ijRASET

International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

www.ijraset.com

Call: ⓒ08813907089 | E-mail ID: ijraset@gmail.com

Unusual Event Detection

Mr. Prashanth H S[1], Ashrit Madhav Vadiraj[2], Deepak S[3], Ajay Girish[4], Saksham Singh[5]

[1]*Assistant Professor, Dept of CSE KSIT College of Engineering, Bangalore, Affiliated to VTU*

[2, 3, 4, 5]*Students, Dept of CSE KSIT College of Engineering, Bangalore, Affiliated to VTU*

Abstract: The protection of privacy against covert video recordings presents a considerable societal challenge. Our objective is to develop a computer vision system, such as a robotic device, that can identify human activities and improve daily life, all while ensuring that it does not capture video content that may violate individuals' privacy. In this paper, we propose a fundamental approach to reconcile these seemingly conflicting goals: the recognition of human activities using highly anonymized video data with extremely low resolution (e.g., 16x12). We introduce a new concept called "inverse super resolution" (ISR), which entails learning the optimal set of image transformations to generate multiple low-resolution (LR) training videos from a single high-resolution source. Our ISR framework learns diverse sub-pixel transformations that are specifically optimized for activity classification, enabling the classifier to leverage high-resolution videos, such as those found on platforms like YouTube, by generating multiple LR training videos tailored to the specific activity recognition task. Through empirical experimentation, we demonstrate that the ISR paradigm significantly enhances activity recognition from extremely low-resolution video data.

I. INTRODUCTION

Numerous initiatives aimed at the intelligent recognition of human actions have emerged as a result of recent developments in the disciplines of computer vision and pattern recognition. A comprehensive understanding of human activity is crucial in a wide range of applications, including surveillance systems and human-computer interactions. Notably, a system for recognizing human activities has the potential to distinguish between unusual activities and regular activities performed by individuals in public spaces, such as airports and subway stations. The automated recognition of human activities is particularly valuable for real-time monitoring of elderly individuals, patients, or infants. Specifically, human action recognition aims to automatically identify and categorize the activities performed by a person, whether it involves walking, dancing, or engaging in other actions. This capability is essential for numerous applications, including surveillance, content-based image retrieval, and human-robot interaction. The complexity of this task arises from the variability in the appearance of individuals, the intricate articulation of poses, changing backgrounds, and the possibility of camera movements. In practical applications, tracking a target in low-resolution video poses a formidable challenge as a result of the loss of discernible visual features in objects in motion. Existing methodologies typically rely on the augmentation of low-resolution (LR) video through super-resolution techniques, albeit at a high computational cost. This cost escalates further when dealing with event detection tasks. This paper introduces an algorithm tailored for the detection of unusual events without resorting to such conversions, making it well-suited for enhancing College/University security, where cost-effective, low-resolution cameras are commonly utilized. The proposed algorithm employs morphological operations with a disk-like structuring element during preprocessing to address the limitations of low-resolution video. It separates dynamic backdrops from foreground elements in a scene using a rolling average background subtraction algorithm. The algorithm is noteworthy for its ability to recognize infrequent occurrences in low-resolution video, like fighting or crowding, based only on statistical analysis of moving objects' standard deviation. Its efficiency stems from its ability to process low-resolution frames swiftly, making it a valuable tool for enhancing College/University surveillance systems that continue to rely on conventional, low-resolution cameras. Importantly, the algorithm obviates the need for classifiers and initial system training.

Various techniques have been employed for scenario recognition, including Bayesian approaches and Hidden Markov Models (HMM) to detect both simple and complex events within scenes. However, this paper demonstrates that a simpler alternative, based on a control rule-based scheme, can be equally effective in recognizing activities in a given context. The action recognition system leverages silhouettes of objects extracted from video sequences for action classification, involving two primary phases: offline manual creation of silhouette and action templates, and real-time automatic action recognition. Actions performed by humans, such as walking, boxing, and kicking, are classified on the basis of temporal signatures of different actions represented by silhouette poses. This paper predominantly focuses on rule-based activity recognition, drawing inferences about action classes based on the recognition of individual atomic actions or sequences of recognized poses. The identified action classes are often considered as atomic or primitive actions, with more intricate activities being perceived as sequences of these fundamental actions.

II.   LITERATURE SURVEY

The impetus for this research stems from the complex security challenges encountered at College/University booths, as delineated in the COPS framework [1]. These intricacies have served as a prime catalyst for the development of an efficient security system. Within the realm of human action recognition (HAR), numerous approaches have been put forth by researchers. For example, temporal templates were first used to identify human actions by Davis and Bobick [3]. Reference [7] presents a thorough analysis of human motion and behavior using Motion History Images (MHI) and its variations. Furthermore, alternative approaches like Random Sample Consensus (RANSAC) and Optical Flow, as explained by [8], clarify methods for representing and identifying human actions. The use of Human Moments, which are popular shape descriptors favored for their ease of use and computational efficiency [7], for feature extraction from temporal templates was first described by Bobick and Davis [6]. Various other descriptors, including Fourier descriptors (FD) and Zernike moments, have also been proposed. However, Fourier descriptors exhibit limitations when dealing with variations in image size, resulting in computational challenges for real-time applications due to the changing number of data points. Zernike moments, while an advanced version of Human Moments with rotation-invariant magnitudes, are hindered by their extensive computational requirements, rendering them less appropriate for processing in real time. [2].

III.   PROPOSED SYSTEM

The system employs a web camera for the real-time video capture, from which individual image frames are extracted. These frames are subsequently subjected to face detection as the initial stage of the processing pipeline. A trained Haar classifier [13] is utilized to discern any anomalous activities as prescribed by a defined algorithm. In the event of detecting such unusual activity, an alert mechanism is promptly triggered.

In the absence of any unusual events, customers may proceed with standard College/University transactions. The proposed system seamlessly integrates a USB camera or the built- in webcam of a laptop or PC, alongside integrated speakers. These components are serially interfaced with an ARM7 LPC2148 processor via UART communication. The live video stream captured by the web camera is processed on the connected PC using custom-developed code to detect any irregular occurrences or events within the College/University environment.

The security system of the college or university is designed to automatically lock its doors and sound an alarm when it detects unusual activity. These notifications are sent via email to security personnel stationed in a central observation room. Furthermore, a panic switch is made available for immediate alerting of local individuals, while a PIR sensor is deployed to automatically activate internal lighting when human motion is detected. The use of GSM modules facilitates message triggering, and the system's internet connectivity allows for automated email notifications to be dispatched to a designated security agency control room.

The operational stages of the system can be summarized as follows:

1) Video Acquisition: This initial stage involves capturing video content through a variety of devices, including Handicaps, mobile cameras, USB cameras, and CCTV cameras.

2) Frame Conversion: Following video capture, the content is converted into individual frames, rendering them suitable for subsequent processing.

3) Preprocessing: The frames undergo preprocessing to minimize noise through techniques such as smoothing, dilation, erosion, median filtering, and opening/closing operations.

4) Background Modeling: Subsequent to preprocessing, background modeling is employed to create an ideal background, either static or dynamic, adjusting to environmental changes. This step may incorporate image subtraction operations and falls into the categories of recursive or non-recursive techniques.

5) Background Subtraction: This pivotal stage identifies significant changes within the image regions concerning the background model. These modified regions' corresponding pixels are marked for additional processing, usually using connected component labeling algorithms.

6) Post-processing: Following background modeling and subtraction, additional techniques are used to improve the results. This is the final refinement. Refining the foreground mask is the goal of several post-processing techniques.

7) Foreground Extraction: In the final stage, the system determines the overall effectiveness of the background subtraction system by extracting moving objects within the frame.

IV. METHODOLOGY

Using low-resolution cameras to detect unusual events intelligently and improve security is a methodical process that involves multiple critical steps. It commences with data acquisition, where low-resolution cameras like USB cameras or built-in webcams capture live video in areas necessitating heightened security, including College/University and public spaces. Video preprocessing follows, where live video is transformed into individual image frames, with noise reduction techniques applied to improve frame quality Background modeling and subtraction are essential because they generate an optimal background model that can be used to distinguish between background and foreground elements over time. Object detection and tracking are used to identify and track objects or people, and they use trained Haar classifiers to detect unusual activity according to preset algorithms. Event recognition examines objects that have been detected and their behavior to differentiate between normal and abnormal events based on predetermined criteria. An alert mechanism responds to unusual events, potentially triggering actions like automatic door locking and alert notifications. Additional sensors, such as PIR motion sensors, enhance human motion detection. Communication modules, including GSM, facilitate alerts, and automated email notifications ensure remote monitoring. User authentication via PIN is provided for normal transactions. Regular testing, optimization, and maintenance ensure system effectiveness, while real-time monitoring in a central observation room enables swift responses to potential security threats.
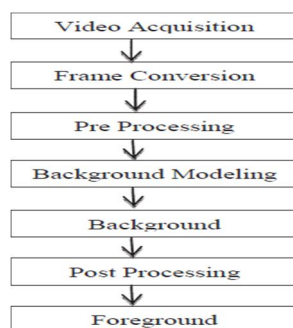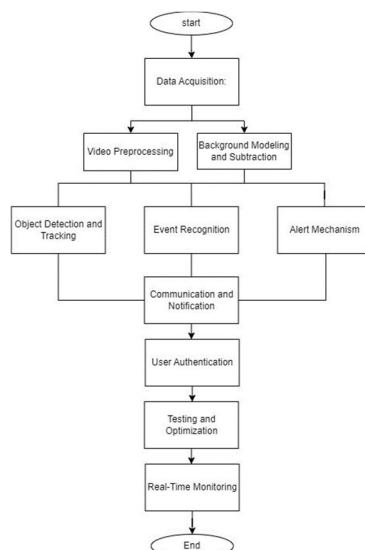
Fig.1. Image processing flowchart

Fig.2 Methodology flowchart

V. CNN ALGORITHM

When it comes to image processing and object detection tasks, Convolutional Neural Networks (CNNs) are crucial. These deep learning algorithms have been intricately crafted to analyze visual data, specifically images and videos. CNNs have a broad range of applications. They stand out in image classification by effectively categorizing objects in images, learning unique patterns and features from labeled datasets. Additionally, CNNs play a vital role in object detection, identifying and locating objects by breaking down the image into regions and determining object presence within each region, employing techniques like R-CNN, Fast R-CNN, and Faster R-CNN. CNNs are also applied in semantic segmentation, which involves pixel-level object classification within images.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XI Nov 2023- Available at www.ijraset.com

Theyexhibit proficiency in transfer learning, making use of pre-trained models to adapt to specific datasetsand tasks. Specialized CNN architectures, such as SSD and YOLO, are tailored for real-time object detection. Lastly, CNNs serve as valuable feature extractors, with their learned features being appliedin various image processing tasks. CNNs have undoubtedly transformed the field of computer vision,boosting accuracy and efficiency in applications spanning autonomous vehicles, medical image analysis, facial recognition, and more.

VI.   HAAR CASCADE ALGORITHM

Object detection is the process of identifying custom objects within an image, and one of the techniques used for this task is the Haar cascade. This algorithm offers a straightforward approach todetect objects in images, regardless of their size or position. Haar cascade detection is not overly complex and can operate in real-time. It can be trained to recognize various objects such as cars, bikes, buildings, fruits, weapons and more. The Haar cascade method employs a cascading window to analyze features within each window and determine whether it corresponds to an object.
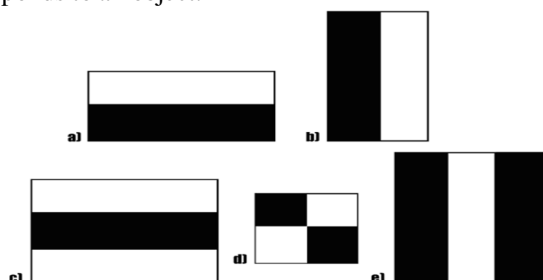

Fig 3. Types of Haar features.
a) Edge Features
b) Line Features
c) Four-Rectangle Features

Haar features are applied in a window-sized fashion across the image to compute and compare features. The Haar cascade acts as a classifier, distinguishing between positive data points, which represent parts ofthe detected object, and negative data points, which do not. While Haar cascades offer real-time performance and efficiency, they are less precise compared to modern object detection methods. They tend to produce a notable number of false positive predictions. One of the advantages of Haar cascade is its simplicity and lower computational demands.

VII.   CONCLUSION

We're working on a College/University model that will have more security features. It will mostly use facial recognition software. Our main goal is to reduce the verification time as much as possible while maintaining the College/University system's high level of efficiency. In order to address concernsregarding unauthorized actions, biometrics offer a promising method for identifying and authenticating persons within the setting of colleges and universities. By utilizing biometric verification, which necessitates the account holder's physical presence, this initiative seeks to address the urgent problem of fraudulent transactions in college and university settings. This will stop illicit transactions occurring without the account owner's knowledge. It is a strong strategy to use biometric traits for identification, andits power can be increased by adding other authentication techniques.

REFERENCES

[1]   U. Mahbub, H. Imtiaz, and M. A. R. Ahad, "Action recognition based on statistical analysis from clustered flow vectors," Signal, Image and Video Processing, vol. 8, no. 2, pp. 243–253,2014

[2]   J. Arunnehru and M. K. Geetha, "Motion intensity code for action recognition in video using PCA and SVM," in Mining Intelligence and Knowledge Exploration, vol. 8284 of Lecture Notes in Computer Science, pp. 70–81, Springer International Publishing, Cham, Switzerland,2013

[3]   Tian Wang, Snoussi, H., "Histograms of Optical Flow Orientation for Visual Abnormal Events Detection", in IEEE Ninth International Conference on Advanced Video and Signal- Based Surveillance (AVSS), 2012, pp. 13-18.

[4]   Lili Cui, Kehuang Li, Jiapin Chen, Zhenbo Li, "Abnormal event detection in traffic video surveillance based on local features", in Image and Signal Processing (CISP), 2011, pp. 362-366.

[5]   Sugandi, B., Hyoungseop Kim, JooKooi Tan, Ishikawa, "Tracking low resolution objects bymetric preservation", in Computer Vision and Pattern Recognition (CVPR), 2011, pp. 1329- 1336.

[6]   Y. Chen, Y. Rui, and T. Huang. "Multicue hmm-ukf for real time contour tracking", IEEETransactions on Pattern Analysis and Machine Intelligence.

INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 ⊙ (24*7 Support on Whatsapp)