International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Implementation of Cars and Pedestrian Detection

Akshay Deshmukh [1], Sanchit Gupta[2], Ambalika Donge[3], Piyush Katolkar[4], Abhishek Tipre[5], Prof. Rakesh A.Bairagi[6].

[1,2,3,4,5]BE Student, [6]Assistant Professor, Depertment Of Computer Science & Engineering Govindrao Wanjari College Of Engineering And Technology Nagpur, RTMNU, Nagpur, Maharashtra, India

A B S T R A C T

Electronic systems that can identify pedestrians in front of a vehicle and forecast vehicle-to-pedestrian collisions must be specified, implemented, and evaluated. Vehicle-to-pedestrian collisions were categorised into eleven different scenarios in this study. The key features of vehicle-to-pedestrian collisions have been established.The statistical behaviours of the various systems involved were modelled (vehicle, pedestrian, environment, and advanced driver assistance gadgets). Then, for crucial automobile and pedestrian road conditions, Monte-Carlo simulations were run. The created simulation tool enables the evaluation and validation of possible innovative systems' performance.Intelligent video surveillance, intelligent transportation, automotive autonomous driving, and driver-assistance systems all use cars and pedestrian detection. For the implementation of automobiles and pedestrian detection in a video segment, we chose OpenCV as the programming tool. This programme will be written in Python and will make use of OpenCV.

KeyWords: OpenCV,MachineLearning,Python,PedestrianDetection,VehicleDetection, Computer vision.

INTRODUCTION

In the subject of computer vision, there has been a lot of progress recently. Using the camera in video sequences, object tracking is the most frequent method for recognising moving things beyond time. Object tracking's major goal is to link the target objects' shape or features, as well as their location, in successive video sequences. As a result, object classification and detection are critical for object tracking in computer vision applications. Furthermore, tracking is the first stage in recognising any moving objects in the frame.Furthermore, the observed objects can be classified as swaying trees, birds, humans, automobiles, and so on. Although object tracking using video sequences is a difficult task in image processing, it is possible. Occlusion of the object to scene, object to object, complex object motion, real-time processing needs, and the inappropriate or distorted shape of the item appear to be responsible for a number of other difficulties.

However, this form of tracking is currently employed in a variety of settings, including traffic monitoring, robot vision, surveillance and security, video communication, and public venues such as subway stations, airports, mass gatherings, and animation, to name a few. As a result, the application requires an ideal trade-off between processing, communication, and network accuracy. The number and type of cooperation performed among cameras for data gathering, dispensing, and processing to validate decisions and eliminate estimation mistakes and ambiguity is what drives computing and communication income.As a result, tracking can be defined as the process of identifying an object's orientation over time as the object moves across a scene.

Because of the proliferation of high-powered computers and the growing demand for automated surveillance systems and other applications, object tracking is becoming increasingly important in thefield of computer vision. It is mostly employed in the areas of automated surveillance, robotics monitoring, human-machine interface, motion-based recognition, vehicle navigation, traffic monitoring, and video indexing these days. A large number of these applications necessitate highly reliable and efficient tracking algorithms that adhere to real-time constraints and are challenging and sophisticated in terms of object movement, scale and appearance, scene illumination, and occlusion.The outcome of object tracking could be influenced by a discrepancy in one of the parameters. A vast variety of ways have been offered to address the above-mentioned challenges and others in object tracking. Cars and pedestrians will be the targets of this object tracking programme. The ability of machines to recognise suspicious objects and further identify their actions in a given context is a critical component of allowing machines to communicate with people in an effective and simple manner.The existing method for analysing and detecting suspicious objects usually necessitates the use of an exceptional marker attached to the suspicious object, which precludes the use of extensive technology. We strive to incorporate automobiles and pedestrian detection in this paper, as well as provide efficiency by including a time library.

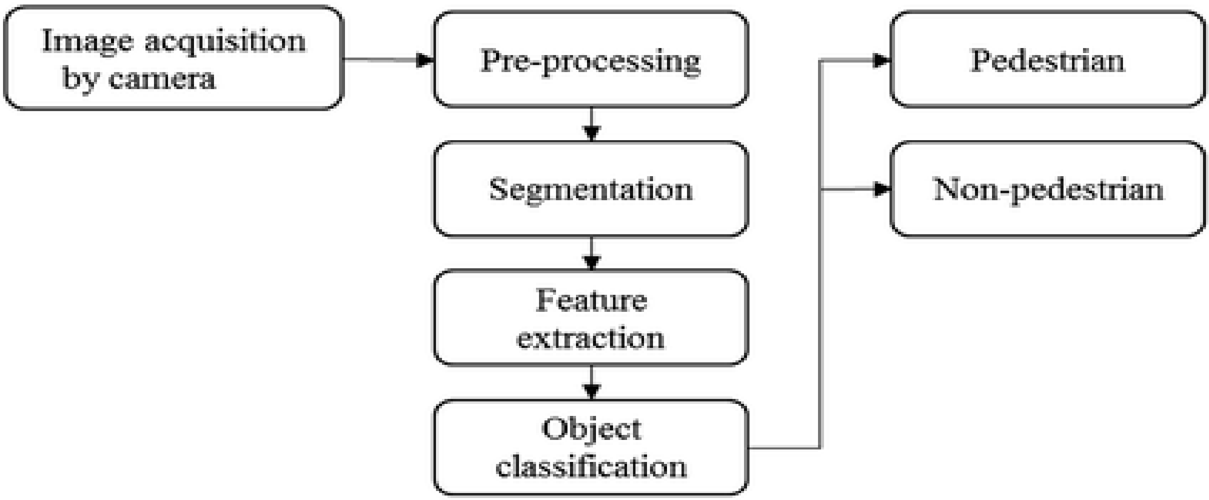
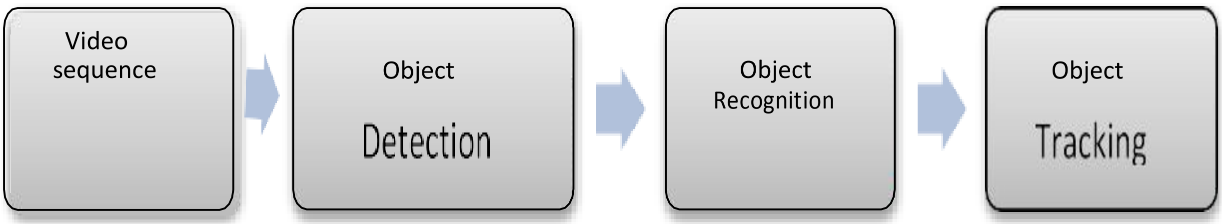


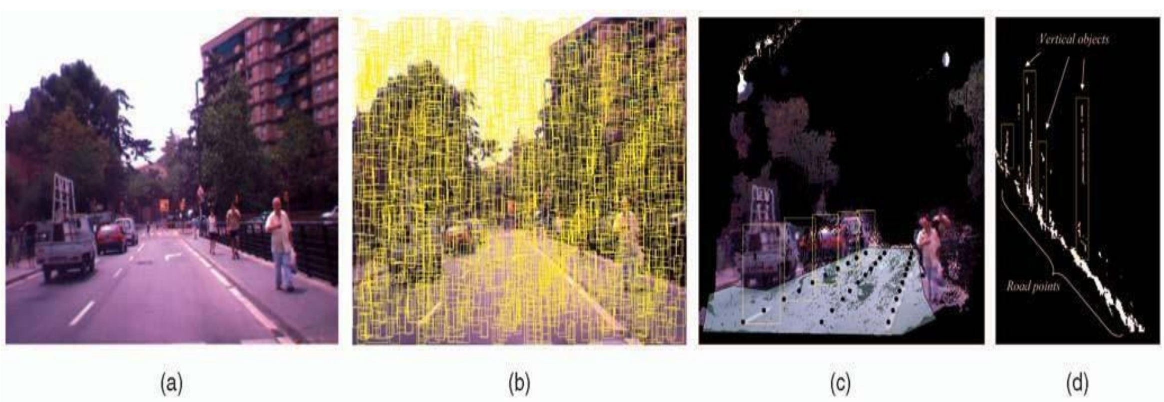
Figure 1 flowchart of system

2 LITERATURE SURVEY AND REVIEW

In the previous study the papers were only focusing on object detection or object tracking. (Ben Ayed et al., 2015; Najva and Bijoy, 2016; Ramya and Rajeswari, 2016; Risha and Kumar, 2016; Shen et al., 2013; Soundrapandiyan and Mouli, 2015; Viswanath et al., 2015) ,Object tracking (Bagherpour et al., 2012; Foytik et al., 2011; Lee et al., 2012; Poschmann etal., 2014; Yilmaz et al., 2006; Zhang et al., 2016) and Object recognition (Chakravarthy et al., 2015; Gang et al., 2010; Ha and Ko, 2015; Nair et al., 2011) for tracking the object using video sequences with the help of camera. These are discussed as follows. The basic flow diagram of an object tracking shown in shown in figure 2.

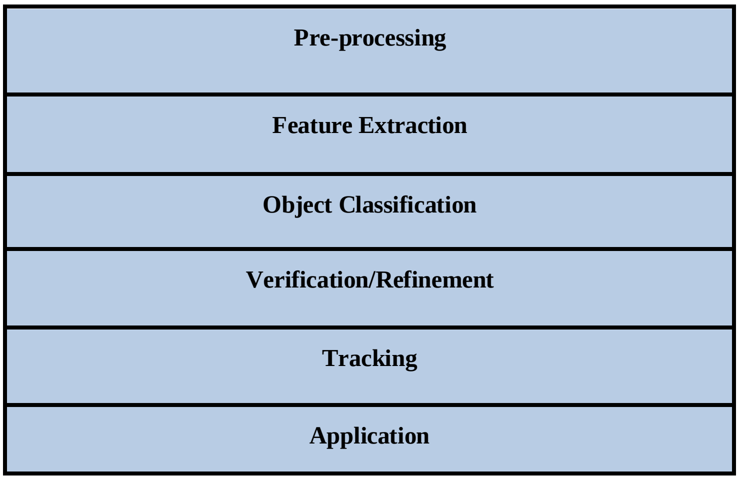
Figure2 TheBasicflowdiagramofObject tracking





The above diagram or figures representforeground segmentation schemes.Here,(a) Original image. (b) Exhaustive scan(just showing 10 percent of the ROIs). (c) Sketch of road scanning after road fitting in Euclidean space . (d) Results of v-disparityappliedtothe same frame.

3 WORKFLOWOFCARS AND PEDESTRIAN DETECTION



- Pre-processing

Pre-processing includes a variety of duties that are required at the start of the process, such as exposure time, gain changes, and camera calibration, to name a few. Some low-level modifications are made that aren't generally documented in ADAS literature, and some researchers have used these systems to target visual improvement. There are two methods for doing so: monocular and stereo vision.

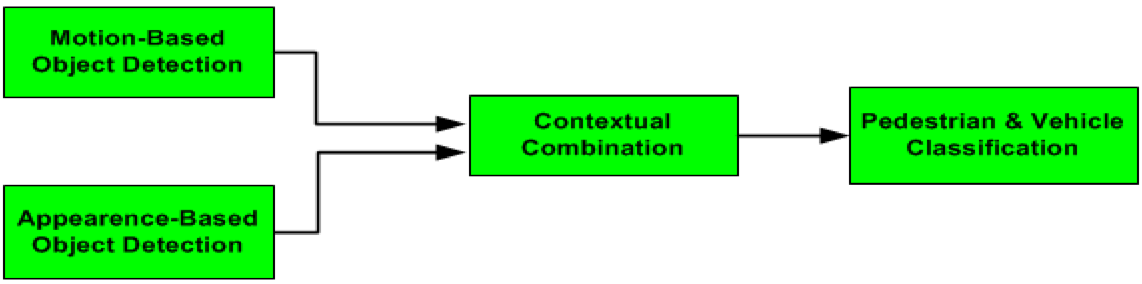
- Feature Extraction

Wecanclassify thedifferentfeaturesas:

- General features: Color, texture, and shape are examples of general qualities that are not application specific. They can be further classified into the following categories based on the level of abstraction: Features at the pixel level: Color and position are examples of features that are determined at each pixel.
- Local features: Local features are those features that are calculated over the results of subdivision of the image bandon image segmentationor edge detection.
- Globalfeatures:Thesearethefeaturesthat arecalculatedover anentireimageorjustregular sub-areaofanimage.
- Domain-specificfeatures:Domain-specificfeaturesarenothingbutapplicationdependentfeaturessuchasfingerprints, human faces and also conceptual features. These features are nothing but a synthesis of low-levelfeaturesfor aspecific domain.

All available features are divided into two categories: low-level features and high-level features. Low-level features are retrieved from the source images, whereas high-level features are extracted from low-level features.

- Object Classification



- Verification/Refinement

A tracking module is used by the majority of advanced systems to track recognised autos and people over time. This step is important for several reasons, including avoiding false detections over time, predicting future positions of cars and pedestrians and thus providing pre-candidates to the foreground segmentation algorithm, and, at a higher level, inferring useful information about pedestrian and car behaviour (e.g., walking/driving direction).

- Tracking

A tracking module is used by the majority of advanced systems to track recognised autos and people over time. This step is important for several reasons, including avoiding false detections over time, predicting future positions of cars and pedestrians and thus providing pre-candidates to the foreground segmentation algorithm, and, at a higher level, inferring useful information about pedestrian and car behaviour (e.g., walking/driving direction).

- Application

Pedestrian detection is an essential and significant task in any intelligent video surveillance system, as it provides the fundamental

information for semantic understanding of the video footages. It has an obvious extension to automotive applications due to the potential for improving safety systems.
For avoiding accidents and also avoiding traffic problems using smart techniques this system is useful.

4 METHODOLOGY AND IMPLEMENTATION

- **Computer Vision**

Computer vision is a field of study which enables computers to replicate the human visual system. It's a subset of artificial intelligence which collects information from digital images or videos and processes them to define the attributes. The entire process involves image acquiring, screening, analysing, identifying and extracting information. This extensive processing helps computers to understand any visual content and act on it accordingly.
Computer vision projects translate digital visual content into explicit descriptions to gather multi-dimensional data. This data is then turned into computer-readable language to aid the decision-making process. The main objective of this branch of artificial intelligence is to teach machines to collect information from pixels.

- o **Object detection:**

Object detection is a computer vision technique that makes a system able to locate objects in an image or video. This bounded box] from the segmentation can help a human or driver to identify and locate the objects in a matter of seconds and quicker than unprocessed images. Therefore, the goal of object detection is to implement this intelligence to a computer. There are different techniques to perform object detection. Popular deep learning–based approaches using convolutional neural networks CNNs, such as YOLO and SSD, which will automatically learn to detect objects within frames. To explain further on object detection, more knowledge on machine learning is presented below. The Figure 1 below shows the difference between image recognition and object detection.

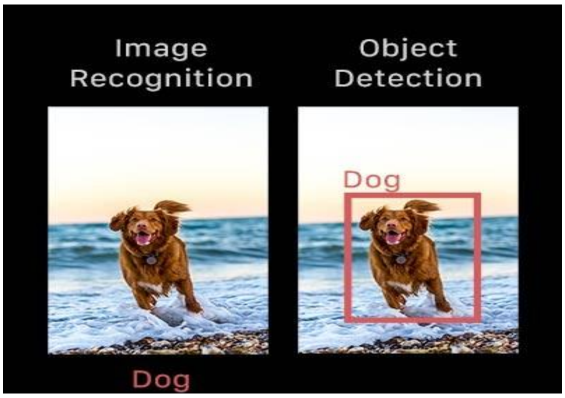


Figure 3The lines around the dog is a boundary b

Now that we know what does Computer Vision mean and some of its application, let's dive into the implementation of it. To implement various examples of computer vision, we are going to use the **OpenCV** library.

OpenCV (Open Source Computer Vision Library: http://opencv.org) is an open-source BSD-licensed library that includes several hundreds of computer vision algorithms.

Haar Cascade Classifiers : We will implement our use case using the Haar Cascade classifier. Haar Cascade classifier is an effective object detection approach which was proposed by Paul Viola and Michael Jones in their paper, "**Rapid Object Detection using a Boosted Cascade of Simple Features**" in 2001.

So, let's see what these Haar Cascade Classifiers are all about. This is a machine learning approach in which a cascade function is trained from a large number of positive and negative images. It is then used to detect the things in the other images based on the training.

So, they're big individual.xml files with a lot of features, and each xml file relates to a very unique type of use case.

- o **Face Detection with the Viola-Jones Object Detection Framework**

The framework consists of two phases: Training and Testing/Application. Let's look at each of them one by one.

1. Training

The goal of this phase is to produce a Cascade Classifier for a face that is able to accurately classify a face and discard non-faces quickly. To achieve that, you must first prepare your training data and then construct a Cascade Classifier by using a modified AdaBoost Algorithm on that training data.

a) Data Preparation

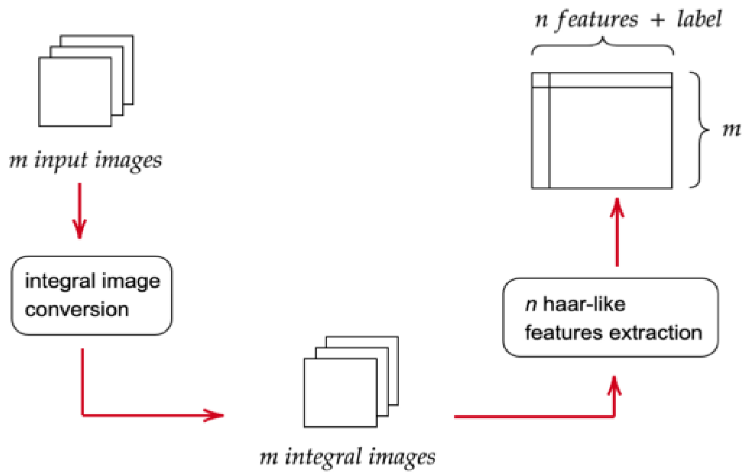


Figure 4The Data Preparation Process

b) Constructing a Cascade Classifier with a modified AdaBoost Algorithm

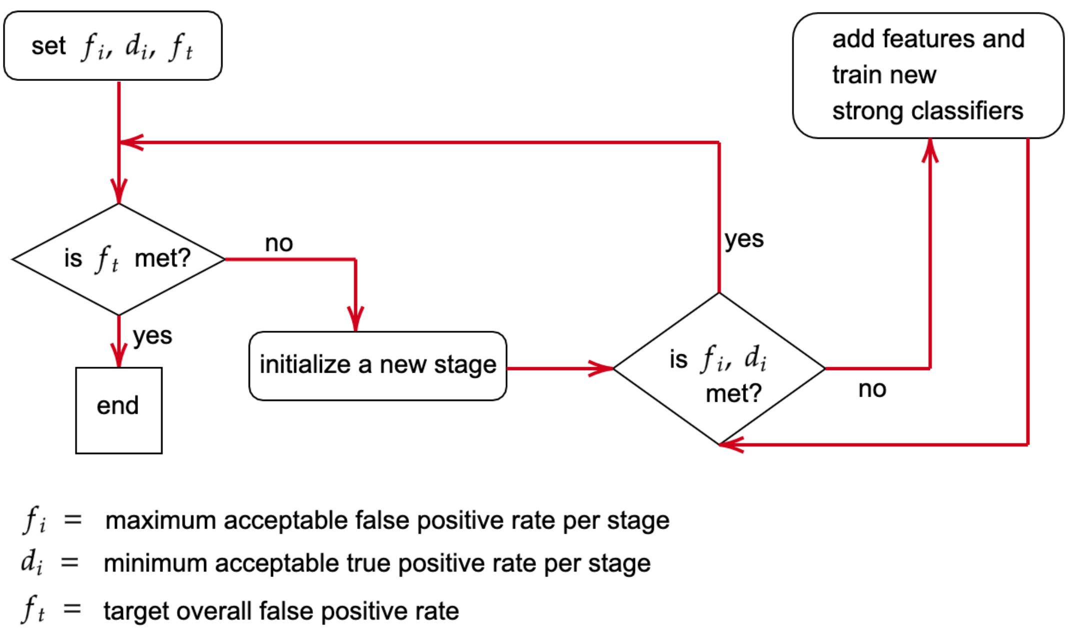


Figure 5 The process to construct a cascade classifier

2. Testing/Application

Figure 6Sliding Windows Detection Process in the Viola-Jones Object Detection Framework

Imagine that we need to detect faces in the above image. Viola & Jones (2001) use a sliding window approach where window of various scales are slid across the entire image. The scale factor and the shifting step size are parameters for you to decide upon. So for the above image, there are *m* sub-windows to evaluate. For a sub-window *i*, the framework resize the image of that sub-window to a base size of 24 x 24 (to match training data), convert it into an integral image and feed it through a Cascade Classifier produced during the training phase. A face is detected if a sub-window passes through all the stages in the Cascade Classifier.

- Machine learning

Machine learning is an area within Artificial Intelligence (AI) that focuses on decision making and predictions. The primary aim is to allow the computer to be further developed without any human intervention, it will train based on observations or data. The machine learning algorithms are often categorised as supervised or unsupervised, where supervised applies what has been learned in the past to new data using labelled examples.

- Supervised learning - Supervised learning is a machine learning task, where some data is labeled and input is given to the algorithm which then can learn from in order to classify the label. The aim of supervised learning is to approximate the mapping function so, that when given new input dataset, it can predict the right output for that dataset. It is a algorithm learning process that iteratively continuous until it achieves an acceptable level of performance. The problem that occurs can be further grouped into classification and regression problems.
  • Regression: is the problem of predicting a continuous quantity output for an example.
  • Classification: is the problem of predicting a discrete class label output for an example.

- Unsupervised learning - Unsupervised learning is a technique of machine learning, where the data provided is not labeled to classify or predict. The algorithm instead detects patterns which may not be obvious, or gains insight of processed data. Therefore, unsupervised learning is very useful when handling big data, but it is less accurate and, a trustworthy method compared with supervised
  learning. The problem that occurs can be further grouped into clustering and association problems.
  • Clustering: Clustering is mainly a task of grouping a set of uncategorized data by finding structure or patterns.
  • Association: Association is a unsupervised technique that is dealing to find relationships between variables in big databases.

Object detection approaches such as template matching and simple part-based models were used in the early days [e.g., Fischler and Elschlager (1973)]. Later, statistical classifier-based approaches (e.g., Neural Networks, SVM, Adaboost, Bayes, etc.) were introduced [e.g., Osuna et al. (1997), Rowley et al. (1998), Sung and Poggio (1998), Schneiderman and Kanade (2000), Yang et al. (2000a,b), Fleuret and Geman (2001), Romdhani et al. (This first successful family of object detectors, all based on statistical classifiers, laid the groundwork for the majority of subsequent research in terms of training and evaluation procedures, as well as classification methodologies.
In order to detect the things showing in the image at multiple scales and positions, most object detection systems use the same basic strategy, known as sliding window: an exhaustive search is used to detect the objects appearing in the image at different scales and places. This

search employs a classifier, which is the detector's fundamental component and determines whether a particular image patch relates to the item or not. Given that the classifier only works at a certain scale and patch size, many downscaled versions of the input image are created, and the classifier is used to classify all possible patches of the given size for each of the downscaled versions.

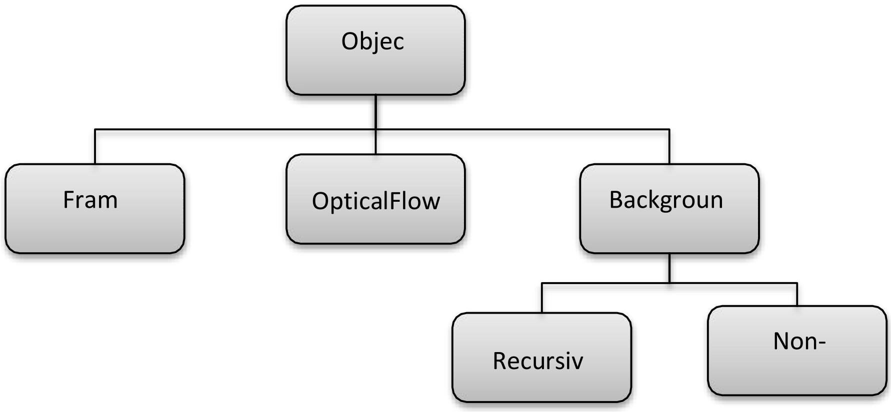


Figure 7 Types of object detection method

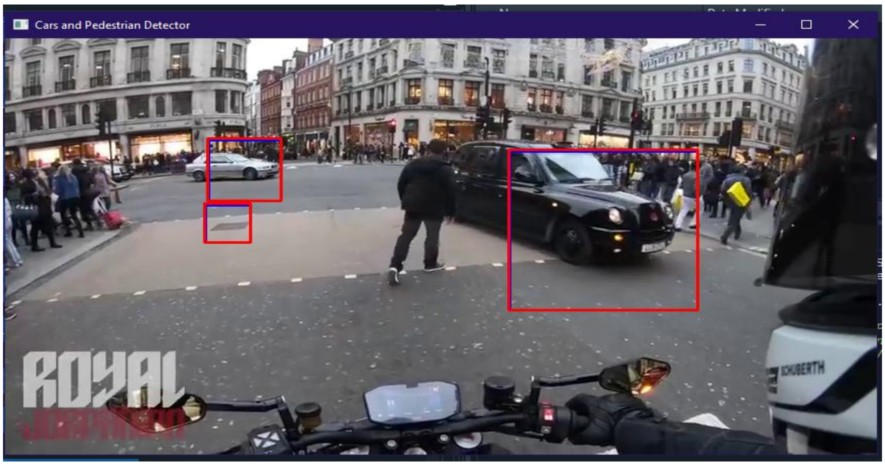
Object detection is important in various applications, but especially in video surveillance applications where things are discovered using video footage (Amandeep and Goyal, 2015). In Figure 6, various forms of object detection are depicted.

Comparative Study of Object Detection technique :

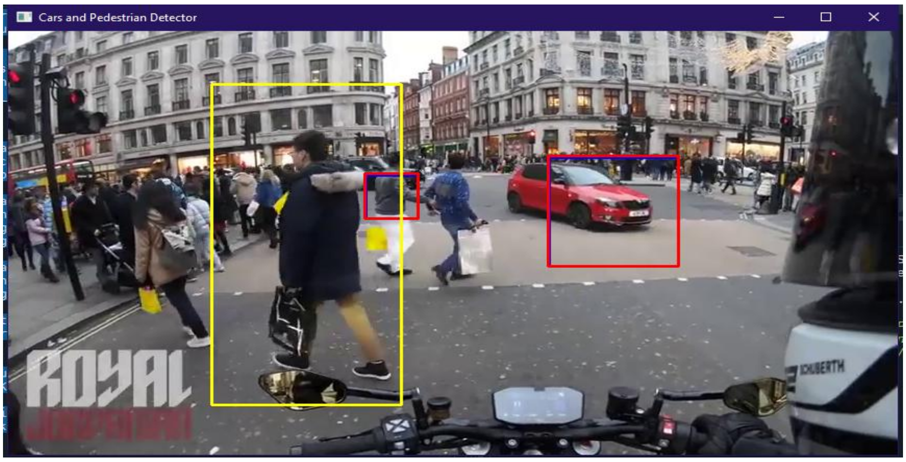
Object D Method		**Basic Principle**	**Computational Time**	**Accuracy**	**Comments**
Temporal Differencing		Pixel-wise Subtraction of Current & Background frame	Low	High	Easy to implement (Chate et al.,2012;MohanandResmi,2014) Sensitive to dynamic changes(Haritaoglu etal.,2000) Needs background frame with still objects (Mohan and Resmi,2014)
Backgro und Subtract ion	Frame Differencing	Current frame is subtracted from Background frame	Low to Moderate	Moderate to High	Simplest background Subtraction(AldhaheriandEdirisinghe,2014; Haritaogluetal.,2000) Cannot be used for real-time applications(MohanandResmi, 2014)
	Approxi mate Median	Simple subtraction between median Frame &test frame	Low to moderate	Moder ate	No need for adequate background modeling (Aldhaheri andEdirisinghe,2014) Requiresa buffer with recent pixel values (Aldhaheri andEdirisinghe,2014)
	Running Gaussian Average	Based on Gaussianprobabilit y densityfunctionofpi xels	Moderatet ohigh	Moder ate	Much suitable for real-time applications (Aldhaheri andEdirisinghe,2014) Statistical calculations consumes moretime
	Mixture of Gaussian	Based on multimodal distribution	Moderate tohigh	Moderate to high	Low memory requirement(Zhiqiangetal.,2006) Cannot cope up with objects aswellasnoise(TaoZhangetal., 2010)
Optical Flow		Usesoptical flow distribution characteristicsof Pixelsof object	Moderate to high	High	This approach offers entire movingdata(Krishnaetal.,2011) however require more calculations

6 RESULT AND DISCUSSION

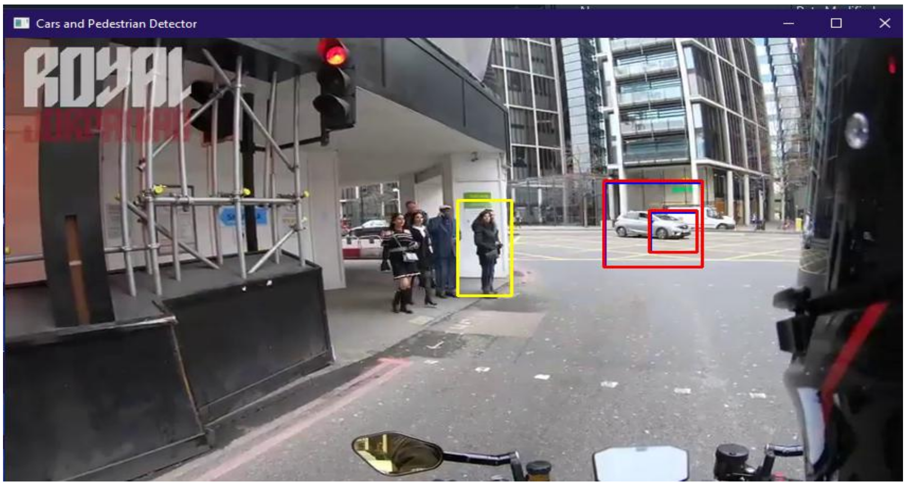
Video(.mp4) file is given to the system as an input where the system will detect the cars and Pedestrian on road.



In above output image, we can see cars are detected with red colour rectangles.



In above output image, we can see one car is detected with red colour rectangle and people are getting detected using yellow rectangles.



Similarily, also in above screenshot of the video which is given as an input, pedestrian and cars are getting detected.

As we have given input a video, similarily live video can also be captured with the build in or external webcam of the system and cars as well as pedestrian can be detected.

7 CONCLUSION

Review of numerous object detection, tracking, recognition algorithms, feature descriptors, and segmentation methods based on video frames and various tracking technologies, as well as our suggested automobiles and pedestrian detection system are presented in this work. This application outperforms similar apps in terms of accuracy and efficiency. We have also explored numerous restrictions and advancements in relation to this proposed use, but it is now far too complex.

FUTURE SCOPE

We could create complicated video sequence simulations and test them using the same tracking algorithm. In the expected case, moving objects are occluded by an item of the same colour or by a larger occlusion with a longer occlusion time. The tracking algorithm will be more efficient and functional if the number of objects is increased.

For each individual pixel's intensity, weight parameters could be applied. In any image, if the intensity value is allocated as foreground based on the current frame, the likelihood that the foreground also has similar pixel coordinates is reduced, allowing the BG weightage for the pixel to be set to the lowest value possible. By adding a weightage that is lower than the beginning value, the previous pixel value can be removed with the least likelihood rather than the evolved scene.

REFERENCES

[1] Aldhaheri,A.R.&Edirisinghe,E.A.,2014.DetectionandClassificationofaMovingObjectinaVideoStream,in:Proc.oftheIntl.Conf. onAdvancesinComputing andInformation Technology.InstituteofResearch Engineersand Doctors.

[2] Aggarwal,A.,Biswas,S.,Singh,S.,Sural,S.&Majumdar,A.K.,2006.ObjectTrackingUsingBackgroundSubtractionandMotionEst imationinMPEGVideos,in7thAsian ConferenceonComputerVision.Springer- VerlagBerlinHeidelberg, ,pp.121–130.doi:10.1007/11612704_13

[3] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11):2278–2324, 1998.

[4] Wei Liu, DragomirAnguelov, DumitruErhan, Christian Szegedy, Scott E. Reed, Cheng-Yang Fu, and Alexander C. Berg. SSD: single shot multibox detector. CoRR, abs/1512.02325, 2015.

[5] Jay F. Nunamaker Jr., Minder Chen, and Titus D.M. Purdin. Systems development in information systems research. Journal of Management Information Systems, 7(3):89– 106, 1990.

[6] Risse B, Mangan M, Del Pero L, Webb B (2017) Visual tracking of small animals in cluttered natural environments using a freely moving camera. In: IEEE international conference on computer vision workshops (ICCVW) Venic, pp 2840–2849

[7] Luo Y, Yin D, Wang A, Wu W (2018) Pedestrian tracking insurveillance video based on modified CNN. Multimed Tools Appl 77:24041–24058

[8] Luo W, Yang B, Urtasun R (2018) Fast and furious: real-timeend-to-end 3D detection, tracking and motion forecasting with asingle convolutional net. IEEE/CVF conference on computer vision and pattern recognition. Salt Lake City, UT,pp 3569–3577

[9] Cen, K. (2016). *Study of Viola-Jones Real Time Face Detector*. Retrieved from https://web.stanford.edu/class/cs231a/prev_projects_2016/cs231a_final_report.pdf

[10] Jensen, O. H. (2008). *Implementing the Viola-Jones Face Detection Algorithm*. PhD thesis, Technical University of Denmark, DTU. Retrieved from https://pdfs.semanticscholar.org/40b1/0e330a5511a6a45f42c8b86da222504c717f.pdf

[11] Mordvintsev, A., &Abid, K. (2013). Face Detection using Haar Cascades. *OpenCV-Python Tutorial*. Retrieved from https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_objdetect/py_face_detection/py_face_detection.html

[12] Papageorgiou, C. P., Oren, M., &Poggio, T. (1998). A general framework for object detection. *Sixth International Conference on Computer Vision*. 555–562. doi: 10.1109/ICCV.1998.710772

[13] Ramsri. (2012, September 16). *Viola Jones face detection and tracking explained*. Retrieved from https://youtu.be/WfdYYNamHZ8

[14] Viola, P. & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1*, I-511 — I-518. doi: 10.1109/CVPR.2001.990517.