VIOLENCE DETECTION AND ALERT SYSTEM TO THE POLICE USING IMAGE PROCESSING TECHNIQUES

A.Priyadharshini[1], R.Priyadharshini[2], Pattuna rajam[3]
final year students, Associate Professor
Department of Electronics & Communication Engineering,
Prince Shri Venkateshwara Padmavathy Engineering College,Chennai 127

Abstract

Violence has become a major threat to the society. It is important to detect the violence and save the victims through an efficient automated alert system to the concern officials. For this, surveillance cameras deployed in sensitive areas like ATMs, Hospitals, Government offices etc can be very much helpful. However, existing systems face the problems of low accuracy, high false alerts and high computational cost in monitoring and analyzing the data set images from surveillance cameras and making the decision about violence in real-time. In this paper an efficient framework is proposed which uses image processing technique to collect the images of human activities which generates the violence related features by KNN classification algorithm. It has some predefined dataset images. The obtained input image will be compared with pre-trained images once the violence image is executed, then the alert message is sent to the police authorities about the violence with the location using GPS and GSM. Arduino acts as brain of the system, because which stores the instructions.All operations displayed on LCD and Buzzer alarm is given to alert the people nearby the violence. In addition, various techniques is tested for improving the accuracy and reducing the error rate.
Keywords: Image processing, violence detection system, automated alert system, GPS, GSM.

1. Introduction

In recent years, violence has become a major issue all around the world. Recent studies from various papers reveal that street harassment is a prevalent problem and observed that 79% Indian women are affected. Violence against women has become a major issue in all countries. In-order to overcome domestic violence, street violence and crimesmany surveillance cameras are installed at almost all the sensitive areas like ATMs, banks, government offices, hospitals, prisons etc. However, these cameras are being monitored manuallywhich require lot of man power and also takes more time to alert the officials about any incident of violence. Thus, there is a need of an automated system which can give automated and accurate alerts of the violence activity without any human intervention, by monitoring and analyzing the video streams collected from the surveillance cameras.

To address such problems, Image processing fields has grown with various techniques for recognizing normal and abnormal activities like fighting, killing, harassment etc. In literature, researchers have proposed violence detection systems based on audio and video. In Audio based approaches, audio information like screams of person, gunshots, energy entropy etc parameters are used to detect violence. In visual based approaches, features shape, color, position of persons were detected. The video streams are sliced into various frames the images with slight change in behavior is obtained as an input image which is further resized and important features of the images are extracted. Using KNN algorithm the images are compared with the pre- trained images and classified.

2. Related works

In olden days, violence control are monitored manually. If violence occur, alert message will not be sent. Now a days, automatic human activity recognition has drawn much attention in the field of video analysis technology due to the growing demands from many applications, such as surveillance environments, entertainment environments and healthcare systems.The Existing approaches can be classified into audio based and video based violence detection.

A. Audio based Violence Detection

In earlier attempts, violence detection have considered only the audio information like gunshots, explosions etc. Cheng et al. [11] had proposed an hierarchical approach based on Gaussian Mixture Models and Hidden Markov Models, in which audio features like gunshots, explosions and car-braking are recognized as violence. Nam et al.

[12] have considered the combination of multiple audio-visual features in videos like flame detection, blood detection and degree of motion to detect violence. A mid-level violence clustering based system [13] has used the combination of audio and visual features with machine learning at segment level. Authors [14] have considered audio features based on time and frequency domain, like short time energy, zero crossing rate, energy entropy, spectral flux are extracted for each segment. These features are given as input to classifier to detect violence. The main problem with audio based systems is that many of the surveillance systems do not provide audio features. Moreover, any silent violence activity cannot be recognized. These problems lead to many research towards consideration of visual features for violence detection.

B. Video based Violence Detection

Violence detection based on video streams perform action detection like punching, kicking etc or use descriptor methods which can detect violence more pragmatically. Many existing methods have considered acceleration motion vector (AMV) and jerk to detect violence actions like kicking and fist fights .And have used adaptive background subtraction method to obtain silhouettes of moving human objects, from which motion trajectory information and orientation information of person's limbs are extracted to calculate AMV and Jerk. However, it may be ineffective in detecting group violence. In addition, it cannot detect violence, if the human object falls down during the fight. To address such challenges, different descriptor methods have been proposed.

Kuldeep Singh [17] had proposed an efficient framework in which they had used Slicing the frames according to motion tracker, Optical flow calculation, Mean calculation of magnitude change vector, Histogram formation, ViF descriptor calculation and Final decision on violence.

As per our analysis, the existing approaches suffer from low accuracy. And it requires man power to monitor

3. Proposedmethod

In this proposed system, Image processing technique is used to detect the violence. In this system, ARDUINO MEGA acts as brain of the system, because the entire system program instruction stored in it. Using image processing, the predefined dataset images are trained in the MATLAB. If the violence image is executed, then the alert message is sent to the police authorities about the violence with the location using GPS and GSM. All the operations are displayed on LCD and Buzzer alarm is given to alert the people nearby the violence.

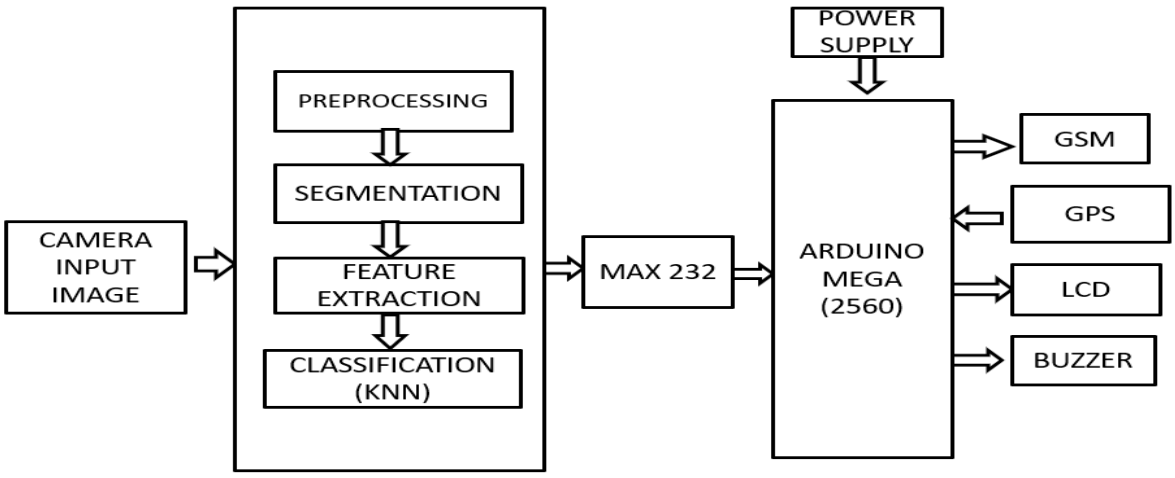


Figure 3.1 Modules of proposed framework

A.CAMERA INPUT IMAGE

Input the images are read from the dataset. The dataset class makes it easier and more efficient to perform tasks with all the image files. After creating a dataset with the image files, Files are decoded into usable pixel data and pre-processed for further enhancements.

B.PRE-PROCESSING:

The aim of pre-processing is an improvement of the image data that suppresses unwanted distortions. Enhances some image features important for further processing. Gaussian filter is used to blur the image and to reduce noise. In image processing, a Gaussian blur (also known as Gaussian smoothing) is the result of blurring an image by a Gaussian function. It is a widely used effect in graphics software, typically to reduce image noise and reduce detail. The visual effect of this blurring technique is a smooth blur resembling that of viewing the image through a translucent screen, distinctly different from the bokeh effect produced by an out-of-focus lens or the shadow of an object under usual illumination.

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

(1) represents 2-D Gaussian distribution.

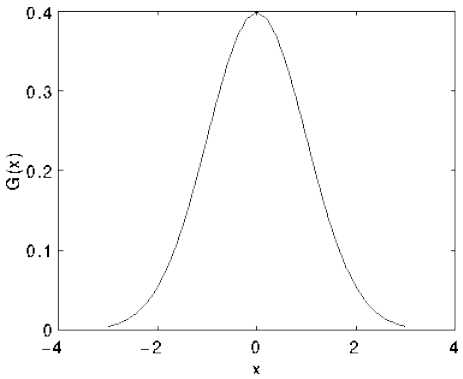


Figure 3.2 1-D Gaussian distribution with mean 0 and σ=1

C.SEGMENTATION:

Image segmentation is a commonly used technique in digital image processing and analysis to partition an image into multiple parts or regions. Region based segmentation separates the objects into different regions based on some threshold value. The edge detection segmentation makes use of discontinuous local features of an image to detect edges and define a boundary of the object. Edge-based segmentation algorithm works to detect edges in an image, based on various discontinuities in grey level, color, texture, brightness, saturation, contrast etc. To further enhancement concatenate all the edges into edge chains that correspond better with borders in the image

Figure 3.3 edge based detection

D.FEATURE EXTRACTION:

KAZE (FEATURE EXTRACTION ALGORITHM)

KAZE Features is a novel 2D feature detection and description methodthat operates completely in a nonlinear scale space. Feature is a piece of information about the content of the image. Feature may be specific structures in the image such as points, edges or objects. Feature extraction is related to dimensionality reduction.

When the input data to an algorithm is too large to be processed and it is suspected to be redundant (e.g. the same measurement in both feet and meters, or the repetitiveness of images presented as pixels), then it can be transformed

into a reduced set of features. Determining a subset of the initial features is called feature selection.The desired task can be performed by using this reduced representation instead of the complete initial data.

E.CLASSIFICATION:

The KNN binary (as two class) is given more accurate data classification which beneficial to select k as an odd number which avoids the irregular data. The KNN procedure is the technique in ML procedures: It is an object which classified through a mainstream selection of its neighbors, with the determination assigned occurrence for most mutual class amongst its k nearest neighbors (k is a positive integer, classically small). Classically Euclidean distance is used as the distance metric; however, this is only suitable for endless variables. In such situation as the classification of text, alternative metric, intersection metric or Hamming distance can be used. KNN is a new process that deliveries all available cases and categorizes novel cases built on an evaluation quantity(e.g., distance functions). KNN procedure is identical simple. It works built on a minimum distance from the interrogation instance to the training samples to regulate the K-nearest neighbors. The information for KNN procedure contains numerous attribute which will be used to categorize. The information of KNN can be any dimension scale from insignificant, to measurable scale.In this method the image is classified whether violence is detected or not.

When KNN is used for classification, the output can be calculated as the class with the highest frequency from the K-most similar instances. Each instance votes for their class and the class with the most votes is taken as the prediction. Class probabilities can be calculated as the normalized frequency of samples that belong to each class in the set of K most similar instances for a new data instance. For example, in a binary classification problem (class is 0 or 1)

p(class=0) = count(class=0) / (count(class=0)+count(class=1)) (2)

it is a good idea to choose a K value with an odd number to avoid a tie. And the inverse, use an even number for K when you have an odd number of classes.Ties can be broken consistently by expanding K by 1 and looking at the class of the next most similar instance in the training dataset.

Figure 3.4 sample violence images

F.MAX232

MAX 232 is connected to the PC is an integrated circuit that acts as a dual transmitter and receiver that typically is used to convert RX, TX, CTS, RTS signals. It acts as a voltage convertor. This IC is used as a hardware layer to converter like to communicate two systems simultaneously. Max 232 is integrated with PC.

4. Results and discussion

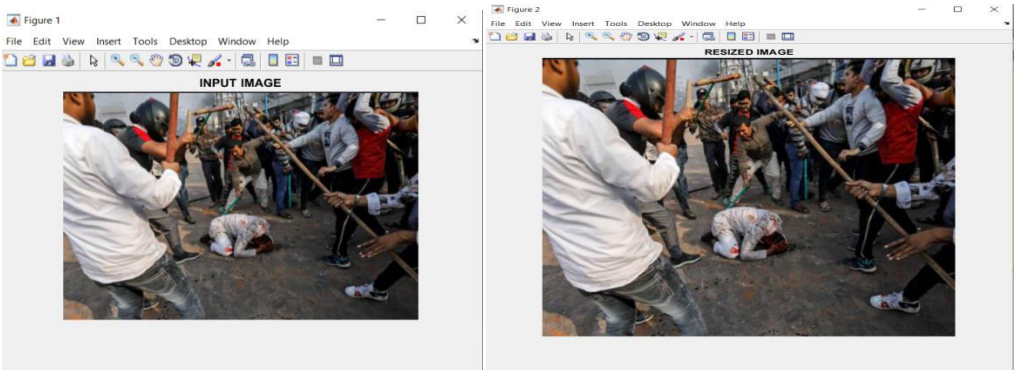
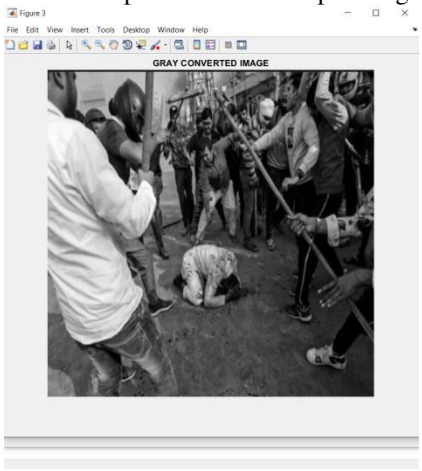


Figure 4.1 Input image **Figure 4.2 Resized image**

The input image is obtained and preprocessed. The image is resized and all the important feature of the image is read and enhanced further. All the input images are resized into same dimensions. If the specified size does not produce the same aspect ratio as the input image, the output image will be distorted.



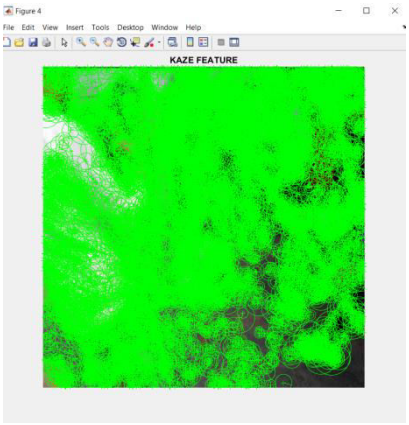


Figure 4.3 grayscale image **Figure 4.4 Feature extraction**

For many applications of image processing, color information doesn't help us. To store a single colour pixel of an RGB colour image we will need 8*3 = 24 bits (8 bit for each colour component), but when we convert an RGB image to grayscale image, only 8bit is required to store a single pixel of the image.
Important features are extracted in fig4.4. Whenever the violence is detected using the GSM module the message will be sent to the nearby higher official location and with the help of GPS the exact location along with longitude and latitude information will be sent to the higher officials.

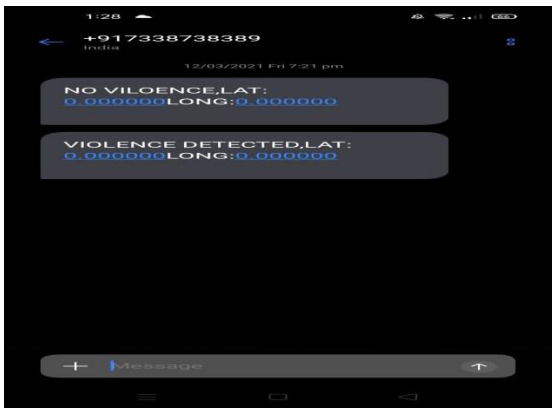


Figure 4.5 output

Thus the message will be sent to the official authorities with the exact location of the violence.
For the feasibility analysis, There are almost various machine learning techniques that are been tested out of that KNN algorithm gives the difference between the Target value and the Prediction value also shows that the value of error is small and the prediction results using KNN Algorithm have an accuracy of 97.27%. According to the results, the KNN Algorithm is effective in predicting the accuracy of image processing This paper also shows that the KNN Algorithm is stable with a small error ratio.

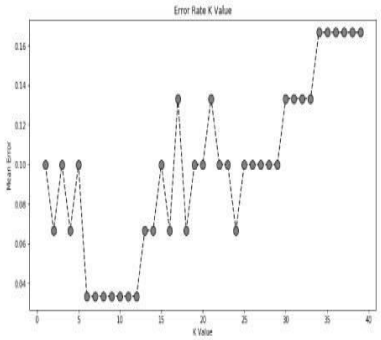


Figure 4.6 error rate of knn

5. Conclusion

This paperfocuses on improved accuracy rate on detection of violence using an effective algorithm. As the human security is very much important it is necessary to save the victims who are affected by any sort of violence. The existing systems faced issues like high false alerts, low accuracy and loss of life. This paper proposes an effective automated alert system for violence in sensitive areas like ATMs, Government officials, Hospitals etc. this alert system gives 97.27% of accuracy rate. Compared to previous existing system which had manual monitoring system The proposed system reduces loss of life. The algorithm used for classification faces some drawback where k should be wisely selected. And large computation cost during runtime if sample size is large.A general difference between KNN and other models is the large real time computation needed by KNN compared to others.In future we can enhance the accuracy rate of the detection system by using various new algorithms.

References

[1] S. Wilkinson. (2016) Meet the heroic campaigners making cities safer for women. [Online]. Available:
[2] T. Hassner, Y. Itcher, and O. Kliper-Gross, " Violent flows: Realtime detection of violent crowd behavior," in Computer Vision and Pattern Recognition Workshops (CVPRW), 2012, pp. 1-6.
[3] P. Bialon, " A linear support vector machine solver for a large number of training examples," Control and Cybernetics, vol. 38, no. 1, pp. 281-300, 2009.
[4] I. Dagher, " Quadratic kernel-free non-linear support vector machine," Journal of Global Optimization, vol. 41, no. 1, pp. 15-30, 2008.
[5] E. M. dos Santos and H. M. Gomes, " A comparative study of polynomial kernel svm applied to appearance-based object recognition," in Pattern Recognition w ith Support Vector Machines, 2002, pp. 408-418. 78
[6] M. Collins, R. E. Schapire, and Y. Singer, " Logistic regression,
[6] M. Collins, R. E. Schapire, and Y. Singer, " Logistic regression, adaboost and bregman distances," Mach. Learn., vol. 48, no. 1-3, pp. 253-285, 2002.
[7] L. Breiman, " Random forests," Machine Learning, vol. 45, no. 1, pp. 5-32, 2001.
[8] C. D. Sutton, " 11 - classification and regression trees, bagging, and boosting," vol. 24, pp. 303 — 329, 2005.
[9] E. B. Nievas, O. D. Suarez, G. B. GarGa, and R. Sukthankar, "Violence detection in video using computer vision techniques," in International conference on Computer analysis of images and patterns, 2011.
[10] M. S. Ryoo and J. K. Aggarwal, " Spatio-temporal relationship match: Video structure comparison for recognition of complex human activities," in IEEE International Conference on Computer Vision (ICCV), 2009, pp. 1593—1600.
[11] W.-H. Cheng, W .-T. Chu, and J.-L. Wu, " Semantic context detection based on hierarchical audio models," in Proceedings of the 5th ACM SIGMM international workshop on Multimedia information retrieval, 2003, pp. 109—115.
[12] J. Nam, M. Alghoniemy, and A. H. Tewfik, " Audio-visual content-based violent scene characterization," in International Conference on Image Processing(ICIP), 1998, pp. 353—357.
[13] S. Goto and T. Aoki, "Violent scenes detection using mid-level violence clustering," in Computer Science. CSCP, 2014, pp. 283— 296.
[14] T. Giannakopoulos, D. Kosmopoulos, A. Aristidou, and S. Theodoridis, " Violence content classification using audio features," in Hellenic Conference on A rtificial Intelligence, 2006, pp. 502—507. [15] A. Datta, M. Shah, and N. D. V. Lobo, " Person-on-person violence detection in video data," in International Conference on Pattern Recognition (ICPR), 2002, pp. 433—438.

[16] C. Stauffer and W. E. L. Grimson, " Learning patterns of activity using real-time tracking," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 8, pp. 747—757, 2000.
[17] Kuldeep Singh, K Yamini Preethi and Chirag N. Modi "Violence Detection in Sensitive Areas using Computer Vision and Machine Learning Techniques", IEEE 2018