IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 3, April 2022
Impact Factor: 6.252

Automatic Video Surveillance System for Pedestrian Crossing Using Machine Learning

A. Mohan Krishna[1], CH. Vinay[1], C. Nagi Reddy[1], P. Mohan Sai[1], Mrs. Benazir Begum[2]
Students, Department of Computer Science and Engineering[1]
Assistant Professor, Department of Computer Science and Engineering[2]
Hindusthan Institute of Technology, Coimbatore, Tamil Nadu, India

Abstract: *Nowadays accidents became very common in this world and the majority of accidents are from the crosswalks. The accidents are caused due to the absence of traffic lights to lead the traffic structure and people. Even though the traffic flow is huge by the traditional traffic accessories has a very particular warning capabilities which is incapable of to satisfy the safety requirements. Like a driver is unable to divine there will be pedestrian crossing the pathway forward, and pedestrians lack of traffic security alertness, which may easily cause of accidents. So, to overcome this issue the video surveillance helps the crippled persons to cross the pedestrians in a riskless way and it helps for independent vehicles. The work recommended an emerging technology to accommodate the crippled persons to cross the pathway just about the traffic signal and it also supports to someone who walks at a leisurely pace with the help of video surveillance. Here the image processing analysis algorithm takes a very crucial role to trace the motion of the object in the certain pedestrian crossing which can accommodate or serve the crippled persons or any senior citizens. They are numerous morphological filtering operations will improve the distinction of the motion of the person in the video. The proposed work improves more accuracy than the previous accuracy methods and also in addition of this we are implementing the crowd detection feature on the pedestrian, whenever the crowd was more in pedestrians.*

Keywords: Image Processing, Video Surveillance, Morphological Operations, Pedestrian Crossing

I. INTRODUCTION

A structure that intelligently dig up a person from an image or a video may be a tough work of this modern era. Over the past few years because of globalization a main change has been arises in various fields such as business, safety, health etc., This has been a very quick growth within the operation of video surveillance systems which can authorize by technological and fabricate the advances that have resulted an ever-widening fulfilment at a low cost. The number of pedestrian deaths and wounds can be decrease by hiring smart frameworks for finding out the persons on road. When we combined within the intelligent spaciousness, the video sensor networks communicate to the smart systems which has the capacity to observe the surroundings, and when it integrates with their smartness, to dig up and recognize the numerous unusual situations which can avoid the accidents. Thus, for the sake of determined by the safety, the video surveillance system was recommended.

The video surveillance system is used to study of the behavior, activity or other contents which is related to the people in a particular place like in shopping malls, public places, industries, banks, hospitals, traffic signals, etc. [1], Video surveillance is a prosperous sector for object tracking. The object tracking is a step-by-step process that trace the movement of one of the diverse of specific objects using cameras and capture the scenario [11-16]. Video surveillance with human being is a time exhausting process. Hence machine has to segment the video and extract the required details for more applications [8-9]. The importance of this arrangement assists the crippled persons and senior citizens to cross the traffic signals very safety manner. Video surveillance significance is appropriate in various sectors like missile tracking, safety and security determinations [2,3].

The figure shows the video surveillance system which is classified based on the application and arrangement of the camera system. The camera system which includes the single camera like camera1, camera2, and multiple cameras like camera N systems. The implementation support system which consists object tracking and remembrance. The application support system which uses in object tracking, identification, motion analysis, personal authentication purpose and behavioral

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 3, April 2022
Impact Factor: 6.252

analysis purpose. The arrangement-based video surveillance structure includes autonomous systems which allows to use programs and data that are stored and managed on the internet on personal computer and appropriate systems [4]. A general frame work of a video surveillance system shown in figure [5-7]. Generally, video surveillance structure is based on numerous cameras, the video from the different cameras is take hold through the network and keep in data sequence.

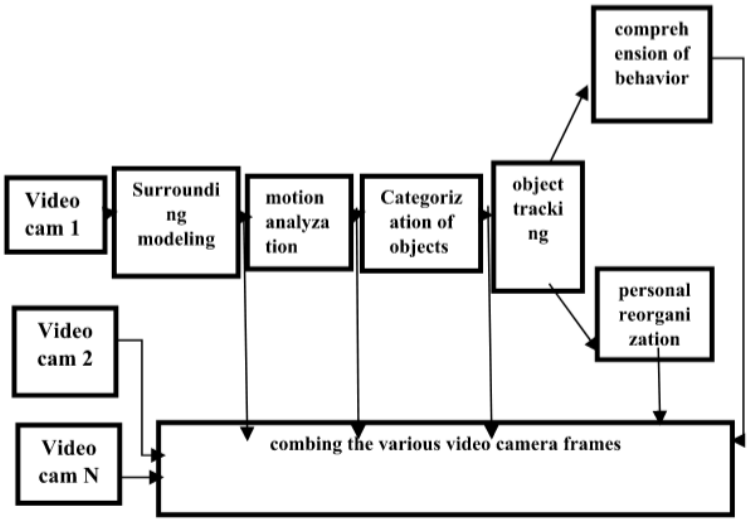


Figure 1: Video surveillance General frame work

Tracking the objects in the video arrangement is illuminating the similar object in a arrangement of frames using the object special attributes described in the form of characteristics. Generally, the recognition and tracking phenomenon is constantly succeed by tracking in video surveillance system. Tracking is accomplished from single frame to another frame using tracking algorithms [10].

This work approves algorithm for observing pedestrian crossings and recognize the human or vehicle and other objects. With the trained pattern matching, the activities are segmented with the help of image processing approaches for analyzation and behavior and motion detection [18-19]. In some situations, it is necessary to examine the behavior of persons and resolve whether their behavior is careful or not such as the response of pedestrian at a crowded area. In this process the movement of the objects is identified from video scene and cause the illustration of the action [20]. The proposed searching and modeling phenomena of human crowds choosing the most comparable approaches names as microscopic, macroscopic and mesoscopic. Microscopic means the pedestrian or any single person detected based on the place, speed and movement specifications are not considered. Macroscopic explains pedestrian are recognized based on the mean pedestrian capacity and motion of the pedestrian which plays major role in human crowd safety and it can be used for fast analyzing in critical conditions. Security of large number of people dependent on the quantity of the pedestrian walk sustainably at various high crowded areas. Mesoscopic means that pedestrian is detected based on the location, speed which depends on the distribution function.

The another action of video surveillance frame work diagram illustrates the personal reorganization. Human face and manner of walking on the necessary features like authentication and biomedical features which can be used for human recognition [6], The main aim of this is to discuss the difficulties affected in video surveillance system. Repeatedly a group of pedestrian detection and tracking methods used for motion in fixed camera into extensive classification to provide an instructive survey of respective techniques in each and every group. The main contribution of person authentication is the provisional approach and analysis of openly accessible data sets of pedestrians with specifications and surrounding restrictions. Determine the challenges of pedestrian detection and tracking and video arrangement pick up by a moving and fixed camera. Classifying the techniques of pedestrian detection and tracking in various types based on the concepts of techniques acceptance to every group and discussed the proposed developments for each technique.

The motion of object is a flexible thing that movement after time in image pattern of a video encapsulated by a fixed and also the motion of the camera. In video surveillance system the province of concern is a pedestrian that requires to be detect and tracked in the video [11]. This is the difficult task to do because of many difficulties and issues collect the video are

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 3, April 2022
Impact Factor: 6.252

light changes, instantaneous moment, complicated environment, shadows, object distortions, etc. The pedestrian detection and tracking challenges are diverse position, crowd compactness and location tracking.

Several factors associated to video accretions techniques, condensation methods, constancy of cameras can directly influence the virtuosity of video sequence [17]. Noise is other factor that can critically decline the virtuosity of video sequence in some instances the gadgets are used for video encapsulates cause constraint for designing object detection and tracking. Moreover, camera vibration causes the reduction of video sequence quality.

To observe this numerous video cameras are placed and information are accomplished and this structure will work for automated vehicle, used to avoid accidents for other vehicles. The study is coordinated as follows. Literature Survey has been represented in section II, section III represents the materials and methods used in the work, section IV represents existing work details and drawbacks of the existing work. Section V represents need for machine learning, The algorithm used represents in section VI, section VII represents proposed work, section VIII represents result which comparison of existing work.

III. Materials and Methods

The investigation expresses a pedestrian crosswalk detection system for operating an intelligent vehicle. This work accommodates a morphological filter ensured by horizontal projection to estimate pedestrian cross walk regions. Then self-sufficient realization method is engaged to analyze the pedestrian walk region. This work improves an accuracy but this system declines to observe the traffic signal and its authority requirements. The average time taken to achieve single frame is about 57. 2milliseconds.The road scene analyzation process has thrive using algorithms which extract the road scene from the 3D scene. The advanced algorithms produce trained labels from general image dataset. From the trained labels, the actual time road scenes can be compared, segmented and acknowledge to the road images. This work combines virtual images with offline images. Compared to the reference line image analysis method, there is improvement of accuracy in the noisy image.

The latest work for vehicle based mobile mapping system for automatic pedestrian crossing. The initialization methods introduce projection filtering, single- eye vision and predict the information analysis. This system gains high reliability, recall rate, effectiveness and precision values. The model can be working in different environmental conditions. The main advantage of this system is to determine completely even degradation and defacement traffic situation of roads. recognition of pedestrians has been done in the work by combining movement information with image intensity information.

This research work more accurate, even in the low resolve of images such as rain, snow and other environment conditions. A proposed method for detecting the pedestrian by combining HOG and optical flow method. An information driven approaches has been developed to detect pedestrian crossing to inspect the images. HOG method worked with SVM algorithm employed. This work can give a navigational management for visually impairment people.

IV. Existing Work

A part of the existing work was concerned to assemble the applicable videos of pedestrians crossing close by the crossway. The work explores the attributes of pedestrian detection and behavior analysis. Pedestrian detection is an imperative and powerful task in any video surveillance system as it gives basic information for morphological comprehension of the video footages. The behavior study introduces various behavior factors like eye gaze, facial gesture, visual communications, pose, body movement and physical movements like hand movements etc., Advanced machine learning algorithms are used to solve these tasks. Some of the drawbacks are present in existing system represented as below.

- In efficient feature extraction
- Due to the use of static reference line, it doesn't give the exact position of the pedestrian.
- It requires more processing time.

The block diagram representation of the Existing system as shown below

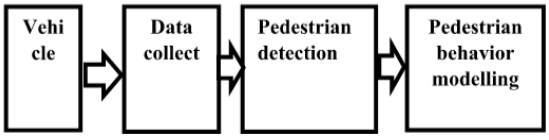


Figure: Existing system Block diagram

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 3, April 2022
Impact Factor: 6.252

V. NEED FOR MACHINE LEARNING

Machine learning is the field of artificial intelligence that uses statistical approaches to learn camouflaged patterns from prevailing data and to make choices about unseen records. The main task of machine learning is establishing a general model on the possible dispensation of training examples and then generalizing experience of concealed examples. The studying process is based on the quality of data display. Smart traffic lights have been all over for a while, but now the wireless has come up with a new innovation hold up to make pedestrians crossings with video surveillance as convenient as possible. More respectively the new system uses the latest generation technology like cameras, Artificial intelligence and machine learning to control the traffic lights and decide as the preference as its importance. If it is in point of fact which detects a certain number of people wants to cross the road, the system can then look into a set of rules to change the traffic lights. The rules can be personalized according to the need of the local authorities and requirements of traffic managers.

Machine learning has the not only replaced the way surveillance video is view, but also the methods it can be leveraged for improving security and operational proficiency it decreases the need for human beings to detect each and every video surveillance footage and recorded video feed and it grants the security staff to react to the situations, analyzing the footage, display trends so that they can use the video footage more efficiently. This has large ranged, positive influence on both security and operations of local governments.

VI. ALGORITHM USED

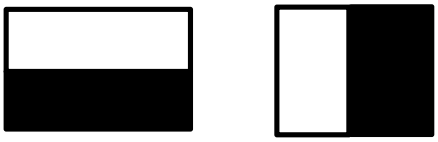
The algorithm used is HAARCASCADE which will be demonstrated by four stages. They are:

- Designing HAAR features
- Generating Integral Images
- Applying Adaboost
- Organizing Cascading Classifiers

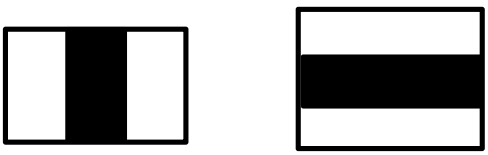
This algorithm contains two types of images. They are Positive images – Consists of Faces to identify and Negative images – Do not consists of faces and objects that will detect

6.1 Designing Haar Features

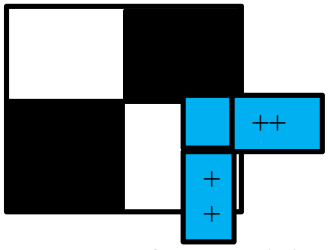
A Haar features is accurate design that will execute on adjoining rectangular location in a window. The designing involves adding the pixel concentration in every region and designing the differences between the sums. It is used for integral images.



a) Edge type of characteristics



b) Line type of characteristics

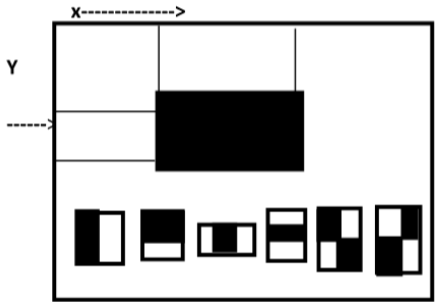


c) Four rectangle type of characteristics

6.2 Generating Integral Images

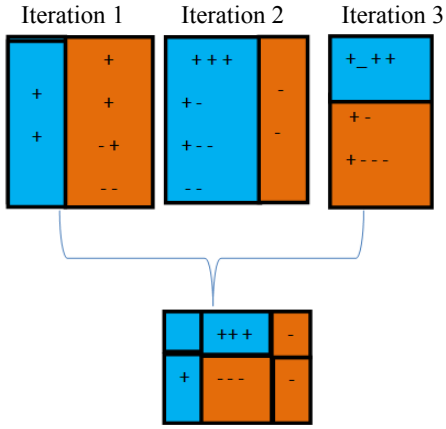
Integral images will increase the Haar features, it directly creates the sub rectangles and generates array reference for every sub-rectangle and it is used for detection.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 3, April 2022
Impact Factor: 6.252



6.3 Applying Ada Boost

Adaboost training consists of lacking classifiers to produce a powerful classifier algorithm which is useful to detect objects. It merges the weak learners into the strong learner using Cascading classifiers



Final Classifier/Strong Classifier Organizing

Cascading Classifiers

The cascade classifier is powered by a sequence of phases, whereas every phase is series of lacking learners. Lacking learners are instructed applying boosting, by which permits for a very error free classifier from the mean projection of all lacking learners. In this low false negative rate should be exaggerated which leads to improvement of detection.

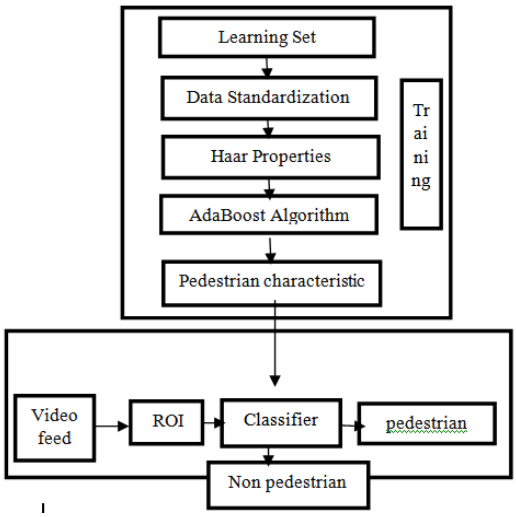


Figure 3: Flow chart of Haar pedestrian detection.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 3, April 2022
Impact Factor: 6.252

Applications of Haar Cascades

In Facial Recognition, Electronic devices and security protocols helps the Haar cascades to decide the authority of the user for secure login.

In Image exploration and Object Identification, In extension of Facial recognition, any collection of objects can be examined by using a Haar cascade computer vision algorithm.

VII. PROPOSED WORK

Deep learning is one of the areas of study of machine learning. By studying the machine, the quality of an image can be obtained automatically. The deep learning images can be inserted straightforwardly into a deep neural network that grasp the characteristics dynamically, which generally requires millions of images to get the be result and also counting in-depth and needs a higher production. This method is most productive, supervised and cost-efficient machine learning approach. These neural networks repeatedly utilize training labels and grasp the algorithms by steadily improving the size of data, competence upgraded.

The operation of this studying procedure relay on two stages, so called training and interference phases. Training phase includes labeling huge amount of data and defining their flexible characteristics. Interference phase is to come to an end that label new and abstraction of data by using their previous acknowledgement with maximum efficiency. This interference phase also known as structured learning, which consist of numerous layers involve nonlinear processing units to change and extract feature. Each and every consecutive layer obtains result from the preceding layer as the given input data.

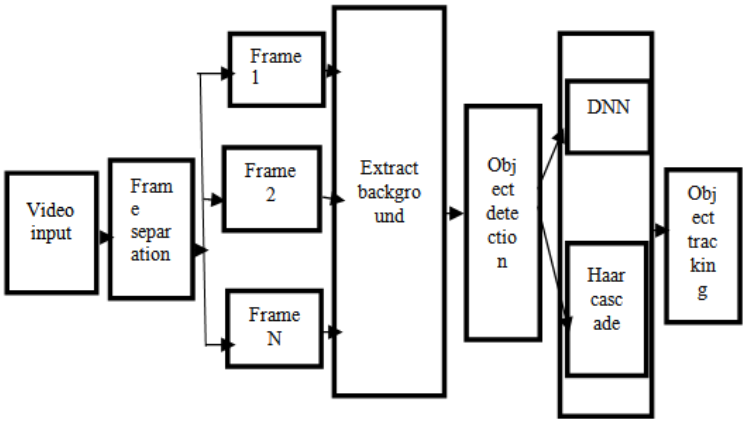


Figure 4: Proposed work Block diagram

Deep neural network uses fundamental counting unit, a neuron that acquire several signals propositionally with weight and transfers the gathered signals to the non-linear task to generate the output. Neural networks require a large number of input data, the more data that is fed into the network, it will better generalize with accuracy. The deep learning utilizes a set of attributes used for numerous features and huge amount of data set for individual features then extracts a categorization and generates an integrated classification to inspect a different application. The major attributes which include the deep learning process depends on the nonlinear processing in multiple layers and supervised and un supervised learning these learning techniques are attached.

In proposed system we are using yolo algorithm. which is used advanced form of image classification and detect the images that belongs to a predefined set of classes, a neural network predicts in an image and points to the form of bounding boxes. using this algorithm Tasks like detection, recognition and find widespread applicability in real world scenarios are possible. yolo algorithm functions by segmenting photograph into the N grids, every grid having an identical dimensional area. Each of these N grids accountable for detection.

7.1 Data set for Training and Testing

We choose coco dataset to train and test our model. It is one of the most widely used pedestrian datasets. It consists about more than 328K images, about 25,00,000 labeled instances and 56,000 unique pedestrians are annotated. Every labelled individual is annotated with an instance and mapping between image pixels that belongs to the pedestrian body and model.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 3, April 2022
Impact Factor: 6.252

The annotations are publicly available only for training and testament of pictures. This database includes thousands and millions of images already labeled for training, which defines 91 classes but data available in 80 classes only. This is a set of challenging high-quality datasets for computer vision used for large scale detection, segmentation.

7.2 Feature Extraction

In this proposed system, we are also implementing the crowd detection feature, whenever the crowd was more in pedestrians and also it helps to maintain social distance to cross the pathway which different scenarios with different color indications. If one pedestrian is near to other pedestrian which describes serious violation which represent red color bounding box. If the pedestrian maintains some what social distance which represents yellow color bounding box which describes abnormal violation and the pedestrian maintain the social distance as much which represents green color bounding box which describes safe level. The proposed work also alerts the abnormal violations, serious violations to track the pedestrians at the traffic signals.

7.3 Tools Used

1. IDLE Software
2. Dnn and Haar cascade algorithms
3. Coco data sets
4. Yolo Algorithm

7.4 Algorithm Steps

- Learn a single simple classifier
- Classify the data
- Look at where the data makes errors
- Reweight the data so that the inputs where we make errors to get larger weight in the learning process
- Now learn a simple classifier on weighted data
- Combine the classifiers in figure and weight the data according to where it makes errors
- Learn a classifier on the weighted data and so on until we learn T simple classifiers
- This process is called Boosting -works very well.

VIII. RESULT

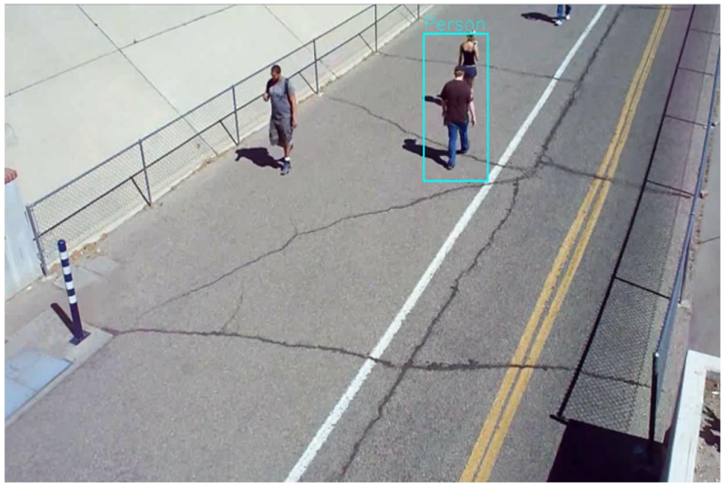


Figure 5: Existing work result

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 3, April 2022
Impact Factor: 6.252

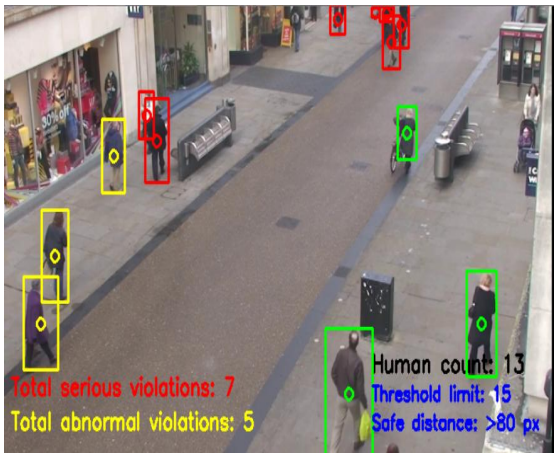


Figure 6: Proposed work result with recorded video

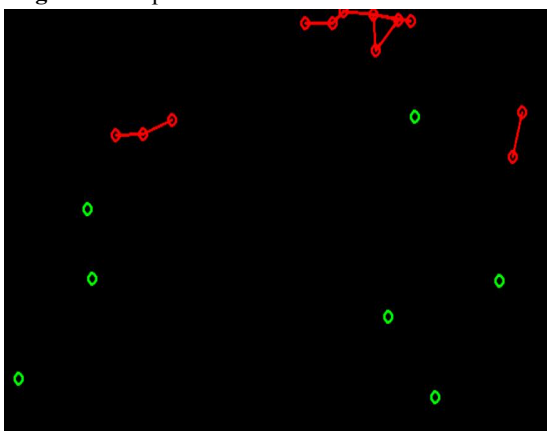


Figure 7: Proposed Work mapping with recorded video

8.1 Comparison Tables

Features	Distance of The Existing work	Distance of The Proposed work
Sample 1	96.7	97.05
Sample 2	97.25	97.59
Sample 3	97.68	97.9
Sample 4	99.05	99.52
Sample 5	99.11	99.89

Features	Speed of the Existing work	Speed of the Proposed work
Sample 1	211.54	213.57
Sample 2	214.73	217.02
Sample 3	211.57	211.68
Sample 4	210.01	210.50
Sample 5	205.95	206.09

Features	Time of the Existing work	Time of the Proposed work
Sample 1	0.456	0.432
Sample 2	0.456	0.432
Sample 3	0.456	0.432

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 3, April 2022
Impact Factor: 6.252

Sample 4	0.456	0.432
Sample 5	0.456	0.432

IX. CONCLUSION

In this work it illustrates and revival of the methodologies, approaches and steps are elaborated in video surveillance. Based on the literature survey many of the existing methodologies proposed by before time which locates the difficulties, issues, at hand solutions and trends for tracing in video surveillance system.

In this proposed system pedestrian detection is proficient through multiple classifiers and detection rates are possible with some of the existing methods to schedule with regards to the performance, speed and counting time. The tracking method is responsible for tracking the human activities by extraordinary features. The methods which are useful in the proposed system will increase the computational speed. Hence it can be used for various applications that involves in object tracking.

REFERENCES

[1]. Minsky M, Kurzweil R, Mann S. IEEE International Symposium on technology and society, Toronto, Ontario, Canada, 27th-29th June; The Society of intelligent surveillance.2013. p.no:13-18

[2]. Foresti GL, Micheloni C, Snidaro L, Remagnino P, Ellis T. Active video surveillance system: The low-level image and video processing methods needed for implementation.in: IEEE Signal Processing magazine, Vol.22(2), March 2005. P.no:25-39

[3]. Gawande U, Golhar Y. Biometric security system: A rigorous review of unimodal and multimodal biometric methodologies. International Journal of biometrics, April 2018. P.no:10(2)

[4]. Zafeiriou S, Zhang C, Zhang Z, A survey on face detection in the wild: Past, Present and future. International Journal of computer vision image understand. Sep.2015. p.no: 1-25

[5]. Teddy K, Lin W. A survey on behavior segmentation and video surveillance applications. Video surveillance.2011. p.no:281-293

[6]. Gawande U, Golhar Y, Hajari K. Biometric based security system: issues and challenges, Computational intelligent techniques in signal processing for multimedia security Vol.660.Cham: Springer; 2017.p.no:150-175

[7]. Gawande U, Zaveri M, Kapur A. The work algorithm for advanced level fusion using SVM classifier for multibiometric based person authentication. Applied Computational intelligence and soft computing, July 2013. P.no: 1-15

[8]. Li X, Flohr F, Yang Y, Xiong H, Braun M, Pan S. A new bench mark for video surveillance-based cyclist detection. In:IEEE intelligent vehicles Symposium, Sweden, 19th-22nd . June2016. P.no: 1028-1035

[9]. Campbell D, Petersson L. GOGMA: Globally optimal Gaussian mixture alignment. In:IEEEconference on computer vision and pattern identification, Los Vegas, USA, IEEE.2016.

[10]. Pellegrini S, Ess A, Van Gool L. Wrong turn-No dead end: A stochastic pedestrian motion model. International Workshop on Socially intelligent surveillance and monitoring, San Francisco, CA, USA,13th-18th June 2010.

[11]. St- Charles PL, Bilodeau Ga, Bergevin R. subsense: A universal change detection method with local adaptive sensitivity. IEEE Transactions on image processing.2015.p.no:359-378

[12]. Cogun F, Cetin AE. object tracking under illumination variations using the spectrum of 2D characteristics of the target. IEEE International workshop on Multimedia Signal Processing.2010. p.no:521-530

[13]. Heikkila M, Pietikainen M.A text-based method for modelling the background and detecting moving objects. IEEE Transactions on pattern Analysis and machine intelligence.2006.28(4) p.no:657-665

[14]. Shen C, Lin X, Shi Y. Moving object tracking under change in the illumination conditions. Pattern Recognition Letters.2006.27(14). Pg.no:1632-1645

[15]. Lee YB. A real-time color-based object tracking robust to irregular illumination situations. In: IEEE International Conference on Robotics and Automation,21st-26th. May 2001. P.no:1659-1665

[16]. Tokmakov P, Alahari K, Schmid C. Learning moment patterns in videos. In: IEEE Conference on Computer Vision and Pattern identification, CVPI.2017.p.no:531-540

[17]. Jepson AD, Fleet DJ, EI- Maraghi TF. Robust online appearance models for visual tracking. IEEE Transactions on Pattern Analysis and Machine intelligence.2003.25(10).p.no:1296-1315

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 3, April 2022
Impact Factor: 6.252

[18]. Athnesious SP. Systematic Illustration on object tracking methods in video. International Journal of Advanced Research in computer Engineering and Technology.2012.1(8).p.no:242-250

[19]. Chaaraoui A, Climent-Perez P. A review on vision techniques applied to human behavior analysis for ambientassisted living. Expert System Application.Sep.2012.39(12).p.no:10873-10890

[20]. Elaiw A, AI-Turki Y, Alghamdi M.A critical analysis of behavioral dynamics of crowd-From a modelling strategy to kinetic theory methods. MDPI Symmetry Journal.July2019.11(851).p.no:1-10