research was made possible thanks to a National Research Foundation of Korea grant (NRF-2020R1A2B5B02002478) financed by the Korean government's ministry of science and ICT.

.

• In this research, we suggest an indoor human activity recognition system based on images. The key difficulties in activity recognition in a sequence of RGB photos include background clutter, poor illumination, image scaling, concurrent and similar activities, rotation, and resolution. Images of ten different actions, including walking, sitting, standing up, picking up, carrying, throwing, pushing, pulling, waving hands, and clapping hands were taken in an indoor setting with a variety of backdrop clutter, lighting conditions, and person image scale. We gathered a sizable amount of data that aids in solving the aforementioned issues. To combat this, a variety of filtering techniques are used to reduce noise and sharpen image edges. [7]–[9] In the suggested strategy, we contrast filtering methods like median filter, Soble filter, and Canny filter.

• The single-layered approach, also referred to as the sequential approach, which uses features taken from sequences to identify human activity. By extracting the features (such as, The degrees of the joint angles) per image frame, sequential methodologies turn a sequence of images into an arrangement of feature vectors. The approach compares the feature vector arrangements created for the person performing the activities after the feature vectors have been obtained. The system will determine that the activities have been carried out if there is a strong similarity between the arrangement and the activity class (i.e., a high posterior probability of the arrangement belonging to the action class). In this research, we offer a mechanism for feature extraction from a picture based on principal component analysis. Using the feature vectors, we subsequently train various machine learning algorithms. In the paper, we compared recognition accuracy performance. The summary of the literature on human activity recognition utilizing numerous devices, various features extraction techniques, and algorithms with the accuracy of various classifiers is presented in Table 1.

II. LITERATURE REVIEW

2.1 DETECTION OF SUSPICIOUS ACTIVITY AND ESTIMATE OF RISK FROM HUMAN BEHAVIOR SHOT BY SURVEILLANCE CAMERA MIWA TAKAI IEEE 2021.

Since surveillance camera systems may be used to monitor from distant locations using Web cameras connected to video monitors through a network, they are quickly becoming the security system of choice in modern society. Likewise, digital devices like hard drives and webcams are mass-produced and offered at inexpensive prices. Also, these digital devices' performance gains happen quickly. The current surveillance camera system displays dynamic photos from some surveillance regions captured simultaneously by several Web cameras. Due to the observer's ongoing exposure to a huge number of dynamic visuals, this system also exhausts the observer's mind and body. Also, this system has a severe issue with an observer that ignores a crime predictor. In order to determine a person's level of activity, this study extracts the motion region from a moving subject and measures the motion quantity. Additionally, this proposed method determines the suspicious activity's detection threshold and calculates its level of risk.

2.2 Automated Invigilation System for Detection of Suspicious Activities during Examination Md Adil; Rajbala Simon; Sunil Kumar Khatri IEEE 2021.

Around the world, offline examinations frequently entail cheating and deviant behaviors like whispering, imitation, or physical touch. This undermines the morality and integrity of a fair testing process. Systems for video surveillance incorporate image and video processing algorithms to track, examine, and process a variety of activities, including the detection of criminal activity and suspicious activity. The goal of this project is to develop a reliable model for a video surveillance system that can track and analyze behaviors during exams, enabling academic institutions to identify and deter unfair or suspicious behavior. This system architecture will use automated video feed to continuously monitor student activity throughout testing. We have talked about the various tools and methods used in surveillance models for video analytics, video processing, and image processing. Pre-processing, segmentation, classification, feature extraction, and their associated video processing algorithms are all discussed in detail throughout the study in a step-by-step fashion. The suggested model works well, uses little computing resources, and is efficient.

2.3 Self-powered event-triggered wireless sensor network for monitoring sabotage activities Chuan Dong; Suiqiong Li; Mengyang Li; Qisheng He; Dacheng Xu; Xinxin Li IEEE 2021.

For observing sabotage activities, a self-powered event-triggered wireless sensor network was created. The network is able to recognize suspicious activity and may instantly send out alarm signals. Self-powered wireless alerting sensor nodes (SWASNs), router nodes, receiver nodes, a GSM network, and a PC terminal make up the sensor network. An innovative vibration-threshold activated energy harvester powers the SWASN (VTT-EH), It only produces a considerable amount of electric energy when the input vibration amplitude exceeds a specific threshold to power the sensor node. The SWASN can transmit alerting signals to the router node using its event identification capability in the event of sabotage activities. The alarm signal is subsequently sent over the GSM network to the receiving node and the control center. The proposed network utilizes the GSM network to enable long distance transmission and the capacity to manage several sensor nodes across a wide area. The created sensor network underwent testing. The outcomes demonstrate that the intended function was accomplished. The suggested sensor network can be applied to numerous outdoor applications of activity monitoring.

2.4 Deep Learning Approach for Suspicious Activity Detection from Surveillance Video C.V Amrutha; C. Jyotsna; J. Amudha IEEE 2021.

Video surveillance is crucial in today's society. When machine learning, deep learning, and artificial intelligence were introduced to the system, the technologies had already evolved too far. Several methods are in place that help to distinguish distinct suspicious activities from the live tracking of footages using the combinations mentioned above. Human behavior is the most erratic, and it can be quite challenging to determine whether it is normal or suspicious. In an academic setting, a deep learning approach is utilized to identify suspect or regular behaviour. If suspicious activity is predicted, the approach alerts the appropriate authority. Consecutive frames taken from the video are frequently used for monitoring. There are two sections to the complete structure.In the first part, the features are computed from video frames and in second part, based on the obtained features classifier predict the class as suspicious or normal.

I. 2.5 WIRELESS REAL TIME SUSPICIOUS ACTIVITY DETECTION USING SMART GLASS SHEFALI SARANG; HARSHAL SHINDE; VAISHNAVI RAUT; SHUBHAM SONJE; GARGI PHADKE IEEE 2021.

The world we live in today is less secure. Despite this, there is always a threat, whether it be robbery or traffic accidents. To address these issues, numerous security measures have been installed. Instead, they take up memory and record video. It makes no inferences regarding the occurrence. A real-time suspicious activity detection system should be created to address these issues. This system will be superior to conventional systems since it will constantly watch the frame from a certain camera. Any field can use this and implement it with minimal hardware. Using image processing, the system we are building is utilized to keep an eye on what is happening in the camera's field of view.In this paper we are using a Raspberry Pi as our main processor to which camera will be interfaced.

III. EXISTING SYSTEM

3.1 PROBLEM DEFINITION

A method for improving an imaging system's resolution is called super-resolution imaging (SR). The technique of turning a low resolution (LR) image into a high resolution is known as single image super resolution (HR). This approach can be used in a variety of fields. It draws attention to or improves the image's quality. Here, a low resolution image will be utilized as input and converted to a high resolution image. The concept of super resolution is now widely discussed in image processing.

The following are the problems in the existing system:

1. Less accuracy rate in classifying the suspected persons in the input video.
2. More time consumption.
3. Cannot be implemented in all datasets.
4. More noise ratio.

3.2 Naïve Bayes (NB):

A family of straightforward "probabilistic classifiers" known as naive Bayes classifiers in statistics are based on the application of Bayes' theorem with strong (naive) independence assumptions between the features. These are some of the most straightforward Bayesian network models, but when used with kernel density estimation, they can produce results with higher levels of accuracy.

The number of parameters required for naive Bayes classifiers is linear in the number of variables (features/predictors) in a learning problem, making them extremely scalable. Instead than using an expensive iterative approximation, which is how many other types of classifiers are trained, maximum-likelihood training can be accomplished by evaluating a closed-form expression, which requires linear time.

Simple Bayes and Independent Bayes are two names for naïve Bayes models that can be found in statistics and computer science literature. All of these names refer to the classifier's decision rule using the Bayes theorem, however naive Bayes is not (by definition) a Bayesian approach.

Naive Bayes is a straightforward method for building classifiers. These models assign class labels to problem cases, which are represented as vectors of feature values, and the class labels are chosen from a finite set. For training such classifiers, there isn't just one technique, but rather a family of algorithms built on the premise that, given the class variable, the value of one feature is independent of the value of every other feature. For instance, if a fruit is red, round, and roughly 10 cm in diameter, it may be regarded as an apple. Regardless of any potential relationships between the variables of color, roundness, and diameter, a naive Bayes classifier assumes that each of these features contributes independently to the likelihood that this fruit is an apple.

Naive Bayes classifiers can be taught very effectively in a supervised learning environment for specific kinds of probability models. It is possible to work with the naive Bayes model without accepting Bayesian probability or applying any Bayesian techniques because parameter estimation for naive Bayes models frequently employs the maximum likelihood method.

Naive Bayes classifiers have performed admirably in a variety of challenging real-world circumstances despite their naïve design and ostensibly oversimplified assumptions. An examination of the Bayesian classification problem in 2004 revealed that the seemingly improbable efficiency of naive Bayes classifiers had solid theoretical justifications. However, a thorough evaluation of various classification algorithms in 2006 revealed that methods like boosted trees and random forests beat Bayes classification.

An advantage of naive Bayes is that it only requires a small number of training data to estimate the parameters necessary for classification.

3.2.1 Existing System Disadvantages:

1. Poor image resolution
2. Less Accuracy
3. Poor lighting and low contrast
4. Higher Computational Cost
5. Lack of standards
6. More time-consuming process

IV. PROPOSED SYSTEM

4.1 MODULES DESCRIPTION

There are 5 components in the system. They are

(i) Image Acquisition
(ii) Image Preprocessing
(iii) Image Segmentation
(iv) Feature Extraction
(v) Classification.

CCTV footage is gathered in real-time during image acquisition. The gathered pictures are reduced in size. Afterwards the image's internal noise is eliminated. Division is the third element. It entails employing median filtering to partition the transformed grayscale images. This aids in eliminating issues with backdrops, lighting, etc. By removing or displaying a section of the split images, feature extraction makes classification simpler. The classification in which Tensor Flow and DBN are utilized is included in the final module.

1. Image Acquisition Image

Image Acquisition is the process of collection of images. These images are downloaded from the online dataset provider called Kaggle.com.

2. Image Preprocessing

Converting color images to grayscale, standard images and resizing them are all included in image preparation. Black and white are combined in grayscale images. Images that are grayscale neutralize the background and help to reduce noise. Moreover, it aids in increasing the image's brightness. The ability to generate more data from limited data and the prevention of over fitting are two advantages of data augmentation, a method of generating new data.

3. Image Segmentation

By using image segmentation, the image is divided into useful parts. It separates digital images into many sections. Simplifying or transforming the representation into a more meaningful image is the aim. It distinguishes between the things we want to look at

more closely and the other things or their surroundings. It entails utilizing K means segmentation to divide up the transformed grayscale images.

4. Feature Extraction

The divided area of the image is extracted or displayed as a feature, making categorization simpler. To discriminate between the photos, features are retrieved. Nearly all machine vision techniques rely on features extraction. The conversion of segmented objects into representations that more accurately reflect their primary characteristics and qualities is the common objective of feature extraction and representation techniques.

5. Classification

4.2 Here, we employ the DBN categorization principle. The classification, for which Tensor Flow and machine learning algorithms will be utilized, is covered in the final module. A matlab-friendly open source toolkit for numerical computing called Tensor Flow accelerates and simplifies machine learning. Dataflow graphs—structures that depict how data flows through a graph or a collection of processing nodes—can be created by developers using Tensor Flow. A mathematical operation is represented by each node in the graph, and each edge between nodes is a multidimensional data array, or tensor.

4.3 ALGORITHM OF PROPOSED WORK

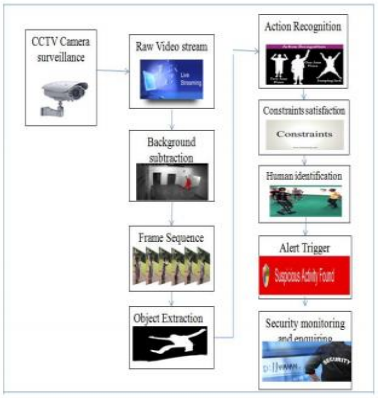
4.4 Deep Belief Network (DBN):

1. Read the source image into input.
2. For pre-processing step, the input image is converted to GREY format from RGB format.
3. GREY format dataset is analyzed into red, green and blue plane which helps analyzing each pixel individually.
4. Histogram is generated which helps in differentiating red, green and blue plane from which net deterministic value for each pixel differentiation.
5. Training parameters are obtained from histogram differentiation which is integrated with intelligence.
6. Various models for various planes are generated.

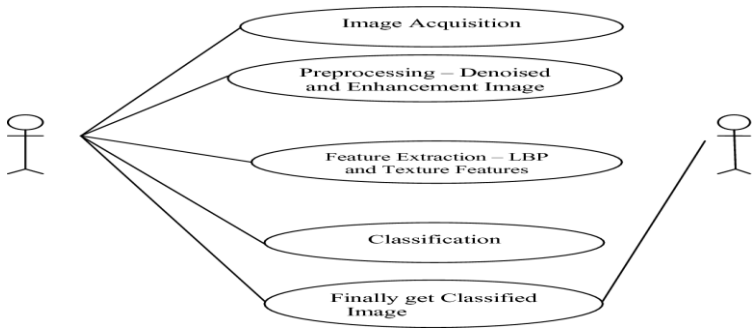
dimension and our color pattern on that particular image 7. A 3D matrix is obtained from results where each dimension refers to particular pixels.

8. Non pixel data generated from convolution is removed then it is integrated with DBN.
9. DBN decides the maxima and minima in convolution models supplied to it as input.
10. DBN technique integrates the similar color pattern on a particular pixel boundary of our convolution model. Similarly it does for other is generated which is our final output.

4.3 SYSTEM ARCHITECTURE



4.4 Use case Diagram



V. CONCLUSION AND FUTURE SCOPE

9.1 Conclusion

The system has demonstrated a revolutionary module that produced an accurate background without the formation of either artificial "ghost" trails or wasteful pixels. The AT module removed the needless investigation of the entire background region once a high-quality background model had been created, which decreased the computing complexity for the ensuing motion detection phase.

The suggested object extraction module used the moving object mask to identify moving object pixels within the triggered alert zone. Also, it starts the creation of a system for the observation and analysis of questionable human behavior. Lastly, this technique has minimal computational complexity and is effective for online (real-time) video processing.

9.2 Future Scope

The technology can be utilized in the future with a highly accessible storage service and can also be used to implement a high-tech method of video recording in monitoring regions. Our suggested method can be installed in roads, traffic light polls, and other locations to track offenders.

REFERENCES:

[1]. Xiaojing Chen, Zen Qin, Le An, BirBhanu"An Online learnedElementary Grouping model for multi-target tracking", University of California, Riverside, In CVPR 2019.
[2]. RohitAgarwal, Sudhir Kumar, Rajesh M. Hegde"Algorithms for Crowd Surveillance Using Passive Acoustic Sensors Over a Multimodal Sensor Network", IEEE Sensors Journal, Vol. 15, No. 3March 2019.
[3]. Md. ShakowatZamanSarker, Tan Wooi Haw and RajasvaranLogeswaran"Morphological based technique for image segmentation" International Journal of Information Technology, Vol. 14 No. 1.
[4]. SeemaKamath"Distributed algorithms for camera network localization and multiple targets tracking using mobile robots" Master of Science Thesis, RPI Troy, New York, Nov 2019.
[5]. Loris Bazzani"Beyond Multi-target Tracking Statistical Pattern Analysis of People and Groups", UniversitydegliStudi di Verona Department of Informatics." May3, 2019.
[6]. Numerod'ordre, Annee"Object Tracking and Re-identification in Multi-Camera Environments",Thesis University Lumiere Lyon, 17 June 2019.
[7].BrankoMarkoski,ZdravkoIvankovic,LadislavRatgeber,PredragPecev,DraganaGlusac"Application of AdaBoost Algorithm in Basketball Player Detection" University of Novi Sad,ActaPolytechnicaHungarica Vol. 12, No. 1, 2019.
[8] "Super Normal Vector for Human Activity Recognition with Depth Cameras", Xiaodong Yang, Member, IEEE, and YingLi Tian, Senior Member, IEEE
[9] C. Lu, J. Jia, and C. Tang, "Range-Sample Depth Feature for Action Recognition", CVPR, 2019.
[10] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R.Moore, A. Kipman, and A. Blake, "Real- Time Pose Recognition in Parts from Single Depth Images", CVPR, 2019.