ADVANCED METHODS FOR MOVING OBJECT DETECTION IN PROBLEMATIC VIDEO FOR INTELLIGENT SURVEILLANCE SYSTEM

A Thesis submitted

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
IN
COMPUTER SCIENCE AND ENGINEERING

By

SNEHA MISHRA
18SCSE3010005

Supervisor

Dr. DILEEP KUMAR YADAV
Professor

SCHOOL OF COMPUTING SCIENCE & ENGINEERING
GALGOTIAS UNIVERSITY
Plot No 2, Sector 17-A Yamuna Expressway
Greater Noida, Uttar Pradesh
INDIA

AUGUST, 2023

CANDIDATE DECLARATION

I hereby certify that the work which is being presented in the thesis, entitled **"ADVANCED METHODS FOR MOVING OBJECT DETECTION IN PROBLEMATIC VIDEO FOR INTELLIGENT SURVEILLANCE SYSTEM"** in partial fulfillment of the requirements for the award of the degree of Doctor of Philosophy in Faculty of Computer Science and Engineering and submitted in Galgotias University, Uttar Pradesh is inauthentic record of my own work carried out during a period from October, 2018 under the supervision of **Dr. DILEEP KUMAR YADAV**, Professor, School of Computing Science and Engineering, Galgotias University, Greater Noida.

The matter embodied in this thesis has not been submitted by me for the award of any other degree or from any other University/Institute.

SNEHA MISHRA

18SCSE3010005

This is to certify that the above statement made by the candidate is correct to the best of our knowledge.

Dr. DILEEP KUMAR YADAV

Supervisor

SCSE

CERTIFICATE

This is to certify that **Ms. SNEHA MISHRA (Reg. No.18SCSE3010005)** has presented her pre-submission seminar of the thesis entitled **"ADVANCED METHODS FOR MOVING OBJECT DETECTION IN PROBLEMATIC VIDEO FOR INTELLIGENT SURVEILLANCE SYSTEM"** before the committee and summary is approved and forwarded to School Research Committee of Computing Science & Engineering, in the Faculty of Engineering & Technology, Galgotias University, Uttar Pradesh.

Dean – SCSE   Dean – PhD & PG

The Ph.D. Viva-Voice examination of **Ms. SNEHA MISHRA** Research Scholar, has been held on __________

Supervisor   External Examiner

APPROVAL SHEET

This thesis entitled "**ADVANCED METHODS FOR MOVING OBJECT DETECTION IN PROBLEMATIC VIDEO FOR INTELLIGENT SURVEILLANCE SYSTEM"** by **SNEHA MISHRA** is approved for the degree of Doctor of Philosophy.

Examiner Supervisor Chairman

ACKNOWLEDGEMENTS

The grace of the GOD is above all. Then come to the blessing of my supervisor Dr. Dileep Kumar Yadav, who always inspires me to take up a new task and supervised me, despite of his busy schedule. He has always given me a valuable guidance, support and encouragement throughout my Ph.D work. He provided me a productive and healthy environment. He supported me patiently and directed me with his vast experience. I sincerely believe that without his guidance this work would not be possible.

I must owe a special debt of gratitude to Hon'ble Chancellor Sir Mr. Suneel Galgtotia, CEO Dr. Dhruv Galgotia, and Hon'ble Vice-Chancellor Dr. K. Mallikharjuna Babu of Galgotias University for their valuable support throughout my research work. I owe my heartiest thanks and respect to Dr. Munish Sabarwal for providing me an inexpressible environment for my research work. Without him this work was not possible because of the productive environment that motivated me to complete this work.

Millions of thanks to Dr. K. Sampath Kumar and Dr. Janarthanan, for helping me out and motivating me at each milestone of review. I express my sincere thanks to Dr. Alok Katiyar, PhD coordinator and Dr. Ruby Dahiya as PhD co-coordinator for their guidance and moral support during my research work and all faculties of School of Computing Science & Engineering who helped me a lot in my course of research work and special thanks to all those who stood behind me.

I would like to express my heartiest love and respect to my friends Ms. Priya Porwal, Ms. Gunjan, Ms. Jayati Mukherjee and Dr. Yogesh.

Last but not the least I would like to express my deep sense of gratitude to all my family members, relatives and friends for their blessings and moral

support during this research work. There are times in this work when the clock beats you time and again you run out of energy and you just want to finish it once and forever, in such situations my parents and husband supported me well. The real motivation, confidence and mental support is achieved from the blessing of parents. I want to give the credit to my husband Mr. Abhishek Pandey whose immense support was always there for me in my odds and evens during this research work.

Once again, I want to mention Dr. Dileep Kumar Yadav, I owe more than what I can mention for teaching and guiding me to see the silver line in every dark cloud.

SNEHA MISHRA

ABSTRACT

In real-time environment, the detection of motion-based object is detected against challenging issues *i.e.* cluttered background. Practically, the desirable results are hard to obtain due to challenging issues such as dynamic background, illumination variation, floating or spouting water etc. The identification, detection, tracking and recognition is a crucial area of research in the area of computer vision that plays smart decision-based roles for intelligent video surveillance system, intelligent transportation, indoor-outdoor, virtual reality, medical diagnosis, robot vision navigation, law and enforcement, security, border monitoring, and many more. With the advent of technological advancements and availability of enormous data in the form of audios, videos, text and other forms. Artificial Intelligence emerged as one of the most significant technologies that simulates human intelligence through computers and several algorithms. Computer vision is one of the prime domains that enables to derive meaningful and crisp information from digital media. It includes several sub domains such as facial recognition, pattern detection, image classification, object detection and many others. Many vision-based applications such as traffic controlling, action recognition, human behavior analysis, industrial inspection, an intelligent surveillance are important issues for the researchers and industrialists in the modern days.

This work presents the study of object detection in videos or continuous sequence of frames captured by the static camera. The thesis will make the clear image of comparison of object detection methods performed in prior researches. The main focus is on the real time application areas of video surveillance and the concerning challenges in the respective areas in near future. All the background modeling techniques and their sub categories studied and implemented by various authors is mentioned in literature. It also depicts the major challenging issues available in real-time environments.

The significant aim of the proposed work is to develop an adaptive method to compute the threshold during run-time and update it adaptively for each pixel in testing phase. It classifies motion-oriented pixels from the scene for moving object using background subtraction and enhanced using post processing. The main objective is to focus and study the problematic video data captured through camera sensor, to handle challenging issues available in real-time video scene and to develop a background subtraction method and update the background model adaptively for moving object detection.

Presently, the vehicle detection and tracking in intelligent transportation system is a highly active area. The proposed work also resolves the above problems and delivers solutions for the enhancement of the transportation system and the automobile industry. So, the work investigates a method using Google's firebase platform (as a cloud service) and a background subtraction method for moving vehicle detection and tracking in a foggy environment. Here, the Google's Firebox storage (Pyrebase API) is used for the storage of video that provides authentication and storage services. The moving vehicle is detected using the background subtraction method for considered video followed by post-processing.

Detection and correctly tracking the moving objects in a video streaming are still a challenging problem. Due to the high density of vehicles, it is difficult to identify the correct objects on the roads. In this work, we have used a YOLO.v5 (You Only Look Once) algorithm to identify the different objects on road, such as trucks, cars, trams, and vans which results that the proposed approach attains improved results as compared to state-of-the-art approaches which is another contribution of the thesis.

Table of Contents

List of Figures

List of Tables

ACRONYMS

SVM	Support Vector Machine
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
R-CNN	Region-Based Convolutional Neural Network
YOLO	You Only Look Once
MOD	Moving Object Detection
GMM	Gaussian Mixture Model
BGS	Background Subtraction
KLD	Kullback- Leiber Divergence
FMOG	Fuzzy Mixture of Gaussian
ML/AI	Machine Learning/ Artificial Intelligence
CCTV	Closed Circuit Television
IP	Internet Protocol
IoT	Internet of Things
DSLR	Digital Single- Lens Reflex
SFO	Stationary Foreground Objects
ABODA	Arizona Band and Orchestra Directors Association
RGB	Red-Green-Blue
KITTY	Karlsruhe Institute of Technology & Toyota Technological Institute
UA-DETRAC	University at Albany Detection and Tracking
CSP	Cross Stage Partial Network
SPP	Spatial Pyramid Pooling
XML	Extensible Markup Language
GPU	Graphics Processing Unit

1 INTRODUCTION

This chapter paraphrases the concepts of the terminologies used throughout the thesis. The conceptual meaning of image processing, computer vision, video tracking, object detection and video surveillance are well expressed. The chapter also describes the real time application areas of object detection and the challenges faced by the cameras to capture the clear image of the object. Later in the chapter the motivation and the objective of this work are explained followed by the organization of the thesis.

1.1 Computer Vision

Digital camera is the term used to capture an image or record the sequence of frames, a video. Object detection or tracking is the term which comes in mind after digital camera. The branch of computer science that comes in place of human vision is named as computer vision which obviously is related to some knowledgeable data extraction from an image or video. The two contradictory terms image processing and computer vision are sometimes considered as the same or rather interchangeable concepts.

Let's consider these terms one by one, they are at times used in a very common way. Image processing focus on processing the images by certain transformations like smoothening of image, increasing brightness or contrast of the image and detecting or highlighting edges of an image. Image processing take input as an image and reverts output as an image only. On the other hand, computer vision is used to retrieve information or features from an image. Computer vision takes input as an image but the output is the information regarding the image which is obviously taken out with the help of image processing techniques. Thus, computer vision is dependent on the image processing techniques. Figure 1 describes the difference between image processing and computer vision.

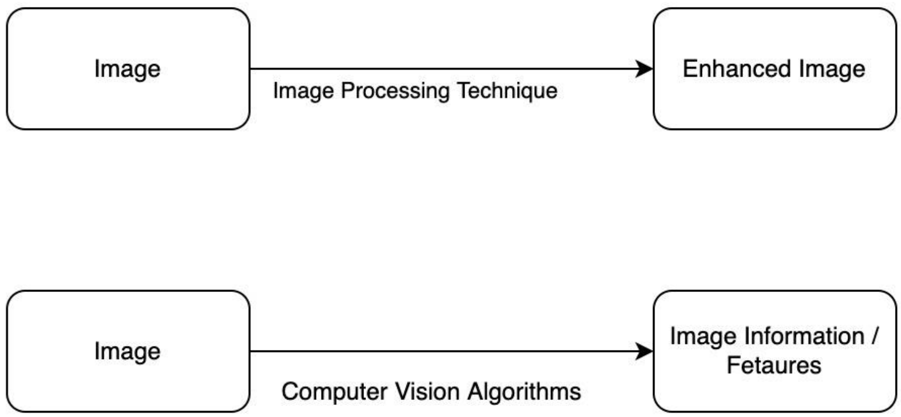


Figure 1.1 - Image processing and computer vision

In many real-time scenarios the computer is taking a dig on human vision, there are areas where human eyes fail to work. Computer Science from few decades is trying to give visual system to machines. Since 1860 when the first photographic camera was invented now it’s almost 200 years and the research is to make a machine work with better accuracy than that of the human. Computer vision is increasingly taking complex challenges by including machine learning, artificial intelligence, deep learning to analyze more complex images with better accuracy.

1.2 Object Detection

Object detection is precised as computer vision and image processing technical approach which allows detecting, locating as well as identifying the objects in an image or video. Every object has its own feature, so classification of the object by analyzing the features of the object is also the part of object detection. The type of the object whether it is a human, car, building, animal, etc., the count of the objects is located in the scene comes under features of the objects. Several methods of object detection are performed by n-number of researchers. Basic background modeling, fuzzy background modeling, statistical background modeling, background estimation are the types of background modeling. Object detection methods are also based on artificial intelligence and deep learning, but background modeling is the most simple and efficient method used to detect the objects. In machine learning the first necessity is to define the features using the techniques like support vector machine (SVM) for classification, while deep learning is based on the techniques like CNN (convolutional neural network), RNN, Faster R-CNN, Retina-net, YOLO are able to so end to end object detection without defining features that much specifically.

1.3 Object Tracking

Traffic surveillance system is the most important aspect of video tracking. Background modeling and foreground object detection plays important role in video surveillance system. Static camera or moving camera captures the videos, inherent changes like waving trees; unconditional weather, water surfaces, etc. may vary as background is not completely stationary. Optical flow algorithms calculate and analysis of each pixel motion within two image sequences or frames. Edge detection detects the regions or edges by edge detector algorithms to detect and classify vehicles in traffic. Figure 2 describes the object tracking method by a flowchart, the video sequences are passed then the machine asks whether it is an initial image or not if the yes the it is a background image otherwise the background subtraction techniques are applied that further do foreground segmentation, then if the object is detected it is tracked but if the object is not detected then again it comes to background modeling or creating the background image.

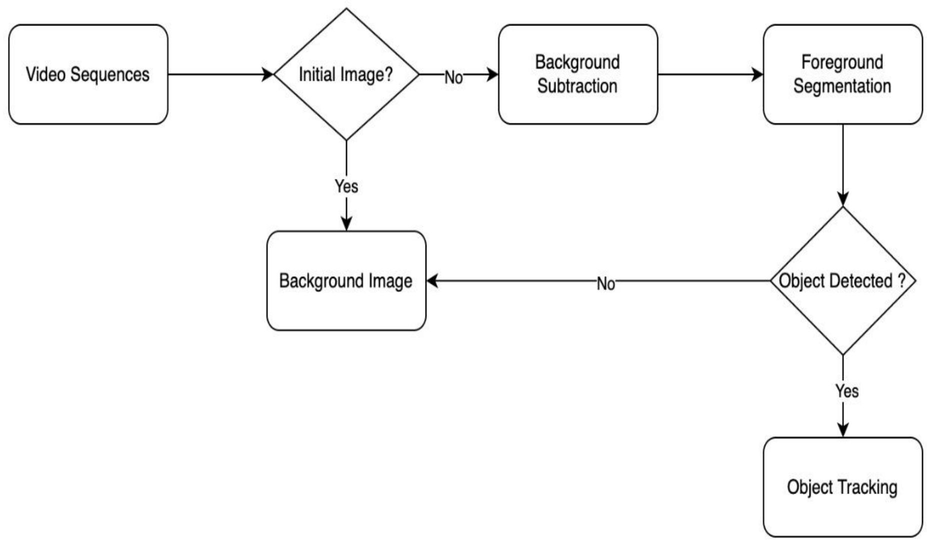


Figure 1.2 Flowchart showing object tracking process

1.4 Application Areas

Today computer vision is playing a revolutionary role in many industries like transport, medical sciences, agriculture, security, retail, banking, etc. Various real time applications that require the object detection and tracking as the part of their work in everyday life are been described as under. The detailed explanations of several real time systems are covered in this section. It can be applied in various applications like:

1.4.1 Indoor outdoor surveillance

For the state of security of any market or industry the examining team is there, that is sitting in one control room and can visualize as many activities as many cameras are connected.

Figure 1.3 The man sitting in control room for surveillance of indoor and outdoor activities

1.4.2 Target detection

Target detection is the very essential task in video surveillance. Target detection technique is followed by object detection and tracking of moving objects. This is to analyze the motion of the moving objects efficiently.

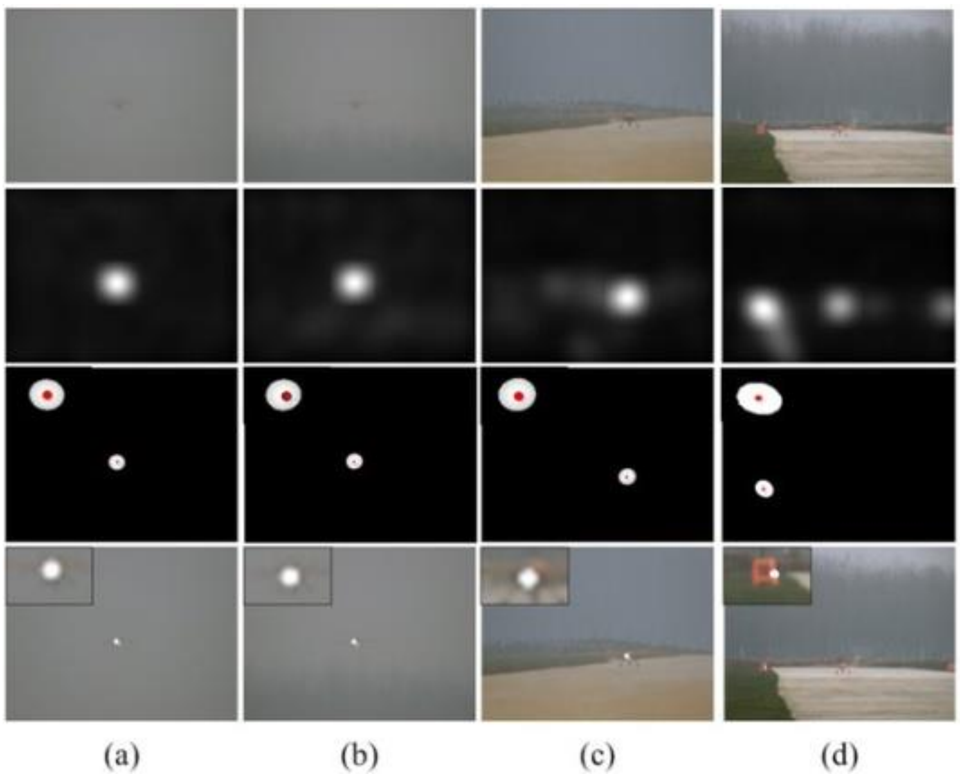


Figure 1.4 Target detection

1.4.3 Traffic monitoring and analysis

Traffic surveillance is the most attracted application area by computer vision. Traffic monitoring needs to detect the vehicles and further to classify them. Traffic analysis gives results like speeding vehicles, gap detection, stopped vehicles detection, detection of vehicles coming from wrong direction.

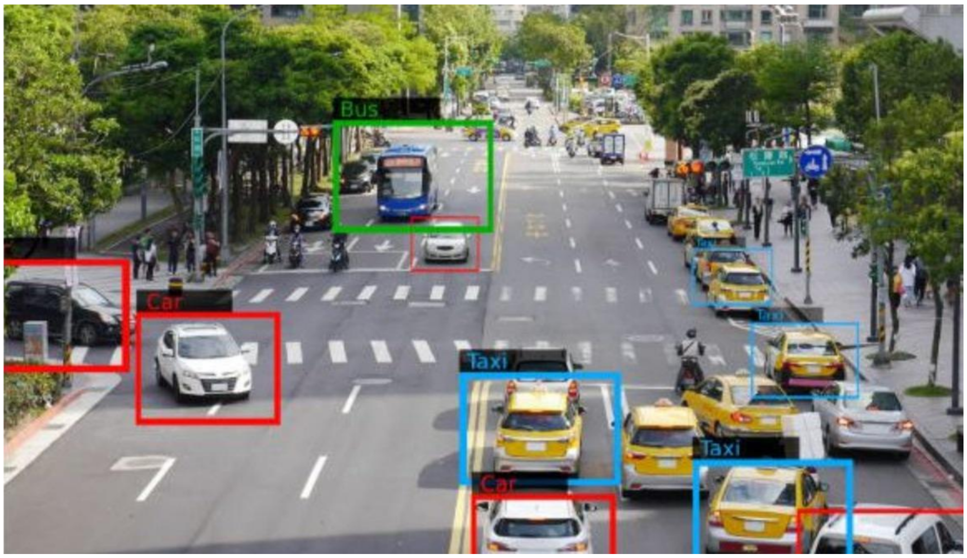


Figure 1.5 The traffic surveillance

1.4.4 Abandoned object detection

From past few decades abandoned object detection has become the wide are of research. In video surveillance the scenario of moving object identifies the stationary objects which are called the abandoned object. This application is very active and desirable in public places like hospitals, railway stations and airports. These are the places at potential risks so the monitoring of abandoned objects is at high priority in such places.

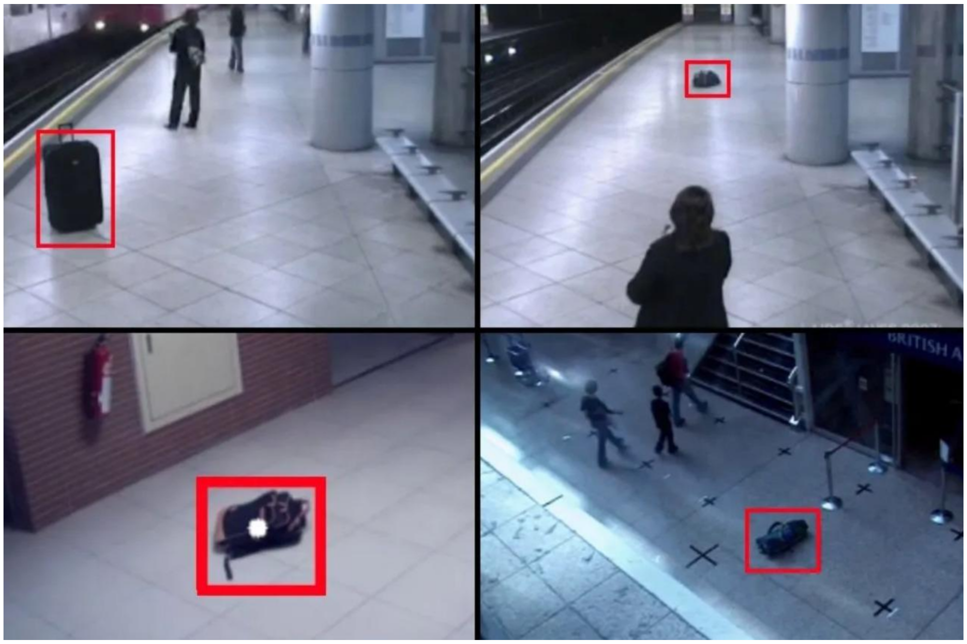


Figure 1.6 The detection of abandoned object

1.4.5 Logo or packaging detection

In packaging industries, the monitoring of the various objects which are been packed is required. Like the packaging of any medicine tablet whether there is no packaging defect in the packet. This area needs high accuracy and efficient algorithms of object detection and tracking to maintain the product quality, because it is the direct and effective carrier for any brand. In industries like cosmetics and food packaging, this is the first impression for the brand.

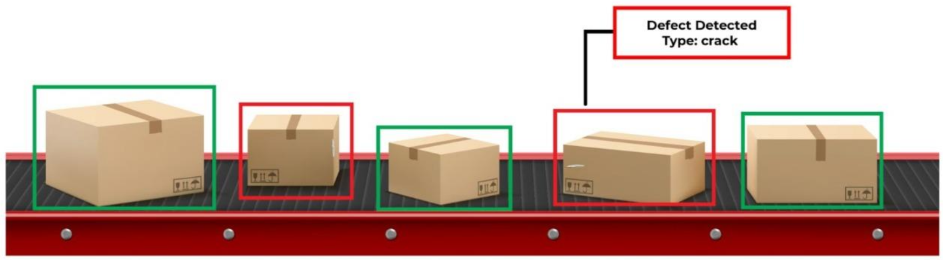


Figure 1.7 The detection of abandoned object

1.4.6 Defense

Military is the most updated version of technology. Every country serves the best technology to its army, so is the computer vision best enabled in this sector. This is used to detect the enemy troops; it helps in target detection and tracking to shoot or take any action or just to spy on enemy. The all-time surveillance is required with the high quality devices and the most advanced object detection and tracking algorithms.

Figure 1.8 The troops and the arms(guns) are detected.

1.4.7 Activity Analysis

Computer vision serves activity recognition as an important application in many areas like healthcare and video surveillance. It is the recognition of activities by regular observations on some particular actions. As example, in a football match the behavioral detection is the part of activity analysis whether the player has hit the other player by mistake or knowingly.

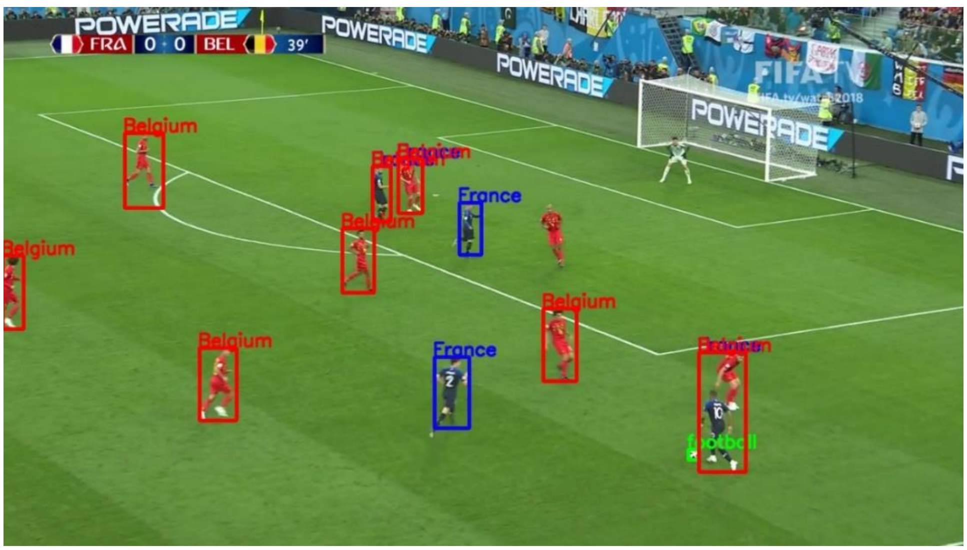


Figure 1.9 The tracking of the activity by the different teams in the football match

1.4.8 Thermal Imaging

This gives the ability to detect in miscellaneous background. It provides ability to visualization in dark environment. Leakage detection in gas pipeline or underwater leakage and leakages in electric wire are also the application of Thermal imaging. Figure below depicts usage of thermal imaging to detect the temperature of human to identify if it is normal or abnormal.

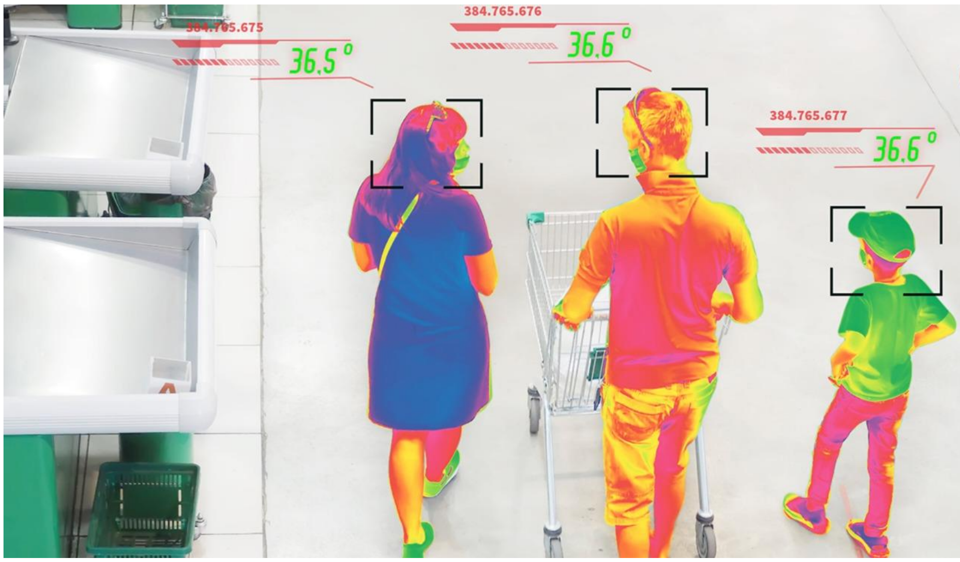


Figure 1.10 Thermal Imaging

This is not the end of the list there are plenty more applications of object detection. In airports, railway station, bus stands and shopping complex, the suspicious activities are tracked by object detection. And thus, object detection plays a crucial role in the real-life scenarios. But there are many challenges to be resolved for detecting the object in adverse conditions which are listed and briefed in the next portion of the chapter.

1.5 Challenges

Video surveillance is the active research area but there are lots of difficulties a capturing device face whether background is static or moving. Challenges may be related to bad weather conditions, light effects, bad quality of camera. Challenging difficulties come in the way to detect the objects in any image or video sequence. The crucial challenging issues in video for surveillance are given as:

1.5.1 Bootstrapping

If background frames are not available then it is very difficult to detect the object because the very first frame is the background. By applying background modeling, if there is no frame available as a background then it is difficult for the algorithms to detect the original object. Thus, it is very big challenge to model a background in such cases.

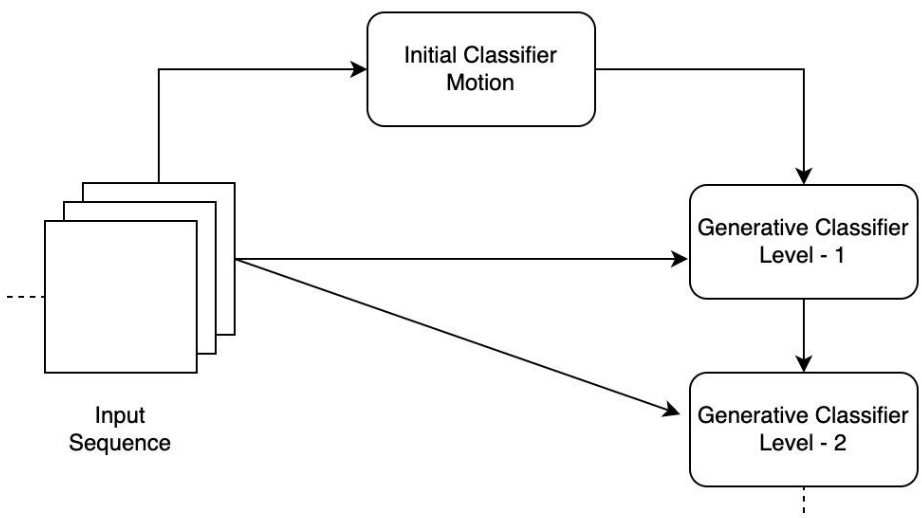


Figure 1.11 Bootstrapping Methodology

1.5.2 Camouflage

Camouflage is a condition when two objects of the same color are being overlapping on each other in the frames and making it difficult to detect the actual object. As an example, in agriculture industry the detection of insects and warms is very difficult to detect as the color of the leaf or flower on which the insect is sitting is of the same color. This makes difficult to detect the insect as an object. Many authors have worked in the area to reduce this by learning representation of moving objects but it's still a big challenge in the area of object detection and tracking.

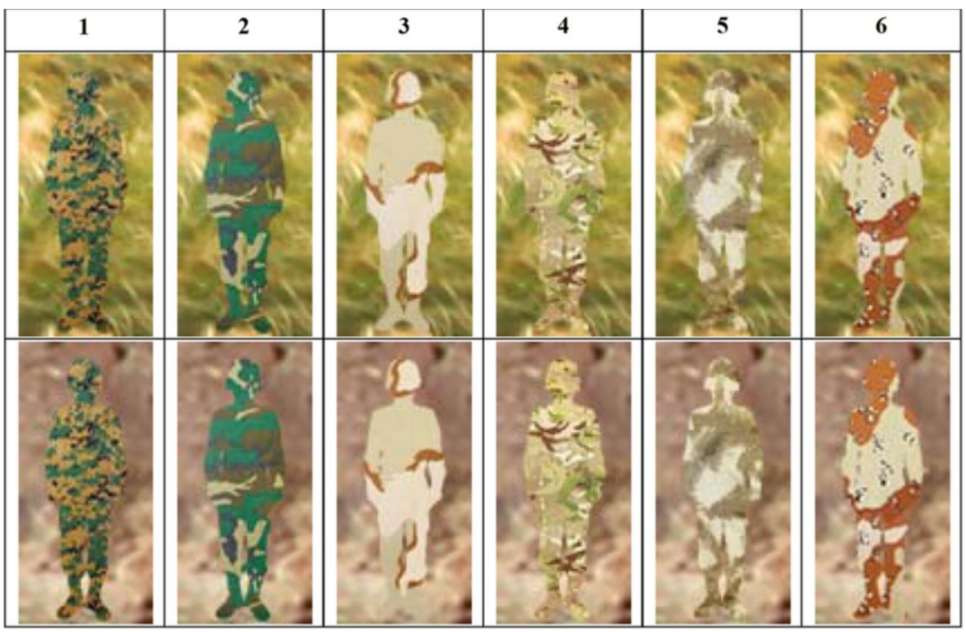


Figure 1.12 Different types of camouflage images.

1.5.3 Illumination variation

At different times of days, the lightening effects may vary due to light sources, reflection from different sources of light, blockage of light from other objects are very challenging difficulties. This lightening impacts the object's appearances which results in false positive detections. Due to these reasons lightening consequences are challenging to be handled. Many authors as discussed in section 2 have tried various approaches in such areas to lessen the challenges of lighting variations.

Figure 1.13 Different times of days have different lighting effects where detection is a challenge

1.5.4 Foreground aperture

In this condition the possibility of false negative increases a lot as foreground region is having the similar colored region and changes inside the foreground region caused by various lightening effects and different issues of motion.

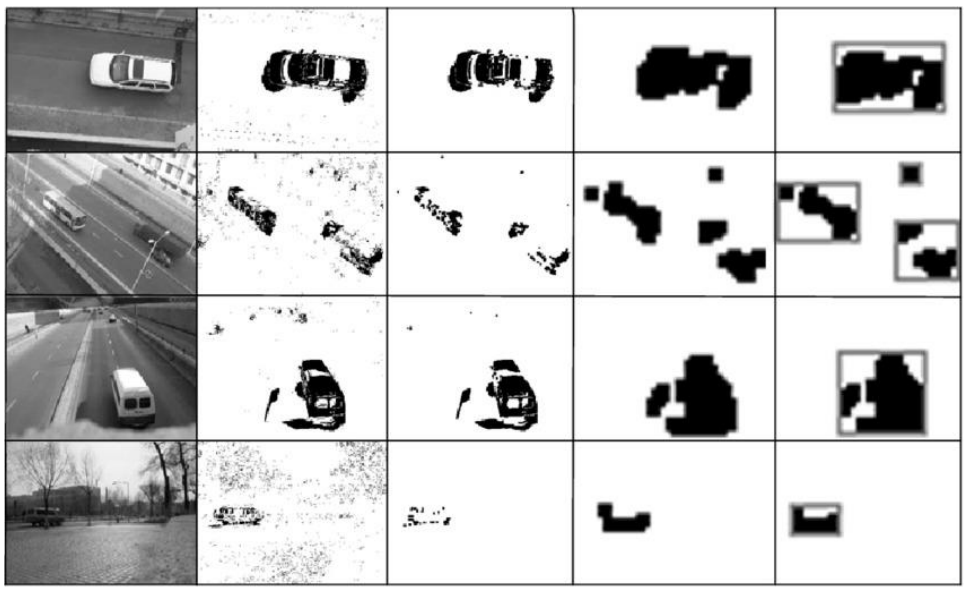


Figure 1.14 Frames with foreground aperture and motion in the background

1.5.5 Motion in background

Background may be of dynamic nature in this condition. There are sequences of frames in which the background is changing that makes object detection and tracking a challenge. Sometimes the swaying trees, sprouting of water from the fountain or rain, moving flags, etc. detects the false positive pixels.

1.5.6 Occlusion

Occlusion is the condition in which an object's some parts are occluded or hidden behind the other object. Occlusion may be of two types- partial occlusion or complete occlusion. If the objects are hidden partially then the term is partial occlusion, while if the object is hidden completely then it is termed as complete occlusion. Let's take an example of a pedestrian walking in the street is occluded with the trees besides the street or by the cars passing away from the street. Thus, occlusion creates hindrance in detecting the actual object, thus makes detection and tracking both a challenging difficulty.

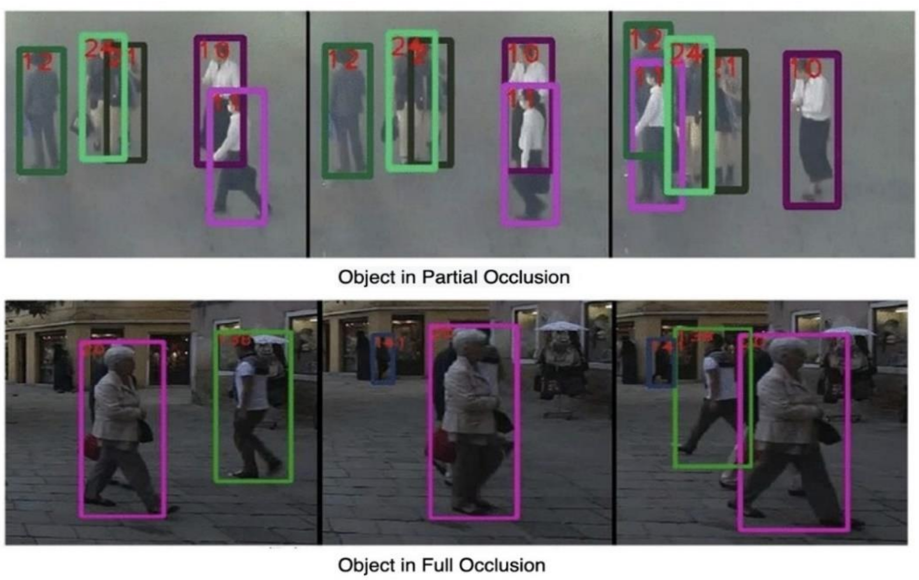


Figure 1.15 The partial and fully occluded images are shown.

1.6 Motivation and Objective

The object detection and tracking in video is a challenging problem and has been extensively investigated in the last decade. It has various applications in numerous fields, such as video compression, video surveillance, human-computer interaction, video indexing, and retrieval etc. The object detected is accompanied with misclassified foreground objects due to illumination variation or motion in background. So, an additional work carried out in this problem will be, to propose a simple but efficient object detection technique, which is invariant in change in illumination or motion in background.

The main motivation evoked from the problematic challenges and requirement of real-time based application are:

- Increasing needs of society for security-surveillance applications.
- Detection and tracking of moving object from problematic video data.
- Real-time based feature extraction from video data for visual surveillance system.

This work contributes to draw a comparison of various techniques of background modeling, their advantages and limitations. The study of existing research work gave idea of various steps used in background subtraction. Background subtraction has four steps – background modeling, threshold generation, background subtraction and background maintenances. The threshold generation may be fixed and dynamic. Background maintenance is threshold maintenance as per the dynamic conditions generated in the scene. Several application areas in the real time which keenly substitute computer vision object detection technique as the key role instead of human are the major attention of the paper. The challenges that come in the way of object detection are well described in this thesis.

Objectives:

- To investigate a method for moving object detection in **dynamic environment / illumination variation**.
- To develop a method for moving object detection under **foggy / snow environment**.
- To improve the existing method for **detection and tracking** of moving object for **feature extraction**

1.7 Organization of the thesis

In this work, chapter 1 explores the introduction of the considered research area in various domains and challenges in the real time environment in videos as well as static cameras. The chapter 2 discusses the background of the area and the various object detection methods. The currently available technologies used in object detection and the data set publicly available with the Indian Universities and the Laboratories which are actively working in object detection. Chapter 3 focuses on the literature review it also provides a comparative study of considered state- of- the- art method. Chapter 4 investigates a method for moving object detection in dynamic conditions such as illumination variation. Chapter 5 proposed the detection of moving vehicles in foggy environment using Google's Firebase. Chapter 6 introduces a new method for object detection and tracking of moving vehicles in the dense traffic area for smart surveillance and chapter 7 concludes the thesis with the future scope followed by references.

1.8 Summary

The chapter summarizes the description of the various parts of computer vision related to the area of research. The revolutionary roles of computer vision are identified and discussed in the chapter. The various applications of object detection in the real time environment are described with suitable examples. Based on the applications of the object detection, the challenges are also enlisted in the chapter.

2 BACKGROUND

This chapter focuses on the various techniques used for moving object detection. Video surveillance, object detection and object tracking are precisely introduced in the first part of the chapter followed by detailed explanation of object detections techniques. Phase correlation, frame differencing, background subtraction, optical flow and probabilistic approach are the major techniques used for object detection and are appropriately elaborated in this chapter of the thesis. Later part of the chapter describes the data set openly available in various institutes and the laboratories for object detection.

2.1 Video surveillance, object detection, object tracking

The major problem in computer vision is moving object detection (MOD) in video surveillance, monitoring, activity recognition for security purposes. So, this is to monitor the physical movement of a particular region. The methods of moving object detection are background subtraction, frame differencing, optical flow and temporal differencing. This particularly separates the moving object from the static component that is called background and the object is considered as foreground. MOD is an elementary as well as an essential step of object detection in computer vision. This paper mainly focuses on the background subtraction approach.

Various *state-of-the-art* MOD methods have been proposed and they also work well in the static camera. But very less work has been carried over in the video sequences. In the current work, we have tried to contribute to overcome the challenges in static as well as moving cameras.

2.2 Object detection techniques

There are various object detection methods or techniques available for different applications. This section explores about the efficient techniques used for object detection in video. The techniques phase correlation, frame differencing, background subtraction, optical flow, block matching and the probabilistic approach are mathematically described below.

2.2.1 Phase Correlation

Phase correlation approach is to enumerate the relative translative motion between the even (similar) images. This approach is estimated by fast Fourier transformation. Mathematical representation of phase correlation approach is described as follows:

Suppose the two input images are, i_a and i_b. If, we apply hamming window function on i_a and i_b, we will get the reduced edge effects. Now, calculate the 2D-fourier transform for i_a and i_b,

$$I_a = F\{i_a\}\ ,\ I_b = F\{i_b\}$$

Now, take the complex conjugate to calculate the cross-power spectrum and then multiply the Fourier transformation element-wise.

$$\mathcal{R} = \frac{I_a \circ I_b^*}{|I_a \circ I_b^*|}$$

Here, ○is the Hamdard product.

Now, the cross-power spectrum for index (j, k) is:

$$\mathcal{R}_{jk} = \frac{I_{a,jk} \circ I_{b,jk}^{*}}{|I_{a,jk} \circ I_{b,jk}^{*}|}$$

Then, we apply inverse Fourier transformation to obtain the cross correlation shown in the following equation:

$$\mathfrak{r} = F^{-1}\{\mathcal{R}\}$$

Location of the peak is determined in terms of 'r'

$$(\Delta x, \Delta y) = argmax_{xy}\{\mathfrak{r}\}$$

This is how the phase correlation methods work. So, on the basis of this we find some advantages and disadvantages of phase correlation approach which are briefed here with.

Advantages of phase correlation approach

Every approach is to overcome the less efficient approach, so it has benefits over other methods. The following are the benefits of phase correlation approach:

- This method is best suitable in occlusion.
- This is used to detect the objects in medical imaging.
- It is more resilient to the noise.

Disadvantages of phase correlation approach

Phase correlation technique does have following imitations:

- Since the consecutive image will be the simple linear shift of the first image, Δr function will not be simple delta function which may reduce the performance of the method.
- It may give ambiguous results in the output for periodic dataset.

2.2.2 Frame differencing method

The most widely used method of object detection is frame differencing method it is the type of background subtraction technique. This can only be performed if the previous frames are available, since buffers require previous frames values. The current frame is subtracted from the past available frames. After comparison with the threshold value the background and the foreground frame are identified.

Mathematical representation of frame differencing method:

For pixel P x and y, the intensity is denoted by I at time t.

The frame difference for the frames at time t and t-1 are represented as:

$$F.D = |I_{x,y,t} - I_{x,y,t-1}|$$

Where F.D is the frame difference, now the value of frame difference is compared with the value of threshold T as determined as follows:

$$P_{x,y} = \begin{Bmatrix} foreground & if\ F.D > T, \\ background & Else \end{Bmatrix}$$

The value of T is the choice of threshold value, which is may be selected by our own parameters and sometimes on hit and trial basis.

2.2.3 Background subtraction

Basic Background Subtraction method is the successor of the frame differencing method. The background modeling may or may not be required in different scenarios of the datasets.

The Background subtraction is very popular and efficient technique.

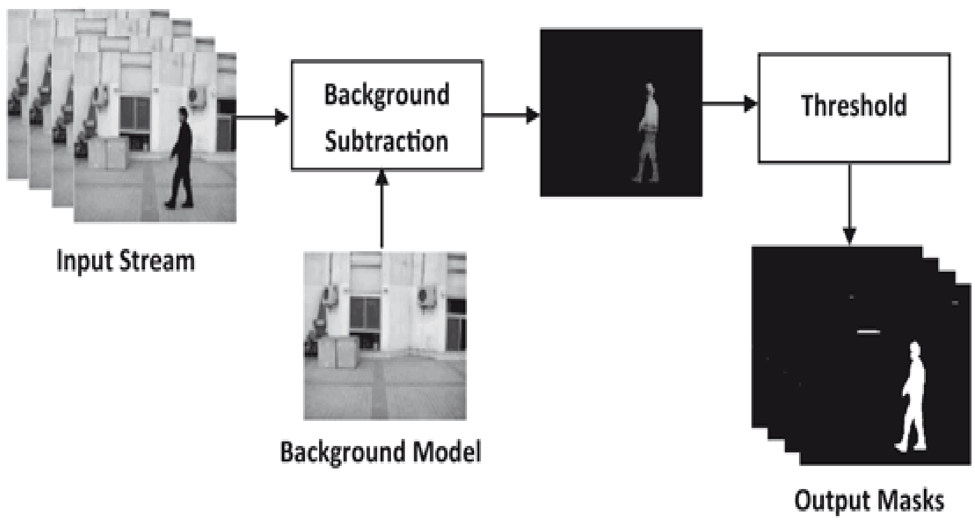


Figure 2.1 Basic background subtraction model

Background subtraction Mathematical approach

- Background modeling
  - ➢ Construction of background frame
  - ➢ B = AVG of initial 100 frames
- Threshold generation
  - ➢ Constant or fixed or global threshold
  - ➢ Dynamic threshold or adaptive threshold
- Background subtraction
  - ➢ Compute difference frame

- Difference (x, y) =abs (C - B)
- Background maintenances
- Pixel Classification
- if (Difference (x, y)>=Threshold
- {

Pixel (x, y) = 1   i.e. moving pixel

}

else

{

Pixel (x, y) = 0   i.e. stationary pixel

}

Advantages of background subtraction method:

- Simple to implement.
- Have very low complexity.
- Very high computational efficiency.

Disadvantages of background subtraction method:

- The foreground aperture is major challenging issue.
- The threshold value is global for all pixels of the frame.

2.2.4 Optical flow techniques

Optical flow is the approach for moving object detection. This for moving objects in an image with stating as well as when the camera is also in motion, such as in drones.

The mathematical representation of the optical flow methods is:

We assume that Intensity I, is constant, in the corresponding frames.

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t)$$

The intensity is taken for two consecutive time frames t and t + δt.

By using Taylor Series and avoiding higher order terms, the above equation reduces to:

$$\frac{dI}{dx} V_x + \frac{dI}{dy} V_x + \frac{dI}{dt} = 0$$

$$I_x . V_x + I_y . V_y = -I_t$$

Here, V_x and V_y are the optical flow vectors at the coordinate (x, y, t)

The value of motion vector is represented as:

$$T = \sqrt{V_x^2 + V_y^2}$$

Benefits and limitations of the optical flow approach:

The jitter in the camera is a major challenge for other approaches of object detection but optical flow algorithms go well with this. Optical flow is very complex for computations and is not costs effective approach.

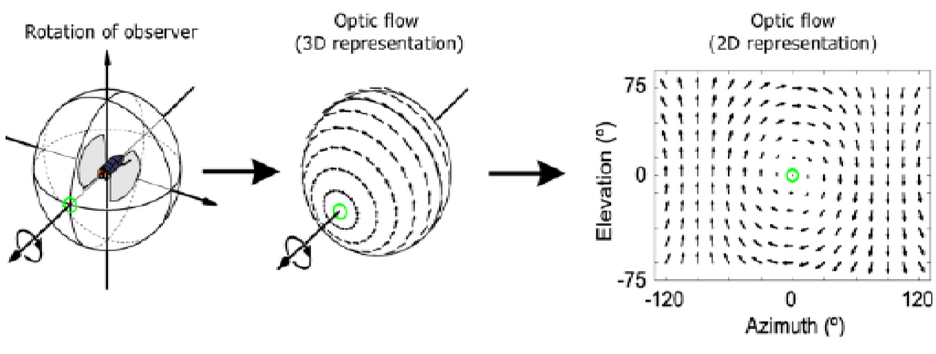


Figure 2.2 Optical flow illustration

2.2.5 Probabilistic approach

This approach accurately quantifies the two types of uncertainties spatial as well as semantic. The probabilistic approach is the extension of conventional object detection approaches.

It requires a detector for the following:

- A category wise distribution of all class labels.
- A bounding box which is based on mean and covariance to describe the top-left and the bottom right corners of the bounding box. Representation of bounding box is:

  $\beta = (N_0 , N_1) = (N(\mu_0 , \Sigma_0), N(\mu_1 , \Sigma_1)$

  Where μ_i is mean for multivariate Gaussians and Σ_i is the covariance. Probability distribution 'P' for the pixels (u,v) is represented as:

$$P(u', v') = \iint_{0,0}^{v',u'} \mathcal{N}_0(u, v)\, du\, dv \iint_{v',u'}^{H,W} \mathcal{N}_1(u, v)\, du\, dv,$$

Here, H is the height and W is the width of the image.

2.3 Publicly available data sets

When we talk about dataset, the availability and accessibility matter a lot. The following datasets are openly available datasets developed by research institutes and the laboratories for the dedicated work on object detection. The table defines the parameters of the dataset like the sequences, frame size, frame type and the availability of required number of background frame.

Table 2-1 List of Openly available Data-set

Dataset	**Sequences**	**Frame size**	**Frame type**	**Availability of required no. of background frame**
Change Detection [Source: University of Sherbrooke, Canada; Changedetection.net]	Library, Dining Room, park, corridor	320 X 240	8-bit gray-scale	Yes
CSIR – CSIO: India [Source: http://vcipl-okstate.org/pbvs/bench/] Dr. Guoliang Fan at Oklahoma State University.	Hot Ambassador Near, Innova	640 X 480	8-bit gray-scale	No
OSU – Pedestrian [Source: Ohio State University, USA; http://vcipl-okstate.org/pbvs/bench/]	OSU-Padestrian_004	360 X 240	8-bit gray-scale	No
Houston-Zoo [Source: OSU-Pedestrian Dataset: http://www.vcipl.okstate.edu/otcbvs/bench/]	HoustonZoo_Rino_l5b1	626 X 360	8-bit gray-scale	No
Microsoft Wall Flower Dataset	Bootstrap, Camouflage, waving trees, time of day	360 X 240	8-bit gray-scale	Yes

2.4 Summary

This section of the work is all about the currently available approaches used in the object detection. The data set been discussed above is the openly available dataset been used by various researchers in their benchmark contributions.

3 LITERATURE REVIEW

The chapter paraphrases the literature review of the existing work been in progress till now in the area of object detection. The chapter comprises of the brief introduction of the area followed by the application areas the challenges and the role of Artificial Intelligence, Machine learning and Deep learning in the area of computer vision. This portion of the work also defines the methods followed by numerous researchers, their advantages, disadvantages and the improvements done by them.

3.1 Introduction

The gaining of high-level understanding from the images or videos is computer vision. It can automate the tasks which the human eyes can do. Somehow, it is always not possible for human visuals to reach at every place to take a track of the work, like in deep mining activities and border security across the border, in any unfavorable climatic condition or unsound geographical areas. For such places video surveillance is required to be performed. Object detection is the key application in computer vision. The object detection may further lead to many more applications of computer vision like object tracking, object recognition, activity detection or behavioral analysis. In real time scenarios, the captured video faces the following problems.

- Climatic changes such as the weather may be rainy, mist, dusty, foggy or sunny.

- Illumination variation in which light may be dark or too bright.

- Background motion which may be due to swaying trees, sprinkling of water from fountain or any other moving sources.

- Noise due to camera hardware.

Nowadays, manufacturing industries and agricultural department are also using video surveillance as there may be many human-made mistakes like inattentiveness, double vision or blurred eyes. Detection, identification, motion and simultaneous tracking of multiple objects or targets in coastal, navy, army, underwater and indoor- outdoor video surveillance system is developing effectively in today's era. Various application areas of video surveillance may be manufacturing industry, sea surveillance, restricted areas, and leakage detections are tried to be described in *Table3.1*.

Table 3-1 Real time applications with description

Application areas	**Description**
Agriculture	To detect the type and size of the material or objects in agriculture is very important for their further classification, so as to detect the diseases or classification of breeds.
Thermal imaging	This gives the ability to detect in miscellaneous background. It provides ability to visualization in dark environment. Leakage detection in gas pipeline or underwater leakage and leakages in electric wire are also the application of Thermal imaging.
Security	In detecting the object which may be a crime suspect, the behavioral tracking of the suspect can give an alarm. The person climbing the fencing can be detected.

Manufacturing industry	The sorting and assembling in manufacturing industry is a very time taking task by human and again carries abnormalities and errors, object detection and tracking can ease the manufacturing world with less or no abnormalities.
Packaging industry	Packaging industry may include medicine packing and many retail objects packing, video surveillance may locate the object with no filling or partial filing of the packet. Human visuals may not be successful detectors in this industry.
Mining	Mining may be very deep under the earth where the human worker or expert is not possibly safe to go inside the deep penetrations for mining of coal or any other minerals, surveillance through camera is the best way to detect the object as much far it is.
Sea surveillance	Sea surveillance by visual cameras is to protect the ship or overseas borders, so the navy is alarmed before any danger.
Restricted zone	Restricted zones may be some deep zones, or dense forest zones where wild animals or terror activities are performed may be tracked by surveillance cameras.
Transportation	Traffic surveillance plays the major role in controlling the accidents and over speeding track of the vehicles. If the camera is mounted on the vehicle then it is much easy for the driver to get alarms as soon as the object, pedestrian or any vehicle is near to the vehicle to save it from accident.
Army	Video surveillance in military purposes can reduce the offences like crimes in the nation, border security is major job of army that can be better controlled by border surveillance by thermal imaging or infrared imaging.

3.2 MOTIVATION

In this chapter the motivation is dynamic background as this requires a change in consecutive frames due to moving objects. Thus, it needs to upgrade the background model every time and then detect the foreground. Even in indoor scenes where the environment is controlled, the shadow or light effects may be undesired. There are various real time application areas of video surveillance like packaging industry, indoor-outdoor surveillance, abandoned object detection, mining industry, manufacturing industry, agriculture and many more.

Object detection in videos requires robust system rather it is simple to detect an object when the background is static. There are plenty of approaches by which object detection may be performed like optical flow, background subtraction, filtering techniques. The background subtraction is the most used technique and so is to be concerned in this paper also. The main goal is to detect the changes between consecutive frames. The idea behind this concept is that no prior information is required to classify the state of a pixel as a foreground or a background object. So, the main objective is to resolve these challenging issues and detect the moving or stationary pixels of the scene.

3.3 LITERATURE WORK

In earlier decades many researchers have applied non-adaptive methods of back grounding which have lots of drawbacks. These methods were useful only in highly supervised and short-term applications without changes in the video scene. It required manual re-initialization when error occurred. Adaptive back ground averages the image over time. This is working in the situations of moving objects (foreground).

Stauffer and *Grimson* introduced Gaussian Mixture Model [1] . His work attempts to model the value of all the pixels as one type of distribution as a mixture of Gaussian. According to the persistence and variance of each, they identified the background and foreground colors. *Lee* developed the statistical framework for GMM techniques which increased the convergence speed of learning mechanism without compromising model stability [2].

Haque further proposed a technique to model each pixel of the scene independently by a mixture of k Gaussian distributions [3]. *Lavanya* worked effectively on the BGS based scheme for critical background. The extraction of foreground pixel was based on Fisher's ratio-based threshold [4].

Yadav has improved the *GMM* by introducing post processing [5] [6]. *Yadav* further improvised for moving object detection in thermal video frames by using statistical parameter based *BGS* method by using different threshold Kullback-Leiber divergence-based method. This work was divided in three respective stages – (i) Training phase (background modeling constructed by trimmed mean based simple averaging method), (ii) Testing phase (generating *KLD* based threshold values, differentiating background reference frame and the current frame, classifying the pixels), (iii) Enhancement phase (applying morphological filters and image processing tools to improve detection quality) [7]. Similarly, Sharma *et.*

al. (2016) [4] has investigated a new concept for moving object detection under thermal environment where static camera has been used for capturing of video.

Allili introduced a new improvement by using Bayesian based estimation and applied infinite Gaussian mixtures [1]

Thierry Bouwman then worked on a new approach fuzzy mixture of Gaussian (*FMOG*) whose parameters were based on fuzzy c-mean algorithm. This focuses on précised estimation of parameters [8]. *Thierry Bouwman* in his survey proposed the statistical method for background modeling is the most used one; he classified the statistical background modeling methods according to three generations. This survey also classified the improvements as intrinsic improvements and extrinsic improvements [3].

Zeng et. al. [9] introduces a *Type-2* fuzzy set in *MOG* which is called *T2-MOG* for uncertainty [10]. Several authors gave their contribution in different background modeling methods.

Steps involved in Background subtraction techniques are described as-

i. Background modeling

ii. Threshold generation

- Fixed (constant)threshold
- Dynamic or adaptive threshold

iii. Background subtraction

iv. Background maintenances (updation)

- Background maintenance

- Threshold updation or adaptive threshold generation or maintenance of threshold as per dynamic condition generated in the scene.

The background subtraction is the process of separation of foreground (moving object) from static information. For background subtraction, background modeling is the first step. *Table* 3.2 categorizes the various background modeling methods with their respective sub-categories and respective authors who contributed in the related researches to provide an improved or innovative idea in background modeling. According to the study every background modeling method have some drawbacks and advantages over other method which are described in *Table* 3.3 to compare the methods. There are numerous advantages and disadvantages of any method but a few and needful are summarized here. This may somehow contribute to the understanding of the existing methods for various background modelling techniques.

Table 3-2 Literature Review of background modelling

Types of background modeling	**Sub categories of background modeling methods**	**Authors**
Basic background modeling	Mean Median Histogram Pixel intensity classification Pixel change classification	Bouwmans [8] Mac Farlane et al. [11] Zheng et al. [12]
Statistical background modeling	Single Gaussian Mixture of Gaussian Single general Gaussian Mixture of general 44Gaussian Kernel Density Estimation Support Vector Machine Subspace learning using PCA Subspace learning using ICA Subspace learning using INMF	Wren et al. [13] Kim et al. [10] Allili et al. [1] Elgammal et al [14] Lin et al [15] Tsai and Lai [16]
Fuzzy background modeling	Fuzzy mixture of Gaussian (FMOG) Type-2 Fuzzy Gaussian mixture model	T. Bouwman [12]
Background estimation	Wiener filter Kalman filter Chebychev filter	Toyama et al. [17] Messelodi et. al. [18] Chang et. al. [19]

Table 3-3 Benefits and Limitations of different background modelling

Type of background modeling	**Advantages**	**Limitations**
Basic background modeling	Easiest method to perform. It performs well in static background.	It does not work in dynamic Background.
Statistical background modeling	Most used to their robustness to the critical situation. Gives good performance.	Presents a tricky parameter optimization problem.
Fuzzy background modeling	FMOG is more robust than mixture of Gaussian.	Works only for dynamic backgrounds (waving Tree).
Background estimation	It has less trouble with frequently changing light (illumination variation), handles better than mixture of Gaussian. Gives more accuracy but fewer computations.	Probably not suitable for most applications. Very less computation is carried.

In computer vision detection of moving object is the first step to be done. On basis of studying a lot of literature there are too many investigating fields contributing a big amount of publications. In *Table* 3.4 these many investigating fields with their respective authors responsible for the representation of their publications are indicated. This will eventually help the readers to carry forward any of the approach of background modeling that works for most robust system and real time application.

Table 3-4 Overview of different techniques of object detection

Techniques	**Contributions**	**Improvements**
Fuzzy logic based	Haque, Murshed, & Paul [3]	-Suitable for infrared videos. -More robust in case of dynamic backgrounds such as waving tree.
Gaussian based	Lucas & Kanade, 1981 [20] Kim, Sakamoto, Kitahara, Toriyama, & Kogure, 2008 [10] Allili et al. [1] Medioni et al. [21] Zhao et al. [22]	-Better performance in presence of gradual illumination changes and shadows. -Intrinsic and extrinsic improvements.
Kernel density estimation based	Elgammal & Davis, 2000 [14] Pahalawatta et al. [23] Orten et al [24] Tanaka et al. [25]	-Deals with multimodal backgrounds (water rippling). -Many works adopted to reduce computation time.
Optical flow based	Sotirios et al. [26] Lucas et al. [20]	-Great work in object identification with high accuracy. -Optical flow analysis may be performed on the series of images to determine a description of motion.
Subspace learning	Oliver et al. [27] Tsai and Lai [16] Bucak et al. [28]	-improvement in presence of multimodal background. -An intrinsic improvement in presence of shadows.
Deep learning	Krishevesky et al. [29] Simonyan et al. [30]	-Background subtraction and convolution neural network. -Combined for anomaly detection in much abnormal object's surveillance.

3.4 ROLE OF MACHINE LEARNING, ARTIFICIAL INTELLIGENCE AND DEEP LEARNING IN COMPUTER VISION

Over a decade background subtraction has been very effectively and frequently used technique for object detection. As discussed in above sections many background modeling methods are and were in existence to detect the objects for further applications like object tracking, classification, extraction or recognition. With the rapid development around artificial intelligence, machine learning and deep learning computers are been trained to build background models to detect the moving objects. *Table* 3.5 describes various techniques and applications of ML, Deep learning and AI in computer vision.

Table 3-5 Methodology and Application oriented study

Approach	Methodology & Application	Details
Machine Learning in Computer Vision	Support Vector Machines	SVM classifier is able to classify the points of the input set into two possible classes.
		SVM is a non-parametric method. Therefore, it has the flexibility to represent complex functions while at the same time being resilient to over fitting.
	Neural networks	Neural networks are organized in layers where each unit receives input from only the nodes of the next higher layer.
		It is possible to represent any continuous function with arbitrary accuracy
		Result is a trained network with adjusted weights capable of processing an input and producing output with computed data class.
	Deep learning	Deep Learning Networks are more complex means of layer interconnection and require strong computational power for training.
		Four main architectures types are: - Unsupervised Pretrained Networks - Convolutional Neural Networks - Recurrent Neural Network - Recursive Neural Network.

	Applications	Machine learning helps computers to understand what they see. It can b used in wide variety of areas like below: - Agriculture (detecting diseases in crops based on visual symptom) - Visual Character Recognition / Generating subtitles for videos - Surveillance etc.
Deep Learning in Computer Vision	Convolutional Neural Networks.	CNN comprises three main types of neural layers, (i) convolutional layers, (ii) pooling layers, and (iii) fully connected layers.
		Every layer of a CNN transforms the input volume to an output volum of neuron activation, eventually resulting in a mapping of the input dat to a 1D feature vector.
		CNNs have been extremely successful in computer vision applications, such as face recognition, object detection, powering vision in robotics, and self-driving cars.
	Deep Belief Networks and Deep Boltzmann Machines.	It's a probabilistic generative model which provide a joint probability distribution over observable data and labels.
		DBNs are graphical models which learn to extract a deep hierarchical representation of the training data.
		DBNs are appropriate for unsupervised learning since they can be trained on unlabeled data.
	Applications	1) Object Detection: Approaches following the Regions with CNN paradigm usually have good detection accuracies. 2) Face Recognition: Google's FaceNet and Facebook's Deep- Face are both based on CNNs. 3) Action and Activity Recognition: In the authors successfully employ a CNN-based approach for activity recognition in beach volleyball 4) Human Pose Estimation
Artificial Intelligence in Computer Vision	Fuzzy logic	It is an extension of conventional set theory that deals with the concept of partial truth. It aims to model the vagueness and ambiguity in complex systems.
		To perform image processing using fuzzy logic, three stages must occu - image fuzzification: To modify the values of a specific data set or image - fuzzy clustering: Used to modify the membership values once image data are transformed from fuzzification - defuzzification: Decoding of the results
	Genetic algorithm	A genetic algorithm operates in three stages: - Build and maintain a population of solutions to a problem - Choose the better solutions for recombination with each other - Use their offspring to replace poorer solutions.
		Typically use the mechanisms of crossover and mutation and work towards obtaining global optimum.
	Applications	Can be used in vide variety of fields like Natural Sciences, Biological Sciences, Industry, Management and Engineering etc.

3.5 CHALLENGING ISSUES

Moving object detection from video sequences is not an easy task. There are lots of difficulties or challenges faced by the camera to capture the sequences. *Thierry bouwman* in his survey presented all the challenges in detail. Some of the challenges faced by the camera are briefed as follows:

- *Illumination Variation*: Lightening situations of the target scene might always not accordingly. The appearance of the object changes due to motion of light sources like weather conditions – sunny, foggy, rainy, mist, windy in outdoor scenes, on and off of light in indoor scenes, reflection from sun or any other bright source.

- *Occlusion*: The target might get occluded by some other objects. Occlusion may be partial or complete, if partial part is occluded or hidden then it is partial occlusion but if full object is hidden then it is called complete occlusion. The pedestrian may be occluded by trees. This is a challenge to detect the object if it is occluded by the background either partially or completely.

- *Complex background*: In outdoor scenes nature may be variable every now and then such as waving trees, sprinkles of water from fountain, movement of clouds is highly variable so these movements are non-periodic and not in control of camera and thus challenging for background modeling.

- *Shadows*: Occurrence of shadow is part of illumination variation; it may be of different size at different times of the day. This is again a complicated task to detect object.

- *Camouflage*: The camouflage is somehow related to partial occlusion but camouflage foreground is subsumed by the background. This causes difficulty to categorize background and foreground.

- *Bootstrapping*: This is a scenario when background frames are unavailable this becomes critical situation to model the background.

- *Foreground Aperture*: This is the condition where foreground region has the uniform color region as that of the background; this increases the possibility of false negative.

3.6 Summary

The background modeling is the basic and most important task in background subtraction technique for object detection. The chapter covers all the advantages and limitations of background modeling methods studied in the literature. Based on these limitations some challenging issues are described in the area od object detection for smart surveillance system.

4 INTELLIGENT TECHNIQUE FOR MOVING OBJECT DETECTION FROM PROBLEMATIC VIDEO CAPTURED THROUGH CAMERA SENSOR

According to the huge demand of surveillance system, the society is towards an intelligent video surveillance system that detect and track the moving objects (person, vehicle *or* animal) using surveillance cameras (CCTV, IP or Spy *etc.*). So, this chapter has developed an adaptive background subtraction method to handle such challenging problems. The chapter focuses the problematic video data captured through camera sensors to handle challenging issues available in real-time video scene. This chapter demonstrates better performance of the proposed method as compare to state-of-the-art methods.

4.1 Introduction

In real-time environment, the detection of motion-based object is detected against challenging issues i.e. cluttered background. Practically, the desirable results are hard to obtain due to challenging issues such as dynamic background, illumination variation, floating or spouting water etc. Various researchers, scientists focus on challenging issue of cluttered background such as waving trees, spouting or rippling water etc. [33], [34]. This chapter work developed a statistical feature-based background subtraction method for moving object.

The identification, detection, tracking and recognition is a crucial area of research in the area of computer vision that plays smart decision based roles for intelligent video surveillance system, intelligent transportation, indoor-outdoor, virtual reality, medical diagnosis, robot vision navigation, law and enforcement, security, border monitoring, and many more [35] [36]. Now a days, various kinds of cameras are available which work in different-2 degrees of movement and autonomy but in reality, various kinds of problems are raised in real time situation such as motion in background, blur cases of motion, illumination variation related issues, occlusion, change of real world 3D information into 2D video frames, hard ware noise and shadow etc. [37] [38] [39]. In most of the outdoor-indoor environment, the major issues are occurred due to light variation, shadow, and motion in background [40] [41] [42]. So, the appearance of moving object diverts and variation effect on precision of outcome. Thus, developing a delight method for motion detection against above challenging issues is very important.

In literature various kind of background subtraction and other method are available. This research work also aims to focus on the background subtraction based solution to cope-up with such challenges [43] [44] [45] [46] . This work also aims to experiment over publicly available datasets which are captured through fixed (static) camera. In this work the background subtraction method focuses on the motion information available in the video frame [47] [48] [49] [50] [51] .

Bigdata in Video based surveillance system:

The proliferation of multimedia devices generates an unprecedented large amount of video data over the Internet of Things (IoT) using camera sensors. The choice of appropriate sensor helps to achieve better quality videos. Consequently, now-a-days, the world has entered into the era of video based Bigdata. Currently, due to the enhancements in recent cloud and distributed computing-based technologies, several researchers, computer scientists are focusing on the video based Bigdata analytics specially in the cloud. The video based Bigdata analytics community and industries are looking towards an effective framework using current technology as per market trends. However, this research work is too limited and providing a brief survey of recent research work on video big data analytics in the cloud and analysis of a large amount of video data, main opportunities, major challenges, and promising research directions [52] [51]. One promising direction for delivering solutions on video big data veracity are various techniques and research methods which are also capable enough to assess the credibility of video data sources so that filtering can be applied over untrustworthy video data. Another way is to is to develop novel ML models which effectively inferences with defective captured video data [52]. Main task is to capture video frames using suitable image sensor inbuilt with CCTV/IP cameras, action cameras, smart phones, mirror less systems, DSLRs and purpose-built video cameras. They all shoot video in completely different ways and some are much better such as DSLRs, IP and CCTV. In case of security surveillance system, a camera continuously captures video sequences and can be considered as BigData that has high volume, high velocity, and variety of scenes.

Still, numerous factors affect video big data analytics performance in the cloud, and moreover experimentation of such a comprehensive evaluation for all considered use cases are almost near impossible in terms of accuracy, storage, speed, and resources etc that effects performance [53] [54] [55]

[56]. The video sequences captured are stored in storage devices or continuously processed for analytical purpose [57] [58] [59] [60] . Moreover, this work focuses developing a solution for moving object detection and tracking over problematic datasets which are available publicly for academic and research purpose [50] [51] [60].

This work focuses on the data captured through camera sensors (Image, Video, Signal etc) and handles the real-time challenges faced by various electronics, computer vision and multimedia industries who are working for developing solutions for surveillance and security. Because, currently everyone is looking for safety, security, monitoring of activities/events through video surveillance. It is used by almost all reputed events (in-house or among countries), industries, indoor-outdoor, border, traffic, transportation, school, colleges, banks, stations, underwater, coal/gold mining etc. So, in brief, this is highly recommended for industrial purpose due to huge demand of society. Now-a-days, cloud is also used to store data stream from sensory devices and we have to do the needful analytics over such data using cloud environment too.

The significant contribution of this work is to develop an adaptive method to compute the threshold during run-time and update it adaptively for each pixel in testing phase. Its aim is to classify motion-oriented pixels from the scene to moving object using background subtraction and image processing techniques. The post processing steps enhance the detection quality of the focused outcomes.

4.2 Literature review

Stauffer *et.al*. [40] has proposed a GMM based concept for background and foreground separation in real time. Here an adaptive mixture model was developed using Gaussian distribution for background pixel classification. This work contains two valuable parameters (i) Learning rate (infinity), (ii) Fixed Threshold T i.e. proportion of data accounted for background. This method deals with repetitive motion and lighting variation in the background. This paper work depicts a robust method for moving object detection.

Mahfuzul *et.al*. [41] has improved the GMM based method investigated by Stauffer *et. al*. In this work, author has improved the threshold for pixel classification and its dependency on learning rate. These methods are more robust and demonstrate better qualitative and quantitative results against its peer methods. It also deals dynamic and lighting variation issues.

Ka Ki Ng *et.al*. [42] proposed a simple background subtraction method using simple mean-based background model & background subtraction. This method performs when there is no illumination variation in background. It also performs and generates better results.

Zhou *et.al.* [61] has developed a DECOLOR method to deal with illumination, motion of background. This method was experimented over various datasets and compared with numerous methods and demonstrates better performance. This method also depicts its limitation along with hardware feasibility and possibility in the practical real time environment.

Rai *et. al.* [46] has developed a method for moving object detection in thermal video frames using Kullback-Leibler divergence. This work [62] demonstrates better performance.

Kajo *et. al*. [47] has focused on few challenging and problematic tasks such as stopped car detection, abandoned luggage detection and stationary foreground objects (SFOs) for detection and classification in real time. This research work decomposes the video tensor spatiotemporally and divides it into foreground and background components using singular value decomposition.

Park *et. al*. [48]) has proposed an efficient method that distinguished stolen objects, abandoned objects, and ghost regions in video scene. This work has been experimented over own dataset focusing on the challenging issues and is more suitable for open environments such as exhibition halls and public parks for delivering security services. Amin *et. al.* [63] has investigated a novel object association technique using convolution neural network and also developed a framework for abandoned object localization along with the identification of the owner. This method was experimented over publicly available datasets and their own newly developed dataset and demonstrates exceptionally well performance.

Kim *et. al.* [63] has suggested a deep learning-based detection method for outdoor environments and is robust to illumination variations and ghost effects.

Liu *et. al.* [9] has developed a YOLO based deep learning model that is experimented over ABODA database for detecting abandoned objects and is more robust to illumination variation available in the scene.

4.3 Proposed Work

In this chapter, the proposed method has been accomplished using the following sections:

- Background model construction
- Automatic generation of threshold
- Background subtraction
- Maintenance of background model
- Post-processing

This section explored the working procedure of the proposed work as mentioned through the following steps depicted in the following flow diagram shown in Figure 4.1.

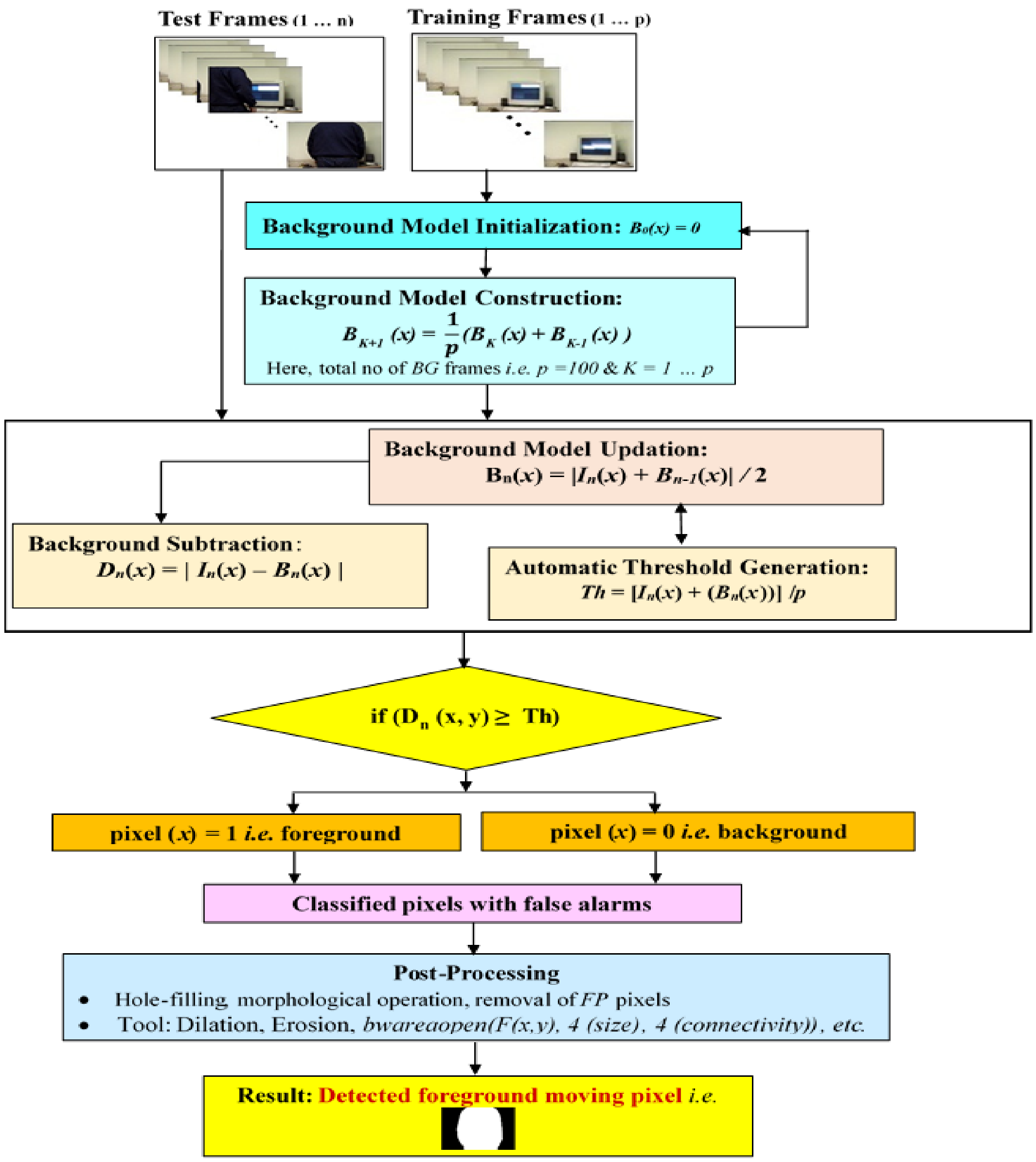


Figure 4.1 Flow diagram depicting the work flow of proposed model

4.3.1 Background Model Construction

The initial step of this paper work is to construct a reference frame or background model. During modeling of background frame is experimented over initial few frames i.e. k = 100 frames. The proposed model is constructed using initial k frames through evaluating an average of considered k frames. So, the constructed background model is developed using the following equation.

$$B_{k+1}(x) = 1/p * (B_k(x) + B_{k-1}(x)) \quad \ldots (1)$$

Here, Bk+1(x) is the constructed reference frame or background model.

4.3.2 Automatic Generation of Threshold

The main task is to compute an appropriate threshold automatically during testing stage at run time for pixel classification. The automatically computed threshold helps for classification of stationary as well as moving pixels from a scene.

$$\text{Th} = [I_n(x) + (B_n(x))] / p \quad (2)$$

- Here, the current frame is represented by $I_n(x)$ and constructed background model is depicted using $BM_n(x)$), these are helpful for evaluating the threshold automatically without external support.

- *Background Subtraction*

In this step, a difference frame (i.e. Dn(x)) is computed by subtracting the test frame or current frame (In(x)) and background frame (Bn(x)). Here, this work focuses on the absolute difference of every pixel of both frames as mentioned in the following equation (3).

$$D_n(x) = |I_n(x) – B_n(x)| \quad .... (3)$$

- Adaptive Maintenance of Background Model

Here, the background model is updated for each pixel of every frame during each iteration during execution process. The maintenance of model cause minor fluctuation, dynamic situations, gradual illumination variations, dust, mist, cloudy, foggy, haze related challenging issues are minimized and can be handled.

$$B_n(x) = |I_n(x) + B_{n-1}(x)| \,/\, 2 \quad ..... (4)$$

So, this threshold helps us to classify the motion-oriented pixels more correctly if there is change in environment, illumination, of dynamic challenge occurred after some time and keep it updated adaptively for every test frame.

4.3.3 Post Processing

The post-processing step is applied to enhance the detection quality after classifying pixels. It minimizes the false classification and improves the accuracy of the classification model. The main steps involved for post-processing are given below.

- Compute labeling, connected component and remove outliers.
- Erosion with structure element of disk type and default connectivity.
- Hole filling operation.
- Dilation operation.
- Erosion operation.

4.4 Experimental Setup: Materials and Methods

In this work, the video data is captured through camera sensors. This data has been developed by Microsoft and change detection community for problematic scenario such as numerous environmental effects, illumination variation, dynamic nature, fog, thermal, rain, etc. The experimental results and analysis of proposed work processed over Microsoft's Wallflower dataset (camouflage, waving trees and foreground aperture sequences) and one Library sequence of change detection dataset. These communities provided such problematic video sequences are freely available for public use for specially for non-commercial purpose. The frame size is 320×240 and available in gray format and the experimental work is analyzed over

Matlab 2014b platform and Windows 8.1 operating system. The proposed work is configured with above mentioned software and is executed over hardware of Intel (R) core-2 duo processor having speed of 2.33 GHz and 8 GB RAM. The results analysis is processed in two steps namely: (i) Qualitative analysis (visual results, Figure 4.2), (ii) Quantitative analysis (numeric values, Table-4.1) and the idea of graphical abstract inherited from Kumar *et. al.*

4.5 Qualitative Analysis

This section discussed about the qualitative analysis of proposed work and its comparison with considered state-of-the-art methods over freely available Microsoft's Wallflower Dataset [64], Change Detection Dataset [65]. This work is experimented over both colored and thermal video datasets and which are converted into gray-scale format. The qualitative analysis and their results are shown in the Figure 4.2, where library (results of thermal data) and other color video frame sequences are considered here. Such work is also applicable in various exciting applications of colored as well as thermal scenarios such as video based smart surveillance, aerial thermography, Traffic analysis, healthcare, restricted zones, Traffic analysis, healthcare, inspections of mechanical components, gas leakage detection, night vision applications, etc. The visual results of this work and considered peer methods are available in Figure 4.3.

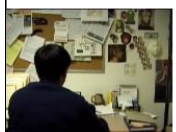
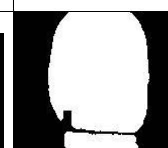
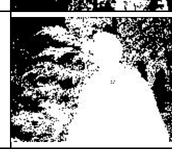
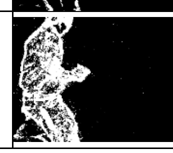
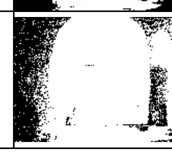
Method/Data set	Foreground Aperture	Waving Tree	Thermal Library	Camouflage
Original Frame				
Ground Truth				
Proposed				
Zhou *et. al.* [61]				
Ka Ki Ng *et. al.* [42]				
Mahfuzul *et. al.* [3]				
Stauffer *et. al.* [40]				

Figure 4.2 Qualitative analysis for object detection (row-wise): (a) Original Image (b) Ground Truth (c) Proposed (d) Zhao et (e) KaKiNg (f) Mahfuzul et. (g) Stauffer et

4.6 Quantitative Analysis

In this section, the quantitative results are evaluated by various parameters and seen in the Table-4.1. The quantitative values of this work and considered peer methods are compared and analyzed which are also depicted in Table-4.1. The quantitative values evaluate the precision, recall and F1-score. The average values of precision, recall and F1-score are computed through the following metrics.

Table 4-1Precision, Recall and F_1-score: Quantitative analysis

Datasets	Methods	Precision	Recall	F1 - score
Foreground Aperture	Proposed	0.9475	0.8432	0.8923
	Stauffer [40]	0.3101	0.6986	0.4296
	Mahfuzul [41]	0.8366	0.4893	0.6174
	KaKiNg [42]	0.7793	0.4456	0.5670
	Zhou [61]	0.5050	0.4271	0.4628
Waving Trees	Proposed	0.9324	0.9836	0.9573
	Stauffer [40]	0.6648	0.9989	0.7983
	Mahfuzul [41]	0.8557	0.6318	0.7269
	KaKiNg [42]	0.9741	0.3672	0.5334
	Zhou [61]	0.9150	0.9934	0.9526
Thermal: Library	Proposed	0.9894	0.9968	0.9930
	Stauffer [40]	0.8781	0.7392	0.8027
	Mahfuzul [41]	0.9075	0.0371	0.0713
	KaKiNg [42]	0.9810	1.0000	0.9902
	Zhou [61]	0.9711	0.9920	0.9814
Camouflage	Proposed	0.9261	0.9607	0.9431
	Stauffer [40]	0.7032	0.9934	0.8234
	Mahfuzul [41]	0.9671	0.7727	0.8590
	KaKiNg [42]	1.0000	0.8164	0.8989
	Zhou [61]	0.5638	0.6809	0.6168

The F1-score metric is evaluated the performance of proposed method and compared with some of the chosen state-of-the-art methods. The precision versus recall based comparison is analyzed shown in Table-4.1 and analyzing these metrics (i.e. precision, recall and F1-score) with the proposed work. According to the quantitative values obtained (Table-4.1: Precision, Recall and F1-score), this work demonstrates better performance of proposed method as compare to state-of- the-art methods where Zhao *et. al*., Ka Ki Ng *et. al*., Mahfuzul *et. al.* and Stauffer *et. al*. also shows better results; these methods achieves lots of miss classification which indicates poor performance. These metrics focus on the classification of pixels which are classified correctly in each frame. Here, experimentally computed average value of these parameters depicts better detection quality and the proposed method demonstrates better performance. The precision-recall curve is shown in Fig. 4.3. that depicts the classification of moving pixels clearly. The proposed method is depicted with the black line curve. Here, the pixels are classified more accurately using the proposed work but Ka Ki Ng *et. al*. and Zhao *et. al*. also achieved good performance but the overall value is lagging.

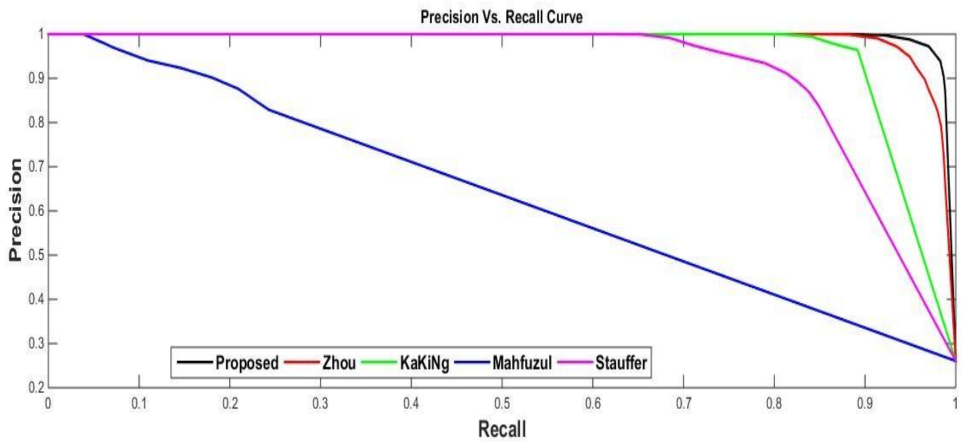


Figure 4.3 Precision-Recall curve

4.7 Results and Discussion

The main contributions and outcomes of this work, discussed using above metrics are mentioned through comparative analysis as depicted below.

- The qualitative analysis of the proposed method is depicted in Figure 4.2.
- The proposed work shows its better performance in terms of precision-recall and F1-score that shows maximum value and can be seen through precision-recall curve (Fig.4.3) also. It is also compared with considered state-of -the-art methods (Figure 4.2, Table-4.1).
- In case of Foreground Aperture dataset as shown in Table-4.1, Mahfuzul *et. al* [41] delivers better performance but its performance is poor as compare to the proposed method.

- In case of Waving Tree dataset as shown in Table-4.1, Stauffer *et. al.* [40] generates better outcomes but its results are lagging as shown in Table-4.1 as compare to the proposed method.
- In case of Time of Day dataset as given in Table-4.1, Zhao *et. al* [61] method is performed better but it gives is poor results against the proposed method.
- Similarly, according to quantitative results given in Table-4.1, in case of Camouflage dataset, Ka Ki Ng *et. al.* [42] method depicted better results but its results are lagging as compare to the proposed method.
- The study of qualitative and quantitative evaluation demonstrated that the results obtained through the proposed method over problematic sequences shown better results and also performs better as compare to considered peer methods.

4.8 Summary

The chapter focuses on the moving object detection in video scenes, this work is helpful for visual surveillance applications. Here, an adaptive background maintenance approach is used that update the background model adaptively and handles various challenging issues. The quantitative evaluation depicts better average values through adaptive method at run-time. The precision-recall curve also shows the better performance and demonstrated maximum average value of F_1-score as compare to considered peer methods.

5 DETECTION OF MOVING VEHICLES IN FOGGY ENVIRONMENT USING GOOGLE'S FIREBASE

This chapter is organized as the introduction of cloud environment, the overview of moving object detection and related literature work. The remaining part of this chapter discusses the main techniques and methodologies of the proposed object detection method using Google's firebase. The experimental results are presented with the proposed task boundaries.

5.1.1 Introduction

In today's real world, technology is enhancing day- by- day for computer vision applications in the field of video surveillance for vehicle detection and tracking. For avoiding real-time adverse situations such as traffic congestion, pile-ups, accidents due to ill management of transportation system, the detection and tracking based methods must be improved. For video surveillance in intelligent transportation system, the traffic cameras are installed on chosen positions. The captured information is passed to the control centre of traffic management system for further analysis and investigations, for such date Pan Tilt Zoom, Gasto or Vector cameras etc may be used for traffic surveillance or detection of motion- oriented vehicles on road or highway. Apart from fog, in reality there may be other challenging problems such as – illumination variation, snow, dust, mist, cloudy environment, drift or occlusion etc. This work mainly focuses on the foggy environment. The main purpose of analyzing computer vision is to detect, track and estimate the features of the object directly with the help of computers and then develop a system that reduces human effort. With the enhancement of technology, society is adopting digital video-based surveillance [64] [65] [66] [67]. The detecting and tracking of target are also a very important application area. Mobile target detection technology distinguishes moving objects and backgrounds and extracts moving targets from video. Such work can be used for transportation, industry and security surveillance [68] [69] [70] .

The detection of moving objects is constantly very difficult in foggy environment. The object detection-based research has offered a variety of traditional methods, including frame differentiation, optical flow and the background extraction to detect moving objects and are recently got a lot of attention. According to the literature, few works has been done by researchers for detection and tracking under foggy environment and still a very tough task to identify, detect and track the moving vehicle on highway or road. So, the proposed work develops a method for such work. The main focus is to experiment such work under Google's firebase cloud platform. It reduces the number of spirits in removing the critical gaps and background in differences between the frames and effectively detects moving objects in the presence of critical background. However, in complicated dynamic situations the detection is not accurate as lots of false positives are classified in most cases. Various pixels may be misclassified due to illumination variations, environmental or weather effects of the background scene.

This work suggests a new direction to detection of moving vehicles. In this work, the background subtraction method is applied over videos stored in Firebase Storage, and the static backgrounds shows the following new attributes: 1) In the proposed method, frames are considered as a sequential order 2) bilateral blurring removes noise as a pre-processing 3) morphological techniques as a post processing are included in the process; 4) Contour detection, that allows to avail more accurate information.

The main motivation behind this work is problems associated with adverse weather while driving. While the entire car body is fogged up, the windshield and windows are most visible and may create dangerous driving situations. In the thick fog, the driver is unable to see beyond the limits of his own vehicle. Driving in such conditions may be very problematic and also cause accidents or pile-up, affects perceived judgments about speed and distance. The effect is a result of reduced contrast. The rain reduces driver awareness and changes visibility through changes to the headlights, windshield, road itself, and road signs. It also affects traffic flow by reducing road capacity, forcing drivers to reduce their driving speed, slowing down travel time, and increasing the overall risk of accidents [71] [72] [20] [73]. Hence, the proposed method works optimally and detects objects on the screen even in an inclement or adverse weather, especially in fog.

5.1.2 Related Work

There are great approaches in the literature for detecting and tracking moving objects in a video stream. Motion-based information plays a very important role in the recognition process. Various approaches and algorithms have been developed on object detection in the literature. Some of them are discussed in this chapter.

R. Li *et. al*. [66] has proposed method for the moving object in the video that has been detected by the cloud server. It also shows the computing ability of the suggested cloud platform and enables utilization of the available resources. It also avoids having to rely on mobile devices for the limited processing capacity.

S. Tuli *et. al.* [68] suggested a framework to deploy deep learning-based applications to deliver better service quality in fog-cloud environments to harness edge and cloud resources. Authors also developed a framework, (*i.e.*EdgeLens) that adapts to user requirements and depicts high accuracy through experimental analysis or low latency modes of services.

Gruyer *et. al.* [69] focuses on the real-time natural issue such as fog, snow, haze, rain, or sun glare and these are very dangerous for drivers. For visibility problem, the driver faces problems while driving. This paper develops a method to improve the visibility in adverse weather conditions.

Rahul Singh *et. al* [70] has investigated a method for detection of moving vehicle in foggy environment. This work also applies a LiDAR to measure the distance from front vehicle and raise a warning message according to the measured distance.

Yadav *et. al.* [74] has developed a Kullack-Leibler Divergence based method for moving object detection in thermal environment and demonstrated better performance against considered peer methods.

Lochabh *et. al.* [75] has given a theoretical application-oriented concept for object detection in video in cloud Environment. This paper work also focuses on the real-time based problematic challenges.

S. Yadav *et. al.* [76] also developed a method for enhancing services in transportation system. This work focused to automate the Indian transportation system through intelligent searching and retrieving mechanism in the Amazon's elastic compute service.

Zhang *et.al*. (2021) [77] has developed a Moving- confidence- assisted Matrix decomposition (MCMD) model that has focused on foreground regularization and configurated to assist real-time moving objects with the moving confidence score evaluated from optical flow methods. This work developed batch processing and online solutions over videos by SkySat and Jilin-1.

Chaojun *et. al*. [78] has proposed a vehicle model detection to improve the accuracy of recognition and vehicle model's speed employing improved YOLOv3 model. The dataset – BIT- vehicle resulted in depiction of improved recognition speed.

Xiaosong [79] improved the method of tracking to detect the increased number of counts for speedy vehicles. This work was based on pairing results over the distinct atmospheric conditions like the dusty and foggy weather.

5.1.3 Proposed Model for Moving Object Detection

In this section, the proposed work explored the work carried out for the development of investigated working model, this work also focuses on the proposed background subtraction method that develop a background independently each pixel relative to the previous pixel value. Here, the bilateral blur filter is applied for improving the model. To ensure uniformity, frames are drawn and converted to grayscale levels. Frame differencing stores different coordinates in grayscale 1 and grayscale 2 frames to detect irregularities or intensity variations in two frames. The bilateral blur is used to smudge the frame for saturation and dilation that make it easier to find contour areas. All of this act as normalization for more accurate motion results. This chapter developed a frame difference scheme using a background subtraction technique for detection of moving vehicles in a foggy environment. Here, input is provided through cloud environment and output is also delivered to the cloud environment. This work is presented in five distinct sections:

- To compute differences between two adjacent frames.
- To change the moving object to grayscale with a threshold and update the background.
- To blur and smooth out the edges and remove noise and the threshold is used to avoid unusual false positives.
- To apply the morphological operators to improve the quality of evidence
- To demonstrate the shape of a moving object.

The final outcome depicts the better detection quality of the moving vehicles in foggy environment. The basic working model of the suggested work is shown in *Figure-5.1* and Similarly, *Figure-5.2* depicts the it's working with Firebase Google's cloud platform.

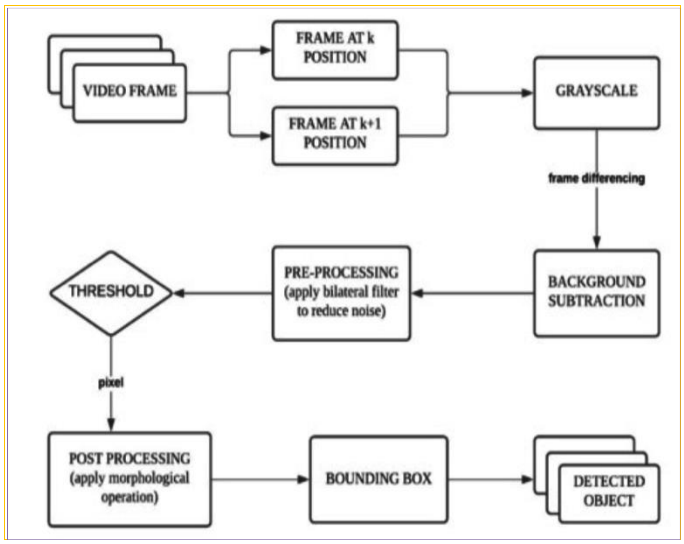


Figure 5.1 Basic work flow of proposed method

5.1.4 Steps and Procedure

The main procedure is performed in two main steps:

- Open the application: Here, the suggested task is started through the application.
- Cloud Data Connectivity: This task is experimented through the connectivity using Google's Firebase platform in the cloud environment. The proposed task is experimented through the following steps. The

forthcoming sections also reveal a brief description about the platform, tools as well as services and methods used to accomplish this work.

5.1.4.1 Firebase

In this work, the Firebase [80] [81] [82] platform is used and its services are acquired from the Google. It accelerates application development through its inbuilt services. It offers BaaS or backend as a service, so Firebase handles cloud infrastructure and all backend requirements. The main benefits of the firebase platform are to allow users to grow and deploy their application faster. The Firebase has some great products also it enables hosting, and has APIs for machine learning tasks such as predictive text, image captions, and more. The working system design using firebase platform is given in fig.5.2.

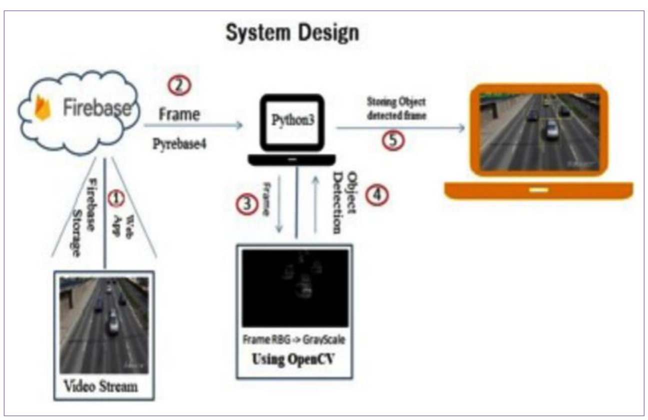


Figure 5.2 System design to demonstrate connection of Firebase to PyCharm

5.1.4.1.1 Firebase console

To work with the Google's Firebase,user/ programmer must have a Google account then, open the Firebase console, register and provide the app with Firebase. It has following project environment according to the requirements: (i) iOS for Apple, (ii) Android, (iii) Network. In the Firebase console click on the web setup link present on the Auth tab to get the authDomain, databaseUrl, apiKey and storageBucket variables needed to link it to the database.

{
"apiKey": "AIzaSyBqDOtkV9ue802-9KonKcdvJRyuL1xzbhI",

"authDomain":"object-detection-project-ce0b4.firebaseapp.com",
"projectId":"object-detection-project-ce0b4",
"databaseURL":"gs://object-detection-project-ce0b4.appspot.com",
"storageBucket":"object-detection-project-ce0b4.appspot.com",
"messagingSenderId":"1050244789749",
"appId":"1:1050244789749:web:1b516c54ec2ef3a8025cbc",
"measurementId":"G-QCL2WVLCQC"
}

5.1.4.1.2 Firebase Storage - Pyrebase

After adding a new project in the Firebase console, it is necessary to import Pyrebase services to the application for using Firebase. The interface of Firebase REST API is Pyrebase i.e. one can use Python to manipulate your Firebase database. Pip is required to install Pyrebase and its dependencies through a package manager to install and manage inbuilt packages of

Python which are acquired from third-party depository. The Pyrebase apps can use many Firebase services:

1. firebase.auth() for Authentication purpose
2. firebase.database() for Database services and management
3. firebase.storage() for Storage management

The architectural diagram of firebase (Google's) platform for cloud storage services is depicted in fig. 3.

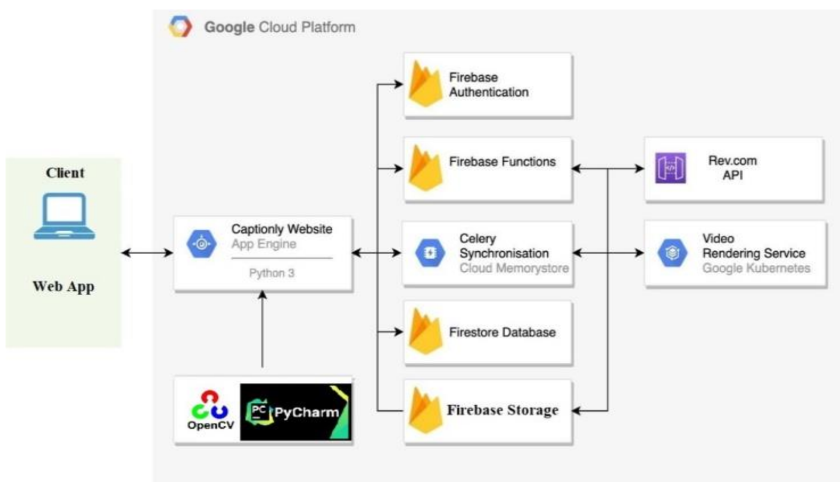


Figure 5.3 Architectural Diagram of Cloud (Firebase)

https://blog.smartandnimble.com/content/images/2019/11/captionly_architecture.jpg

5.1.4.1.3 Firebase Storage - Read video through frames

The Cloud Storage provides additional protection for Google's Firebase when uploading and downloading files from the Firebase app, irrespective of network quality. Here, SDKs are used to accumulate video, images, audios, or other content generated by the user, the Google Cloud Storage is used to store and retrieve the same data on the server.

With the help of the root or child method, the paths may be built to data with the Storage service and the download method takes the path to the saved database file.

5.1.4.1.4 OpenCV

Here, an OpenCV (an open-source library for computer vision, machine learning, and image processing) plays a vital role in real-time work in this system. It can be used to refine videos and images to detect faces, features, objects or even the handwriting of a person. To perform object detection, the following modules must be imported.

1. import cv2 - python library for solving computer vision problems.
2. import imutils - a set of functions to facilitate basic image processing operations and viewing Matplotlib based outcomes.
3. import time - useful for knowing and analyzing each object found in the video.

5.1.4.2 Grayscale

Here, the grayscale format is converted from other RGB spaces. It represents shades of black and white from 0 to 255 that contains 8-bit/pixel (bpp) data. This reduces the size and complexity of the model and is important because other algorithms are adapted to work only with grayscale images.

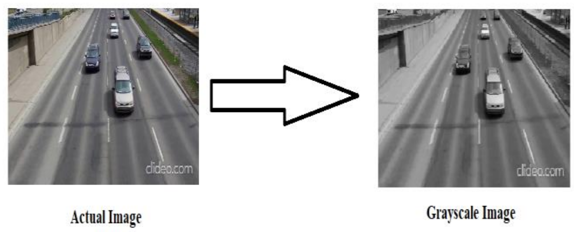


Figure 5.4: Result of the execution of grayscale dependent

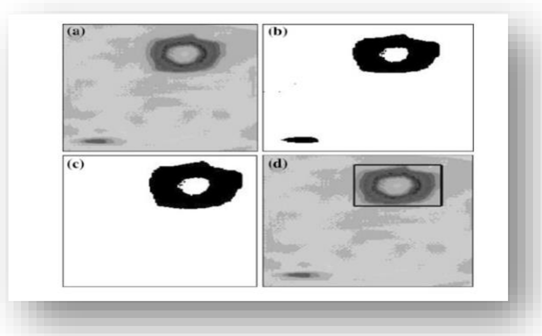


Figure 5.5: Detection result. (a) Grayscale image; (b) Grayscale threshold image; (c) Grayscale threshold image after clutter removal; (d) Grayscale image with object outlined

5.1.5 Background Subtraction

The background subtraction method allows to detect the moving objects or vehicles from video captured by fixed/ static camera. Object detection in this approach would be based upon identification of changes between sequence of frames. The Background model works as a reference here. Reference and identified changes on the new frame are used to detect movement of objects. At the time of object detection, state could be static or moving. This method compares pixel differences between the adjacent frames to identify the state (static or moving) of object.

In this section, cv2.absdiff() function is applied to determine the absolute difference between the pixels of two image arrays. We can extract pixels from moving objects with this. To use cv2.absdiff, first convert the image to grayscale (grayscale is the grayscale area that changes from black to white). The frame differencing function can be computed by using the given equation (1) and detects object between two frames.

$$\boldsymbol{F}_i(x, y)=|\boldsymbol{F}_{\boldsymbol{k+1}}(x, y) - \boldsymbol{F}_{\boldsymbol{k}}(x, y)| \quad (1)$$

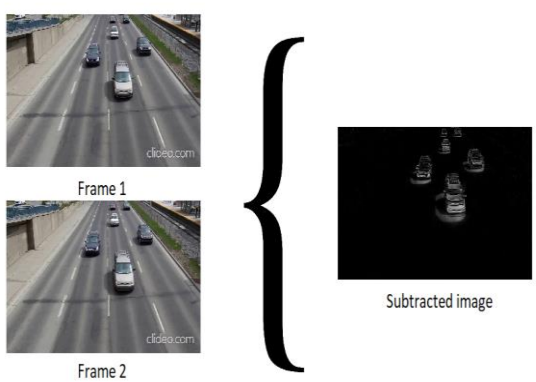


Figure 5.6 Result of frame differentiation dependent on picture successions

5.1.5.1 *Blurring – Bilateral Filter*

Image blur refers to making an image less vivid. To accomplish image blur various low-pass filter cores are used. The bilateral filter is a non-linear image smoothing filter, preserves edges and reduces noise. It replaces each pixel's intensity with a weighted average of nearby pixels' intensity values. Similar to Gaussian filter, the bilateral filter also considers adjacent pixels with a set weight. This weight is made up of two sections, the first of which is the same as the one utilized by the Gaussian filter. The second factor considers the intensity difference between the neighbors and the estimated pixels [127].

$$\boldsymbol{BF}\,[\mathbf{I}]_{\boldsymbol{p}} = \frac{\mathbf{1}}{\mathrm{W}_{\boldsymbol{p}}} \sum\nolimits_{\boldsymbol{q} \varepsilon \boldsymbol{S}} \mathbf{G}\boldsymbol{\sigma}_{\boldsymbol{r}}\left(\left|\left|\boldsymbol{p}-\boldsymbol{q}\right|\right|\right) \mathbf{G}\boldsymbol{\sigma}_{\boldsymbol{r}}\left(\left|\mathbf{I}_{\boldsymbol{p}}-\mathbf{I}_{\boldsymbol{q}}\right|\right)\mathbf{I}_{\boldsymbol{q}} \quad (2)$$

cv.bilateralFilter() is very effective at removing noise while keeping the edges sharp. However, it works slower than other filters. The cv.bilateralFilter() function accepts the following parameters.

- The first parameter is the source image.
- The second parameter is the diameter of each quarter of the pixels used during filtering. If it is not positive, sigmaSpace will calculate it.
- The third parameter is sigmaColor which is sigma filters in the color space. A higher parameter value means that the colors further next to the pixel are blended, resulting in a larger area with semi-flat colors.
- The fourth parameter is sigmaSpace which is filter sigma in coordinate space. Higher parameter values mean farther away pixels are affected as long as they are close enough. If d > 0, the size of the

environment is determined independently of the sigmaSpace. Otherwise, d is proportional to sigmaSpace.

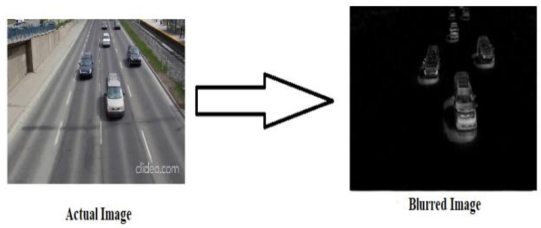


Figure 5.7 Result of the execution of bilateral blurring dependent on picture successions

5.1.5.2 Threshold

The threshold (T) is a more adopted segmentation technique used to separate an object from the particular frame. The threshold [83] [84] is determined by comparing the value of each pixel in the image (pixel intensity) to a specific value of threshold. This separates the input image's pixels into two groups:

1. Pixels with a value of intensity less than the threshold.
2. Pixels with a higher intensity more than the threshold value.

Now, these two groups receive different scores depending on the type of segmentation.

The pixel value is set to 0 if it is less than the threshold [83] [4] [85] else it is set to the utmost value. The cv.threshold function is used to set the threshold value.

- First parameter is an input frame on which Gaussian Blur operation is applied.
- Second parameter is classifying pixel values by threshold used.
- Third parameter is the utmost value which is assigned to a value of pixel above the threshold.
- Fourth parameter is various types of thresholds indicated, offered through OpenCV.

The basic traction described above is performed with type cv.THRESH_BINARY. The classification function can be computed by using the given equation (3), this function classifies the pixel during runtime and also uses threshold (T) for classification for computing maximum value of pixel.

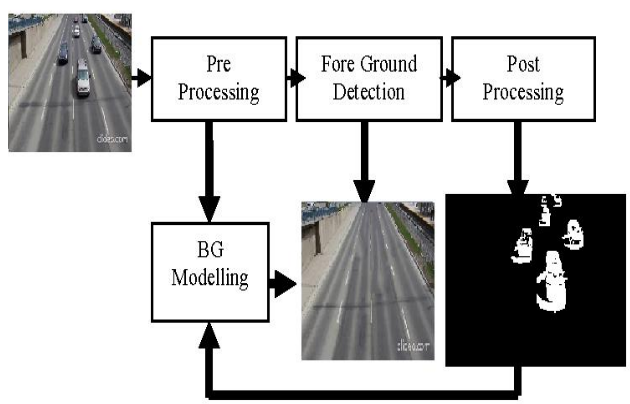


Figure 5.8 Working model of proposed method with outcome

5.1.5.3 Dilation

Dilation is a technique that adds pixel to the boundaries of an object in an image. It is performed on binary images. The main effect of dilating a binary image is to continuously enlarge the boundary of the foreground pixel area [32] [86] [85]. Thus, the pixel area in the foreground expands as the holes in this area get smaller.

The dilate () function accepts the following parameters.

- The first parameter is a source image with a number of channels (all processed independently of each other)
- The second parameter is the kernel element, the origin of this is determined by the anchor (standard (-1, -1), that is, in the middle of the structuring element.

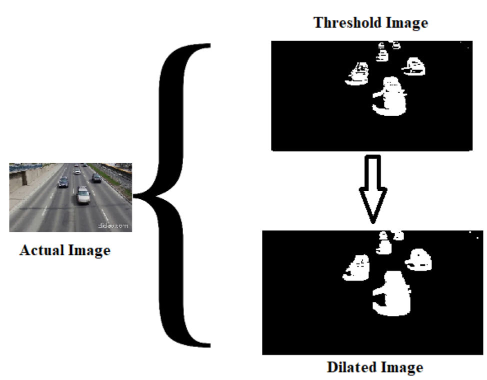


Figure 5.9 Result of the execution of dilation dependent on picture successions

- The third parameter iteration shows how often the process should be repeated. This is at times useful to orient the image to edge pixel readout or when the image is irregularly shaped. This is done using the "bordertype" and "borderValue" arguments.

$$\mathbf{cfn}(\mathbf{x},\mathbf{y}) = \begin{cases} \mathbf{maxValifsrc}(\mathbf{x},\mathbf{y}) > T \\ \mathbf{0otherwise} \end{cases} \tag{3}$$

5.1.5.4 Bounding Box

A bounding box [87] [88] is an imaginary rectangle that serves as a starting point for finding an object and creating a collision box for that object. To draw these rectangles on the image, sketch the objects of interest in each image, and determine their X and Y coordinates, this work, uses contour recognition. This makes it easier for machine learning algorithms to find what they are looking for, identify collision paths, and saves valuable computational resources. A contour is a closed curve that joins all continuous points of a certain color or intensity. It represents the shape of the object found in the image.

In order to correctly identify the bounding box, we need to convert the image to a monochromatic color format (such as grayscale) and then apply a binary threshold. Applying binary thresholds makes the object completely black and white. The object's edges are completely white with the same color intensity. In fact, this is necessary for the contour recognition algorithm to function properly. It detects the boundaries of objects with white pixels (and of course with the same intensity, since each white pixel has a value of 255). Black pixels with a value of 0 are used as a background and ignored.

The cv.findContours() function, which store coordinates (x, y) at the boundary of a particular form, and takes three arguments:

1. Original image,
2. Contour extraction mode.
3. Contour approach method.

And shows modified images, contours and hierarchy.

A. Finally, the algorithm will detect the objects in the frames.

B. And in the end the processed data will detect object and send the feedback to the user so that the user can take the appropriate action immediately.

C. Close the application.

5.2 Experimental Setup and Result Analysis

To test the efficiency of the proposed method, the analysis deals with three video datasets in this paper work. The input videos are randomly captured and taken from. The results are shown in Figure 5.9, where (a) the time is normal; (b) foggy and snow weather; (c) binary output. It selects three images from each video as shown in the results. In Figure-: (a) and (b), in each Figure, the first row depicts video frames and the second row focuses on the tracking of moving vehicles using the proposed method. In Figure-

5.9 (c) computes the detected results of moving vehicle on highway in binary form. In this figure, first row depicts original frame, second rod consist of ground truth frame and third row clearly shows the extracted information of moving vehicle in binary form. The binary results demonstrate better qualitative observations and shows that the proposed method is suitable for moving vehicle detection under the Google's Firebase cloud environment.

Model	Frame-1	Frame-2	Frame-3
Original Frame			
Detected Frame			

Model	Frame-1	Frame-2	Frame-3
Original Frame			
Detected Frame			

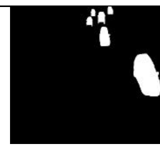
Model	Frame-1	Frame-2	Frame-3
Original Image			
Ground Result			
Detected Result			

Figure 5.10 Detection result. (a) normal weather (b) Snow/foggy Weather; (c) Binary output

Further, the performance is evaluated using various performance metrics such as precision, recall F-Measure defined in equations (1)-(3).

$$\text{Precision} = \frac{\text{Total Identified images}}{\text{Total Correct images present in Database}} \quad (1)$$

$$Recall = \frac{\text{Total Identified Corrected images}}{\text{Total Correct images present in Database}} \quad (2)$$

$$F\text{-}Measure = 2 \times [(Precision \times Recall) / (Precision + Recall)] \quad (3)$$

Table 5-1Precision, Recall and F-score

Sequence	Prec	Rec	F-Meas	FP_Err	FN_Err	Total_Err
Highway	0.8887	0.9439	0.9155	2.4414	1.2096	3.6510
Foggy	0.84	0.824	0.831	1.100	0.909	2.009
Snow	0.8261	0.7851	0.8051	1.2359	0.8090	2.0449

This work also solves the problem related to adverse weather conditions and detect moving vehicle on road against challenging issues [73] [89] in real-time environment.

5.3 Summary

In the real-time environment, during foggy weather condition, it's very difficult to drive safely. In adverse weather condition such as fog, the accidental cases happened on the highway in foggy or snow environment due to poor vision or blurred vision. So, this chapter presents a background subtraction-based method for moving vehicle detection in the foggy environment. All this work has been carried out in Google's Firebase platform that provides a new direction for video surveillance application in the area of computer vision.

6 Vehicle Detection in High Density Traffic Surveillance Data Using YOLO.5

In this chapter, a YOLO. v5 (*You Only Look Once*) algorithm is used to identify the different objects on road such as truck, car, tram and vans. YOLO.v5 is the latest algorithm in the family of YOLO. To train the YOLO. V5, KITTY dataset is used having 11682 images having different objects in traffic surveillance system. After training and validating the dataset, three different models have been constructed setting various parameters. All the models have been evaluated using three performance metrics such as precision, recall and mean average precision (MAP). Final model has attained the best performance as, 93.5% precision and 90.7% recall and 0.67 MAP for different objects. The results depict that the proposed model has improved results as compared to state-of-the-art approaches in term of performance and also reduces the computation time and object loses. Organization of the chapter is as follow, Section 6.2 describes the past literature review and shortcoming of the past methods. Section 6.3 describes the complete methodology and framework of the proposed approach; next section shows the implementation phase and Section 6.5 shows the results obtained by the proposed approach. Last sections have conclusion and future direction.

6.1 Introduction

With the advent of technological advancements and availability of enormous data in the form of audios, videos, text and other forms, Artificial Intelligence emerged as one of the most significant technologies that simulates human intelligence through computers and several algorithms. Computer vision is one of the prime domains that enables to derive meaningful and crisp information from digital media. It includes several sub domains such as facial recognition, pattern detection, image classification, object detection and many others. Object detection is used in several applications such as advanced driver assistance system [90] [91] [92] motion-based identification (Srigrarom, Sutthiphong, & Kim Hoe Chew, 2020), access control [93] , traffic violation monitoring [94], vehicle tracking [95] [96] [97], counting vehicles on highways [98] [99] (Fachrie, 2020) , classification of vehicles [100] [101] [102] and parking spot detection and monitoring [103] [104] (Zhang & Du, 2020) . Object detection comprises of three main methods including: detection of an object, frame to frame object tracking and evaluation of results. In past decade, traditional approaches used for object detection comprises of two main steps: i) extraction of different parts in image using sliding window of various sizes ii) classifying the objects to their respective classes. These methods require high computation cost and multiple stages to classify an object and having low accuracy.

In India, real time traffic monitoring and surveillance is more arduous and challenging task as compared to the other countries. Figure 6.1 depicts the complexity of images extracted from traffic monitoring cameras for vehicle detection. The vehicle detection from traffic images is a critical task as images are affected due to several factors such as lightening conditions, weather changing situations such as fog, excessive heat, occlusions, motion-perspective and having different types of objects such as auto-rickshaw, carts and bicycles. As per recent study, it was found that road density in Delhi is about 155 Km/100,000 population with 80 vehicles/km. It was also found that cycle time at intersection ranges from 120 to 180 seconds leading to long queues mainly in peak hours. Thus, vehicle detection from such heavy traffic monitoring leads to several challenges: vehicle face captured with distinct angles (front, back, side), thus differs in vehicle size computation, high overlap and cut-off of vehicles due to heavy traffic.

Figure 6.1. Traffic Congestion on Indian Roads

With an increase in traffic in all urban and sub-urban areas, enormous traffic problems occurred and to solve these problems & to improve the efficiency of moving traffic, consistent and accurate real time traffic information should be obtained. This real time information needs to be extracted automatically from real time traffic monitoring video cameras. Therefore, to efficiently extract such information and detection of moving vehicles in videos, an accurate, reliable and precise traffic surveillance system is needed and object detection plays a significant role [90]. Using the past methods [105] [106] [93], such real time monitoring is not possible. In recent times various neural network-based algorithms such as Convolution Neural Networks (CNN) [91], Faster R-CNN(Ross-CNN) [107], YOLO (You Only Look Once) [108] and YOLO. v3 [91] has been used and having the better results as compared to the traditional methods, low operating cost and much faster response time.

However, for further improvement, recently YOLO.v5 has been introduced which has better accuracy and can be train on the small dataset with less computation time. YOLO.v5 is a single-stage object detector algorithm which uses cross stage partial network (CSP) on a backbone to extract key features from an input image. Then neck model is used to generate feature pyramids which scales the object with different sizes and improves the performance on unseen data. So, objective of this chapter is to detect traffic objects such as Car, Truck, Van and trams. in real time video system, using YOLO.v5 for better accuracy and applicability in high traffic system.

6.2 Related work

Moving object detection had always been a challenging task due to various factors such as camouflage, misclassification, dynamic background and various objects on the roads. Various traditional approaches have been used for object detection in traffic surveillance videos such as frame differencing, temporal differencing, background subtraction and optical flow [109]. Dong *et.al.*[115] proposed an approach which determines object and its shadow separately based on RGB color space with edge ratio. It consists of three steps: In first step, distinct features of object and its shadow were analyzed in RGB color space which were further segregated based on brightness distortion and chromacity of pixels of current and background image [79] . In another work, choi et.al. proposed an approach based on fast illumination changing condition. The approach was based on chromacity and brightness ratio model [110]. Stauffer and Grimson proposed an approach based on Gaussian Mixture Models (GMM) which evaluated the state of each pixel, detecting many false negatives [40]. This approach was improved by filtering and learning methodology by Lee [111]. In another work, Hague *et.al*. proposed a novel learning method [41], thus yielding improved results as compared to [111] [40] . Chaudhari and porwal proposed a method for thermal imaging based on multi-target detection method [112]. In this work, remote sensing-based technique was proposed for real-time object detection for defence domain by Dordevic *et. al.* [113]. Apart from these techniques, few authors have worked on background subtraction methods. Akula *et. al.* proposed a method in infrared frame sequences that reduces variations in the background [114]. Dong *et.al*. [115] proposed a method in thermal video frames which extracts feature points using difference of gaussian filters and classify them into foreground or background points using R means clustering [115] . H. Tao *et.al*. [116] focused on detection of smoky vehicles based on vehicle rear detection and multi-feature fusion. The approach was evaluated on 157,550 frames including 4 videos with non-smoky vehicles and 98 short videos with one smoky vehicle. The results were evaluated using precision and recall metrics [67] .

These approaches have required high computation time, can't be trained on large dataset and the accuracy lies between 46-53%. To overcome these issues, lead to emergence of deep learning technology in moving object detection. In past decade, deep learning techniques are widely used for object detection due to its high accuracy and low computation time. Ross Girshik's et.al proposed a R-CNN model to find the optimal regions of images based on non-maximal suppression method. The drawback of the approach was high computational time and complex training steps [110]. To further improve upon, Ross Girshik's et.al. improved upon their R-CNN model and proposed Faster R-CNN which iterates the computation of computation of candidate windows [107].

In recent research of object detection, Apart from convolution neural network, Joseph Redmon et.al. propose YOLO algorithm which utilizes image segmentation, attain candidate regions and improves the speed of detection and accuracy. Lu, et al., 2019 [117], Hong, et al., 2020 [118] proposed a YOLO based real time object detection for videos. The author incorporates Google Inception Net to improve the YOLO architecture. The approach was evaluated on videos of smart cities from Xiamen municipal transportation bureau consisting of 8000 images. The results depict that the approach outperforms state-of-the-art techniques [118]. Hong *et. al.* proposed an algorithm to detect vehicles in complex scenarios-based YOLO.v3. This version works efficiently on the smoky data. A codec module was developed to detect vehicles of different scales. The approach was evaluated on KITTY and UA-DETRAC datasets using precision and recall performance metrics [117]. In another work, Kim *et. al* [91] integrated additional prediction layers using spatial pyramid pooling into YOLO- V3. The result was evaluated on precision, recall and frames recognized by second performance metrics. The approach attains 76% precision which outperforms other state-of-the-art techniques when evaluated on benchmark dataset of UA-DETRAC. In this latest work Jain *et. al*. applied YOLO.v3 algorithm on *Karlsruhe Institute of Technology and Toyota Technological Institute* (*KITTY*) dataset consisting of 80 classes among which car, bus, truck, motorcycle and train were 5 major classes [119] with a precision *86.7*% and Recall *81.7*%.

In this research paper authors have implemented the *YOLO.v5* with three trained models using various parameter settings for better accuracy, to reduce the computation time and check the applicability in dense traffic conditions for Indian traffic surveillance system. Authors has used the *KITTY* dataset for the training, validation and testing purpose having *11682* images. The models are trained for *4* different classes such as truck, car, trams and Vans.

6.3 Proposed Framework

In this section, the framework of the proposed approach is presented. The main strategy was accurate prediction of vehicles from traffic surveillance videos of highly populated and congested metropolitan cities of India.

6.3.1 YOLO.v5 (You Only Look Once. Version 5

Convolution neural network, one of the techniques of deep learning was selected due to its availability, less computation cost and improved performance (Fig. 6.2).

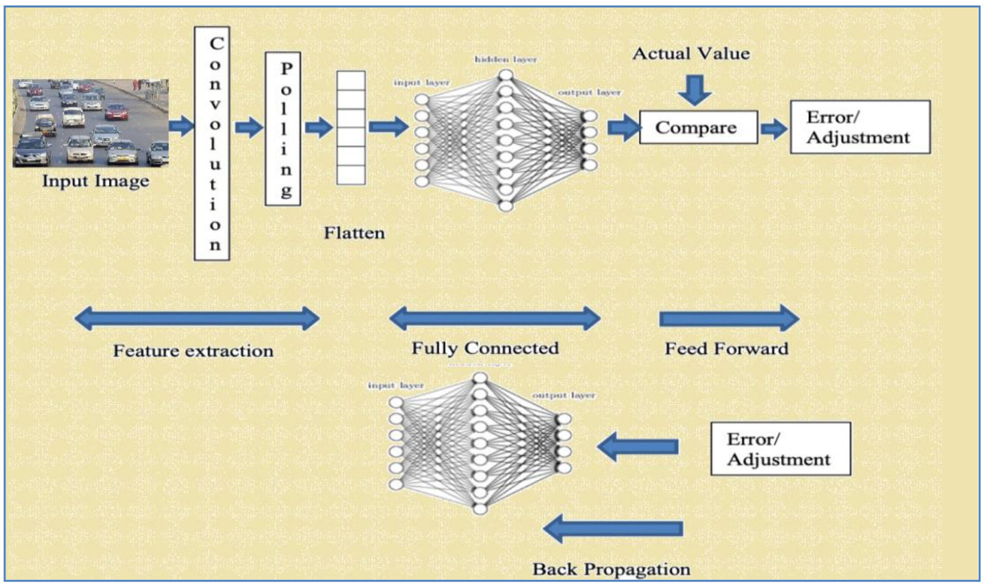
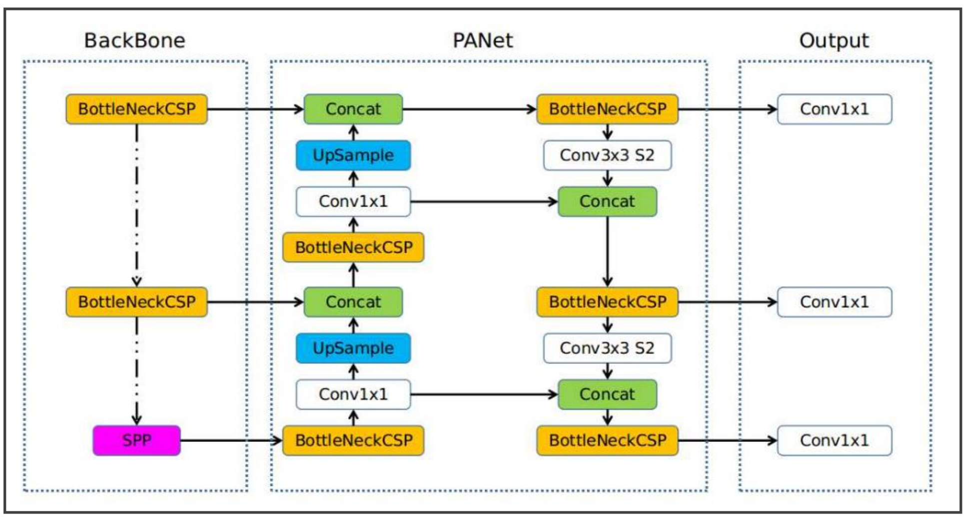


Figure 6.2 CNN Model for YOLO.v5 in traffic surveillance system (Choi, Chang, Yoo, & Choi, 2012)

The YOLO algorithm was selected as it is a unified model with all phases of neural network [32] and it computes the features and predicts all objects at the same time. The YOLO object detection algorithm is used to detect vehicles in high density traffic. Yolo algorithm has various versions such as Yolov1, YoloV2 and YoloV3 [91] [107] . These algorithms had some major drawbacks which leads to the evolution of new version 5. The major challenge with Yolo.v3 was it was unable to detect various sized vehicles such as car, van bus, truck and others in real time traffic [91]. Therefore, to overcome this drawback, Yolov5 is used in this work. Yolov5 is a single-stage object detector algorithm which uses cross stage partial network (CSP) on a backbone to extract key features from an input image. Then neck model is used to generate feature pyramids which scales the object with different sizes and improves the performance on unseen data. In this work, PANet is used as neck model. At last, model Head is used for object detection. It creates several anchor/bounding boxes with various class probabilities and object ness scores which improves with every epoch.

Also, it was written in python, that is easy to installed and integrated with all devices including IOT devices. Also, in place of Darknet, it incorporates Pytorch which has many advantages in object detection in real time. Fig 3 shows the architecture of YOLO v5.



SPP- Spatial Pyramid Pooling, CSP- Cross Stage Partial Network

Figure 6.3 Architecture of YOLO.v5

In this research paper different traffic objects such as Car, Truck, Van and trams are detected in real time video system using YOLO.v5 for better accuracy and applicable in high traffic system.

6.3.2 Dataset Description

In this work, the proposed framework is applied on KITTY dataset which has 80 classifiers and is generally used for multiple object detection [119]. The five main objects or classifiers used are car, bus, truck, motorcycle and train. It consists of total 11682 number of images which are split into 9736 as training images and 1946 testing images. some sample images from KITTY dataset are represented in Figure 6.4

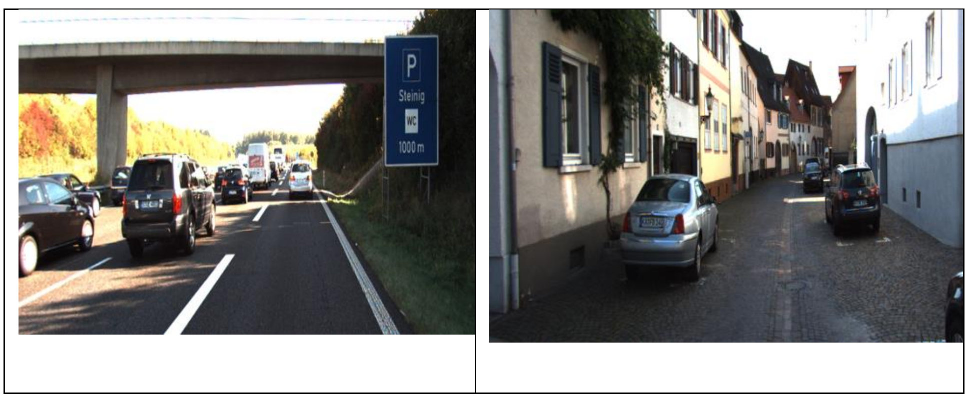


Figure 6.4 Sample images of KITTY Dataset

6.3.3 Methodology

In this work, we have built three versions of YOLO v5; namely Baseline model, model V2, model V3. In baseline model, the model was trained from the scratch. It does not require any pre-trained model. The number of epochs set to 25 as compared to default (300) as time to execute each epoch was

10-12 mins. Due to limited runtime, number of epochs changes from default (300) to 25. It works on default image size of 640 pixels.

In model V2, epoch size is same as baseline model and image size is changed from default (640) to 1280. In model V3, it incorporates parameters of both the baseline model and model V2, along with some additional changing parameters. The epoch size and image size are same as model V2 i.e. 25, 1280. Also, hyp default is hyp.scratch-low.yaml is changed to hyp.scratch-high.yaml. The prime difference is that instead of using no weights and starting training from scratch, YOLO.v5s.pt is used as pre-trained model weight.

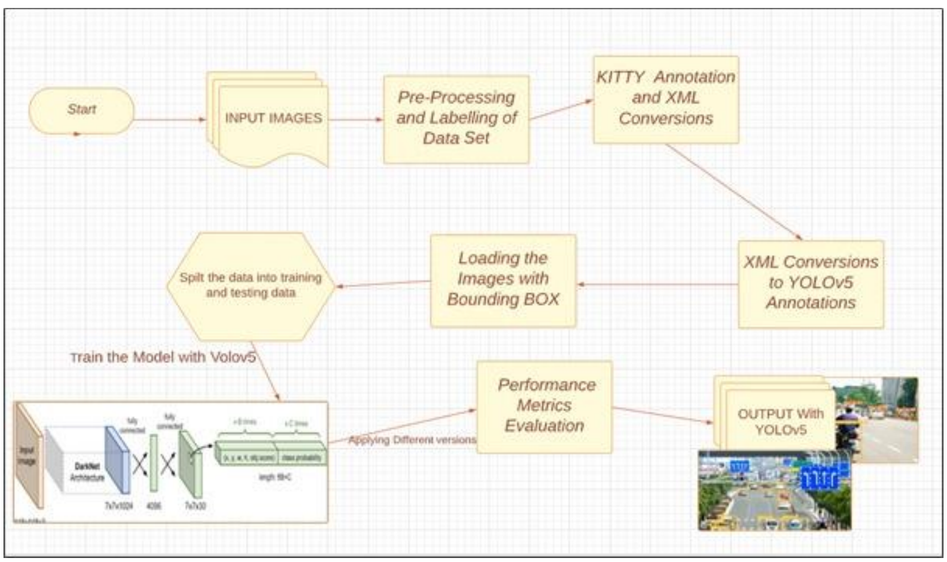


Figure 6.5 Framework for Object detection using YOLO. v5

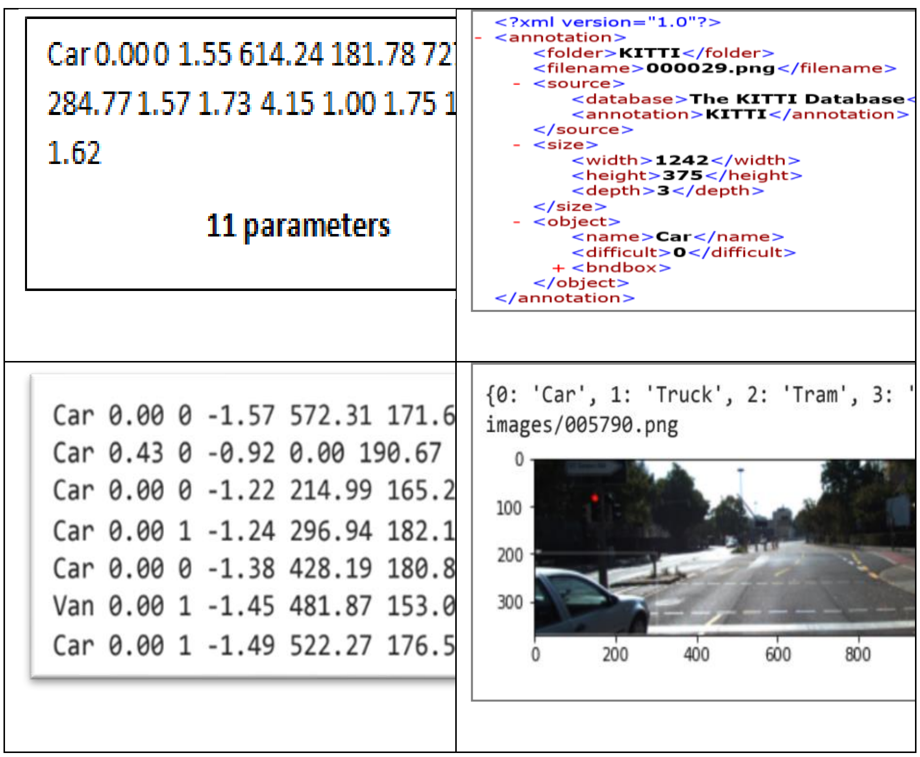
The framework of the proposed framework is shown in Fig 6.5. The model was trained using Google colab providing free access to Runtime GPUs with no extra configuration. Initially, KITTY dataset is downloaded and all the images are extracted from the zip folder. After extraction of images, the images are pre-processed and annotations and labelled images are extracted. After pre-processing, KITTY annotations with 11 parameters are converted into XML notations with 11 parameters. After this, XML notations with 11 parameters are converted into YOLOv5 notations with 5 parameters. After conversion into YOLO notations, images with bounding boxes are loaded and the dataset is divided into training, testing and validation set. The model is trained on three versions of YOLO algorithm, namely Baseline model, Model V2 and model V3. After evaluating the models, the performance of these models is compared through three performance metrics namely, precision, recall and mean average precision. After evaluation of results, relevant objects are detected from images.

6.4 Implementation

The proposed approach is implemented in Python in COLAB environment. As computational time of YOLOv5 is high, the GPU runtime selected. The complete implementation procedure is explained as follows:

- To start with,YOLO. v5 is cloned into COLAB from Github.
- After cloning, various libraries with several functionalities are imported such as Pytorch, Random, Shutil, Sklearn, ElementTree, Numpy, Matplotlib and Image.
- After importing all the libraries, KITTY dataset is downloaded on COLAB itself. As it is bulky dataset, normal downloading is not appropriate.

- After downloading the dataset, the images in the dataset are preprocessed. For preprocessing, the dataset is converted into training dataset and labels have been changes to appropriate annotations and are unzipped for further processing.
- KITTY annotations are first converted into XML Notations with 11 parameters.
- After this, XML notations are converted into YOLOv5 notation
- After conversion into YOLOv5 notations, the images folder is organized into Training, Testing and validation images folders as it necessary for YOLOv5 to provide appropriate paths of folders.



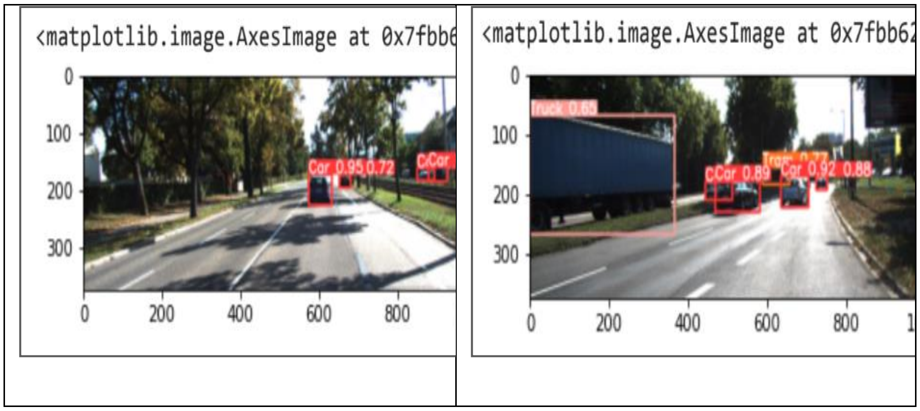


Figure 6.6 Implementation Steps of the Approach

- After creation of appropriate folders, .yaml file is created which comprises of i) source link of the respective folders i.e. training, testing and validations, ii) the names of classes such as car, van, truck and tram with their numbers (fig 6.7).
- After this, wandb is installed for better visualization of performance metrics graphically.
- After this, the data is trained, tested and validated for three models namely, baseline model, model V2 and YOLOv5 model. Then converting annotation images into training, validation and testing data. Data is split in 60:10:30 ratio respectively. All the three models are compared with respect to various parameters in Table 6.1.

Table 6-1 Comparison of Three versions of YOLO.v5 model

Parameters	Baseline model	Model V2	FINAL model
Number of epochs	25	25	25
Image size	640 px	1280 px	1280 px
Training weight	No weights	No weights	Pre-trained model weights
Hyper_parameters	Hyp_default	Hyp.scratch-low.yaml	Hyp.scratch-high.yaml
Training size	640	640	640
Testing size	749	749	749
Validation size	748	748	7482
Yolo layers	213	270	270
For all classes, Precision	0.83	0.89	0.91
Recall	0.83	0.85	0.86
MAP	0.89	0.94	0.95

Fig 6.7 (a) shows the test results of an image showing car as object and with the confidence level 0.95 to 0.64 using Yolo trained model. In the next phase model V2 is trained and shows the results in 6.7 (b). Similarly, last model is trained and tested on KITTY dataset.

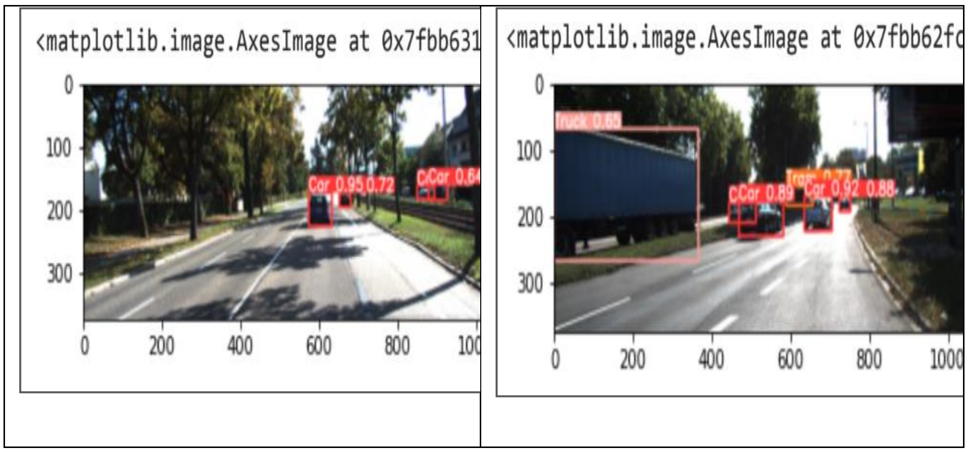


Figure 6.7 (a) & (b). Tested images after training and validation

6.5 Results and Discussion

In this proposed approach, different images of KITTY dataset have been trained and tested using YOLO.v5 algorithms. In implementation phase 3 models are constructed using various parameter setting namely, baseline, model v2 and final model. After training and testing phase, images are validated in the form of batches, where each batch comprises of 16 images. teach batch consists of labelled images and predicted images. Predicted images depicts the probability of finding each object in an image.

For illustration, three batches of each model are depicted in Figure 6.8, 6.9 and 6.10.

Labelled Images	Predicted images
	

Figure 6.8 Labelled and predicted images of Baseline model

Labelled Images	Predicted images
	

Figure 6.9 Labelled and predicted images of Model V2

Labelled Images	Predicted images
	

Figure 6.10 Labelled and predicted survey of Final model

Further, the performance of all the three models of YOLOv5 are evaluated using various performance metrics such as precision, recall and mean average precision defined in equations (1)-(3).

$$Precision = \frac{\text{Total Identified images}}{\text{Total Correct images present in Database}} \quad (1)$$

$$Recall = \frac{\text{Total Identified Corrected images}}{\text{Total Correct images present in Database}} \quad (2)$$

$$Mean\ average\ precision = mean\ of\ average\ precision \quad (3)$$

The values of the performance metrics for various classes of objects in images are depicted in table 6.2.

Table 6-2 Results obtained by Proposed Model

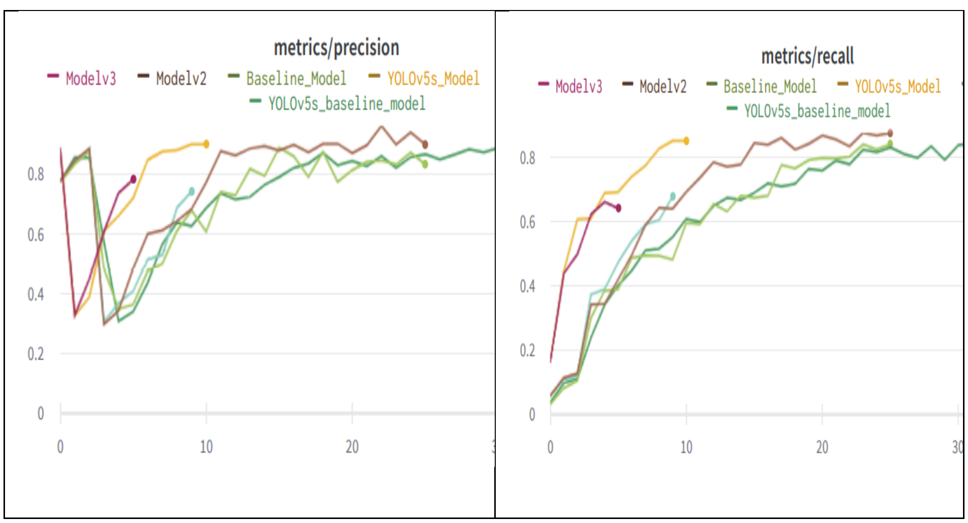
Class	Images	Label's	Precision	Recall	MAP@0.95
All	749	3298	0.918	0.859	0.67
Car	749	2845	0.941	0.891	0.73
Truck	749	91	0.93	0.899	0.72
Tram	749	61	0.875	0.77	0.56
Van	749	293	0.927	0.87	0.66

Figure 6.11 further illustrates performance of different models trained through YOLO.v5 for each epoch. In terms of precision, final model attains stabilized results at the end of 5th epoch as compared to other models which attain stabilized results at or after 10th epochs.

In terms of recall and mean average precision, final model attains stabilized results at the end of 4th epoch which were attained by other models at-least after 8th epochs.

Further, to evaluate the loss at training set, class-loss and object loss were stabilized at 6th epoch which got stabilized after 10th epochs for another model.

The result concludes that final model i.e. modelv3 with pre-trained weights attains best results as compared to other models of YOLOv5.



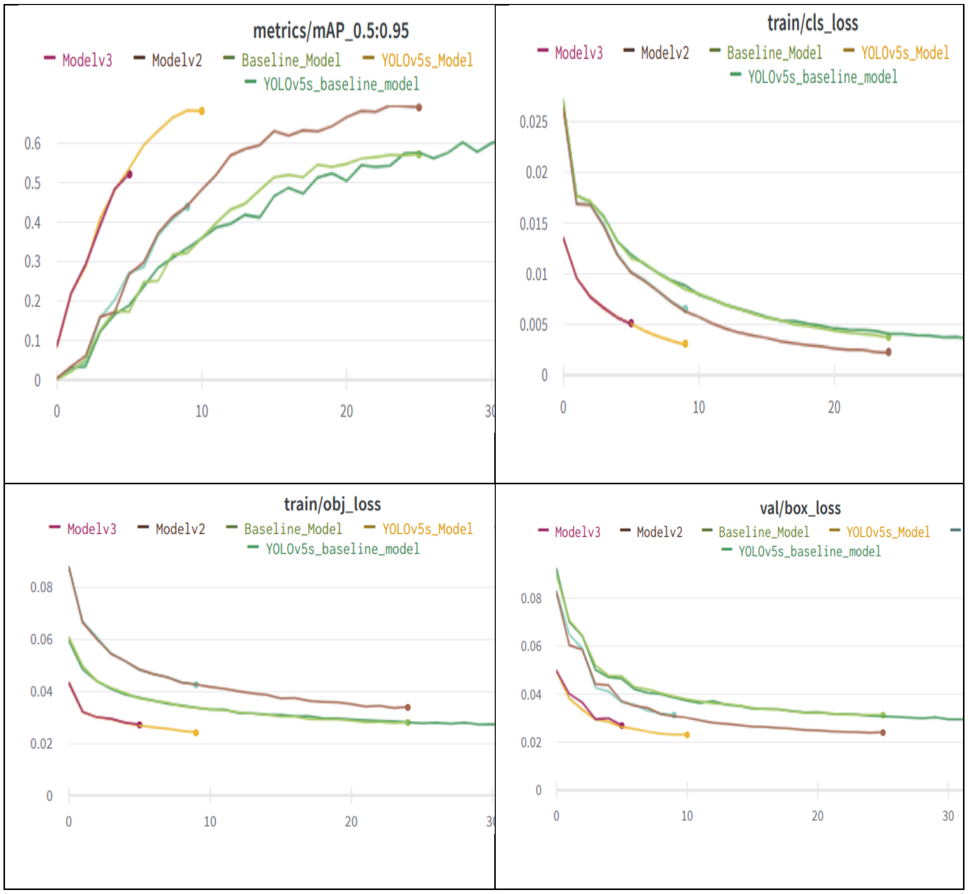


Figure 6.11 Performance of various versions of YOLOv5 models in terms of performance metrics

Further, the results attained by our proposed model was compared with the state-of-the-art approaches on KITTY dataset. The comparison of results is presented in Table 6.3**.**

Table 6-3 Comparison of proposed and state-of-the-art approaches

Class / performance metrics	**Proposed Approach**		**State-of-the art Approach**	
	Precision	**Recall**	**Precision**	**Recall**
All	**0.918**	**0.859**	0.87	0.82
Car	**0.941**	**0.891**	0.89	0.84
Truck	**0.93**	**0.899**	0.85	0.82
Tram	**0.875**	**0.77**	0.83	0.72
Van	**0.927**	**0.87**	0.86	0.79

6.6 Summary

Major breakthrough has been achieved by deep learning techniques to derive meaningful information from digital media such as images, videos and other visual inputs mainly in object detection domain. In the past different approaches were presented in the literature for detecting and tracking moving objects in a video stream having low precision and accuracy. Also, in all urban and sub-urban areas of India, enormous and congested traffic has been monitored nowadays due to which traffic surveillance has low efficiency using deep learning techniques. To improve the efficiency of moving traffic system YOLO.v5(You Only Look Once) algorithm has been applied on KITTY dataset. This chapter have implemented Yolo.v5 model for object detection for the traffic management system. Yolo.v5 is the latest version in the Yolo family, it has fast computing method and can also trained the model on the small dataset, where as in previous model's sample training required is more than lakh images. Yolo.v5 has fast computing method and can easily run on the local machines. India has huge traffic in mid-level and large cities, to detect object in real time has low efficiency as compared to other global countries. To overcome these issue author has trained the model attain better results as compared to the state-of-the-art techniques. Final model has attained 93.5% precision and 90.7% recall and 0.67 MAP in all cases as explained above. The results conclude that the proposed approach attains 20% improved results as compared to state-of-the-art approaches.

7 CONCLUSION AND FUTURE SCOPE

The research work focuses on the considered objectives of proposed work discussed in Chapter 2. This section provides a brief conclusion of each chapter of the work and briefly described according to organization and flow of the research work.

The chapter-1 focuses on the definition of the area, the concepts like object detection, object tracking, video surveillance, computer vision, smart surveillance and video surveillance. The introduction also covers the challenges and the application in the real-life application areas.

The chapter-2 'Background' is the collection of all the methods used by various authors in their various works. The computer vision came into existence long time back; the chapter covers the description of the methods. Almost all the used methods till now are consolidated in the chapter. For every research to be followed there is the need of data set on which the methods can be applied so that portion is very beautifully collected in the chapter with the frame size and with the appropriate source like the name of the university and the link.

Then chapter 3 covers the literature study of various methods and techniques proposed by different authors. The literature is also about the improvements, the advantages and disadvantages of the methods followed by varied researchers. The study also includes the role of machine learning, and deep learning in the area of object detection. The methodologies used in the area of object detection for the smart surveillance system are enlisted in this chapter.

7.1 Objectives Achieved

The motivation and the objective in the chapter define the interest of the author in the direction of this work and make this work to get fulfilled. The objectives after the literature and the background study headed towards the achievement of the objectives described in further chapters.

Chapter -4 has developed an adaptive background subtraction method to handle the various challenging problems such as illumination variation. The chapter focuses the problematic video data captured through camera sensors to handle challenging issues available in real-time video scene. This chapter demonstrates better performance of the background subtraction method by using adaptive background subtraction method as compare to state-of-the-art methods.

Chapter-5 has investigated a method that is developed using background subtraction technique. In this work, the data is stored on the cloud environment *i.e.* Google's firebase. The data set used in this chapter was the collected by EFCON INDIA-Yamuna Expressway and Highways Dataset in the adverse conditions of foggy weather.

Chapter – 6 have implemented Yolo.v5 model for object detection for the traffic management system. Yolo.v5 has fast computing method and can easily run on the local machines. India has huge traffic in mid-level and large cities, to detect object in real time has low efficiency as compared to other global countries. To overcome these issue, this work trained the model to attain better results as compared to the state-of-the-art techniques.

7.2 Future Scope

The existing methods for object detection and tracking with the combination of new technologies like IoT, deep learning and cloud computing may form a better way. The security and storage are a big concern that will surely be overcome in the near future. The challenging difficulties discussed over in the thesis will be tried to be eradicated in the coming future. Involvement of Artificial Intelligence and IoT in computer vision may serve a better exploration. The future will try to overcome the challenges discussed in the thesis by applying the better aspect of background modeling technique.

Following are the future enhancements which could be focused in subsequent research work:

- ✓ To focus on moving object detection and Tracking with features from online video captured through static/moving camera in real-time system
- ✓ To develop a GUI based application.
- ✓ To apply Cloud environment

❖ Finally, to develop a GUI based Intelligent Surveillance System using the Cloud environment.

8 References

[1] M. Allili, N. Bouguila and D. Ziou, "A Robust video Foreground Segmentation by using Generalized Gaussian Mixture Modeling," in *CRV 2007, 503 – 509,*.

[2] B. Lee and M. Hedley, "Background estimation for video surveillance," *Image and Vision Computing New Zealand,* p. 315–320, 2002.

[3] M. Haque, M. Murshed and M. Paul, "On stable dynamic background generation technique using gaussian mixture models for robust object detection," in *5th Int. Conference on Advanced Video and Signal Based Surveillance, IEEE,* 2008.

[4] L. Sharma, D. K. Yadav and A. Singh, "Fisher's Linear Discriminant Ratio based Threshold for Moving Human Detection in Thermal Video," in *Infrared Physics and Tech.*, 2016.

[5] D. Yadav, L. Sharma and S. Bharti, "Moving object detection in real-time visual surveillance using background subtraction technique," in *14th International Conference on Hybrid Intelligent Systems, IEEE,* 2014.

[6] D. Yadav, "Efficient method for moving object detection in cluttered background using Gaussian Mixture Model," in *International Conference on Advances in Computing, Communications and Informatics (ICACCI),* 2014.

[7] D. K. Yadav and K. Singh, "A combined approach of Kullback–Leibler divergence and background subtraction for moving object detection in thermal video," *Elsevier, Infrared Physics and Technology,* vol. 76, no. ISSN 1350-4495, pp. 21-31, 2016.

[8] F. Baf, T. Bouwmans and B. Vachon, "Type- 2 fuzzy mixture of Gaussians model: Application to background modeling," in *International symposium on Visual Computing,* Dec,2008.

[9] J. Zeng, L. Xie and Z. Liu, " Type-2 fuzzy gaussian mixture models," in *Pattern Recognition- 41(12):3636–3643,* , December2008.

[10] H. Kim, R. Sakamoto, I. Kitahara, T. Toriyama and K. Kogure, "Background subtraction using generalized Gaussian family model," *IET Electronics Letters,* no. 44(3), p. 189–190, January2008.

[11] N. McFarlane and C. Schofield, "Segmentation and tracking of piglets in images," *British Machine Vision and Applications,* p. 187–193, 1995.

[12] J. Zheng and Y. Wang, "Extracting roadway background image: A mode

based approach," *Transportation Research Board, TRB2006.*, 2006.

[13] C. Wren and A. A. Pfinder, "Real-time tracking of the human body," *IEEE transactions on Pattern Analysis and Machine Intelligence,* vol. 19(7), p. 780 –785, 1997.

[14] A. Elgammal and L. Davis, "Non-parametric model for background subtraction," in *6th European Conference on Computer Vision*, 2000.

[15] H. Lin, T. Liu and J. Chuang, "A probabilistic SVM approach for background scene initialization," in *ICIP 2002*, September2002..

[16] D. Tsai and C. Lai, "Independent component analysis based background subtraction for indoor surveillance," *IEEE Transactions on Image Processing,* vol. 18(1), p. 158–167, January 2009.

[17] K. Toyama and J. K. Wallflower, "Principles and practice of background maintenance," in *International Conference on Computer Vision ICCV 1999*.

[18] S. Messelodi and C. Modena, "A kalman filter based background updating algorithm robust to sharp illumination changes," in *International Conference on Image Analysis and Processing, ICIAP 2005*, 2005.

[19] R. Chang, T. Ghandi and M. Trivedi, "Vision modules for a multi sensory bridge monitoring approach," in *IEEE Conference on Intelligent Transportation Systems, ITS 2004,* 2004.

[20] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *7th lnternational Joint Conference on Aliificial Intelligence*, 1981.

[21] A. Francois and G. Medioni, "Adaptive Color Background Modeling for Real –time Segmentation of Video Streams," in *International Conference in Imaging Science, Systems and Technology*, June.1999..

[22] M. Zhao, N. Li and C. Chen, "Robust automatic video object segmentation technique.," in *International Conference on Image Processing*, September,2002.

[23] P.Pahalawatta, D. Depalov, T. Pappas and A. Katsaggelos, "Detection, Classification, and Collaborative Tracking of Multiple Targets using Video Sensors," in *International Workshop on Information Processing for Sensor Networks* , April,2003.

[24] B. Orten, M. Soysal and A. Alatan, "Person Identification of Surveillance Video by Combining Mpeg-7 Experts," in *WIAMIS 2005*, April,2005.

[25] T. Tanaka, A. Shimada, D. Arita and R. Taniguchi, "Object Segmentation under Varying Illumination based on Combinational Background Modeling," in *Proceeding of the 4thjoint Workshop on Machine Perception and Robotics*, 2008.

[26] Diamantas, Sotirios, Alexis and Kostas, "Optical Flow Based Background Subtraction with a Moving Camera: Application to Autonomous Driving," 2018.

[27] N. Oliver, B. Rosario and A. Pentland, "A Bayesian computer vision system for modeling human interactions," in *International Conference on Vision Systems*, Jan 1999.

[28] S. Bucak and B. Gunsel, "Incremental subspace learning and generating sparse representations via non-negative matrix factorization," in *Pattern Recognition, 42(5) 2009.*, 2009..

[29] A. Krizhevsky, I. Sutskever and G. Hinton, "ImageNet classification with deep convolutional neural networks," *Adv. Neural Inf. Process. System,* p. 1097–1105, 2012.

[30] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv:1409,* p. 556, 2014.

[31] M. Yazdi and T. Bouwmans, "New Trends on Moving Object Detection in Video Images Captured by a moving Camera: A Survey," *Computer Science Review,Elsevier,* pp. pp. 1-66, 2018.

[32] D.K.Yadav and K. Singh, "Adaptive Background Modeling Technique for Moving Object Detection in Video under Dynamic Environment”," *International Journal of Spatio-Temporal Data Science, Inderscience, vol. 1, no. 1,* pp. pp. 4-21, Jan, 2019..

[33] D. K. Yadav, Detection of Moving Human in Vision based Smart Surveillance under Cluttered Background: An Application for IoT, Taylor & Francis, pp 1-296,.

[34] M. Hashemi and M. Hall, "Detecting and Classifying Online Dark Visual Propaganda”,," *Journal of Image and Vision Computing, Elsevier,* pp. vol. 89, pp. 95-105, 2019.

[35] T. Bouwmans, F. Porikli, B. Höferlin and A. Vacavant, "Background Modeling and Foreground Detection for Video Surveillance," 2014.

[36] A. Sanin and C. Sanderson, "Shadow Detection: A Survey and Comparative Evaluation of Recent Methods," *Elsevier, Pattern Recognition, vol. 45, issue 4,* pp. pp. 1684-1695, 2012.

[37] T. Haines and T. Xiang, "Background Subtraction with Dirichlet Process Mixture Models," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Apr. 2014..

[38] P. Dollár, R. Appel and S. Blondie, "Fast Feature Pyramids for Object Detection," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014.

[39] P. Tokmakov, K. Alahari and C. Schmid, "Learning Motion Patterns in

Videos," in *", IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[40] C. Stauffer, "Adaptive background mixture models for real-time tracking,," in *Int Conf. on Computer Vision & Pattern Recognition*.

[41] M. Haque, M. Murshed and M. Paul, "On Stable Dynamic Background Generation Technique using Gaussian Mixture Models for Robust Object Detection," in *5thInternational Conference on Advanced Video and Signal Based Surveillance*, 2008.

[42] K.K. Ng and E.J. Delp, "Background subtraction using a pixel-wise adaptive learning rate for object tracking initialization," in *Proceedings of SPIE Digital Library Visual Information Processing and Communication II*, January 2011.

[43] C. R. Jung, "Efficient Background Subtraction and Shadow Removal for Monochromatic Video Sequences," in *IEEE Transactionson Multimedia*, April, 2009.

[44] D. K. Yadav and K. Singh, "A Combined Approach of Kullback-Leibler Divergence Method and Background Subtraction for Moving Object Detection in Thermal Video," *Infrared Physics and Technology, Elsevier, vol. 76*, pp. pp. 21-31, Feb 2016.

[45] L. Sharma and D. K. Yadav, "Histogram based Adaptive Learning Rate for Background Modelling and Moving Object Detection in Video Surveillance," *", International Journal of Telemedicine and Clinical Practices, Inderscience, vol. 2, issue 1,* pp. pp. 74-92.

[46] M. Rai, R. Sharma, S. C. Satapathy and D. K. Yadav, "An Improved Statistical Approach for Moving Object Detection in Thermal Video Frames," *Special issue 1212: Deep Learning Techniques for Infrared Image/Video Understanding, Springer,* p. vol 81, 2022.

[47] H. Park, S. Park and Y. Joo, "Detection of Abandoned and Stolen Objects Based on Dual Background Model and Mask R-CNN," in *IEEE Access, vol. 8*, 2020.

[48] F. Amin, A. Mondal and J. Mathew, "A Large Dataset With a New Framework for Abandoned Object Detection in Complex Scenarios," in *IEEE Multieedia*, 2021.

[49] W. Liu, P. Liu, C. Xiao and R. Hu, "General-purpose Abandoned Object Detection Method without Background Modeling",," in *IEEE, International Conference on Imaging Systems and Techniques (IST)*, 2021.

[50] D. K. Yadav, K. Singh and S. Kumari, "Challenging Issues of Video Surveillance System using Internet of Things in Cloud Environment"," in

ICACDS-2016, Ghaziabad, 2016.

[51] S. Mishra, D. K. Yadav, F. Tabassum and D. Kumar, "Detection of Moving Vehicle in Foggy Environment through Google's Firebase Platform," *Turkish Online Journal of Qualitative Inquiry,* pp. Vol. 12, Issue 6, pp. 9892-9901, July-2021.

[52] A. Alam, I. Ullah and K. Lee, "Video Big Data Analytics in the Cloud: A Reference Architecture, Survey, Opportunities, and Open Research Issues," *IEEE Access, vol. 8,* pp. pp. 152377-152422, 2020.

[53] S. Pouyanfar, Y. Yang, S.-C. Chen and M.-L. Shyu, "Multimedia big data analytics: A survey," *ACM Computing Survey,* pp. vol. 51, pp. 10, Jan. 2018.

[54] "Video Analytics Market,," Dec. 2012. [Online]. Available: https://github.com/apache/rocketmq.

[55] "YouTube Statistics," [Online]. Available: https://www.youtube.com/about/press/.

[56] "Netflix," [Online]. Available: https://www.netflix.com/.

[57] I. A. T. Hashem, I. Yaqoob and N. B. Anuar, "The rise of 'big data' on cloud computing: Review and open research issues," *Information Systems, Elsevier, vol. 47,* pp. pp. 98-115, Jan. 2015..

[58] N. Elgendy and A. Elragal, "Big data analytics: A literature review paper," in *Int. Conf. Data Mining, Springer,* 2014.

[59] D. M. Jiménez-Bravo, A. L. Murciego and A. S. Men, "Multi-object tracking in traffic environments: A systematic literature review," *Neurocomputing, Elsevier, Vol. 494,* pp. pp. 43-55, 2022.

[60] N. H. Abdulghafoor and H. N. Abdullah, "A novel real-time multiple objects detection and tracking framework for different challenges",," *Alexandria Engineering Journal, Elsevier, Vol. 61, Issue 12,* pp. pp. 9637-9647, 2022.

[61] X. Zhou, C. Yang and W. Yu, "Moving Object Detection by Detecting Contiguous Outliers in the Low-Rank Representation",," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* pp. vol. 35, issue-3, pp. 597-610, March, 2013.

[62] I. Kajo, N. Kamel and Y. Ruichek, "Incremental Tensor-Based Completion Method for Detection of Stationary Foreground Objects," *IEEE Transactions on Circuits and Systems for Video Technology,* pp. vol. 29, no. 5, pp. 1325-1338, May 2019..

[63] D. Kim, H. Kim, Y. Mok and J. Paik, "HLDNet: Abandoned Object Detection Using Hand Luggage Detection Network," *IEEE Consumer Electronics Magazine,* pp. pp 1-1, 2021.

[64] "Microsoft's Wallflower Dataset," [Online]. Available: https://www.microsoft.com/en-us/research/publication/wallflower-principles-and-practice-of-background-maintenance/.

[65] "Change Detection Dataset," [Online]. Available: www.changedetection.net.

[66] S. Rosaline, J. J. J. Sheela, A. M, A. T. Anasuya and M. Velmurnrgan, "Fog Density Detection Based Automotive Vehicle," *European Journal ofMolecular & Clinical Medicine, vol. 7, issue 3, ,* pp. pp. 1633-1639, 2020.

[67] H. Kasban, O. Zahran and M. El-Kordy, "False Alarm Rate Reduction in the Interpretation of Acoustic to Seismic Landmine Data using Mathematical Morphology and the Wavelet Transform," *Sensor Imaging, vol. 11, ,* pp. pp. 113- 130, 2010.

[68] R. Li, X. Xie, P. Wang and C. Jin, "Cloud-based moving object detection for mobile devices," in *20th Conference on Innovations in Clouds, Internet and Networks*, 2017.

[69] R. J. Guo, N. Liu, S. Li, F. Liu, B. Chen, J. Cheng, X. Duan and C. L. Ma, "Pixel-Wise Classification Method for High Resolution Remote Sensing Imagery Using Deep Neural Networks," *ISPRS, Int. Journal of Geo-Information, issue 3,* pp. no. 110, pp. 1-23,.

[70] S. Tuli, N. Basumatay and R. Buyya, "EdgeLens: Deep Learning based Object Detection in Integrated IoT, Fog and Cloud Computing Environments," *4th Int. Conf. on Information Systems and Computer Networks,* pp. pp. 496-502, 2019.

[71] R. C. Miclea, V. I. Ungureanu, F. Sandru and I. Silea, "Visibility Enhancement and Fog Detection: Solutions Presented in Recent Scientific Papers with Potential for Application to Mobile Systems," *Sensors, vol. 21, issue. 3370,* 2021.

[72] R. Singh, S. Singh and N. Kaur, "A Review: Techniques of Vehicle Detection in Fog," *Indian Journal of Science and Technology, Vol 9, issue 45,* pp. pp 1-4, 2016.

[73] C. Wi.irthner, *Object Detection and Tracking with Firebase ML Kit and camera - ML Product Search (Part 3),* 2019.

[74] H. Dwivedi, *Building a real-time object detection app on Android using Firebase ML Kit,* 2019.

[75] D. F. e. al., "Deep Multi-Modal Object Detection and Semantic Segmentation for Autonomous Driving: Datasets, Methods, and Challenges," in *", IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 3,* March 2021.

[76] D. K. Yadav and K. Singh, "A Combined Approach of Kullback-Leibler Divergence Method and Background Subtraction for Moving Object Detection in Thermal Video," *Infrared Physics and Technology, Elsevier, vol. 76,* , pp. pp. 21-31, Feb- 2016.

[77] K. Lochab, D. K. Yadav, M. Singh and A. Sharma, "Internet of Things in Cloud Environment: Services and Challenges," *SERSC-International Journal of Database Theory and Application, vol. 10. No. 5,* pp. pp. 23-32, May, 2017.

[78] S. Yadav, D. K. Yadav, A. K. Budati, M. Kumar and A. Suri, "Automating the Indian transportation system through intelligent searching and retrieving with Amazon Elastic Compute," *Special Issue: Aliificial Intelligence based Network Security and Computing Tech.*

[79] J. Zhang, X. Jia, J. Hu and K. Tan, "Moving Vehicle Detection for Remote Sensing Video Surveillance with Nonstationa1y Satellite Platform," in *IEEE, Transactions on Pattern Analysis & Machine Intelligence, vol. , no. O1,* .

[80] X. Chaojun, Y. Qing, L. Jianxiong and L. Liang, "Research on Vehicle Detection Based on YOLOv3," in *2nd International Conference on Information Technology and Computer Application*, Guangzhou, China, 2020.

[81] K. W. a. G. J. Xia Dong, "Moving Object and Shadow Detection Based on RGB Color Space and Edge Ratio," in *IEEE 2nd International Conference on Image and Signal Processing*, Oct. 2009.

[82] "Google's Firebase Notes," [Online]. Available: https://firebase.google.com/docs/ml-kit/object-detection.

[83] A. Yilmaz, O. Javed and M. Shah, "Object tracking: A survey," *Acm computing surveys (CSUR), vol. 38, issue 4,* , pp. pp. 1-45, 2006.

[84] Piccardi and Massimo., "Background subtraction techniques: a review," in *IEEE, International Conference on Systems, Man and Cybernetics* , 2004.

[85] Lee, S. Uk, S. Y. Chung and R. H. Park, "A comparative performance study of several global thresholding techniques for segmentation," *Computer Vision, Grap.hies, and Image Processing, vol. 52, issue 2,* pp. pp 171-190, 1990.

[86] D. K. Yadav, "Detection of Moving Human in Vision based Smart Surveillance under Cluttered Background: An Application for IoT," in *From Visual Smveillance to Internet of Things: Technology and Applications*, Taylor & Francis, March 2019, pp. pp. 161-173.

[87] M. Yazdi and T. Bouwmans, "New Trends on Moving Object Detection in Video Images Captured by a moving Camera: A Survey," *Computer*

Science Review, Elsevier, pp. pp. 1-66, March 2018.

[88] L. Sharma, D. K. Yadav and A. Singh, "Fisher’s linear discriminant ratio based threshold for moving human detection in thermalvideo," *Elsevier, Infrared Physics and Technology*, 2016.

[89] Y. Youssef and M. Els.henawy, "Automatic Vehicle Counting and Tracking in Aerial Video Feeds using Cascade Region based Convolutional Neural Networks and Feature Pyramid Networks," *Transportation Research Record,* pp. pp. 1- 14, March 2021.

[90] S. Maqbool, M. Khan, J. Tahir, A. Jalil, A. Ali and J. Alunad, "Vehicle Detection, Tracking and Counting," in *IEEE, 3rd International Conference on Signal and Image Processing*, 2018..

[91] A. N. e. al., "Multimodal Features for Detection of Driver Stress and Fatigue: Review," in *IEEE Transactions on Intelligent Transportation Systems,* June 2021. .

[92] B.-G. Han, J. T. Lee, K.-T. Lim and Y. Chung, "Real-time license plate detection in high-resolution videos using fastest available cascade classifier and core patterns," *ETRI ., vol. 37, no. 2,* p. pp. 251–261, 2015.

[93] K.-J. Kim, P.-K. Kim, K.-T. Lim and Y.-S. Chung, "Vehicle color recognition via representative color region extraction and convolutional neural network,," in *10th nt. Conf. Ubiquitous Future etw. (ICUFN)*.

[94] Prisacariu, Victor Adrian, Radu Timofte and Karel , "Integrating object detection with 3D tracking towards a better driver assistance system," in *20th International Conference on Pattern Recognition*, 2010.

[95] Hu, Ling and Qiang Ni., "IoT-driven automated object detection algorithm for urban surveillance systems in smart cities," *IEEE Internet of Things Journal 5,* pp. 747-754., 2017.

[96] W. Wu, O. Bulan, E. Bernal and R. Loce, "Detection of Moving Violations. In Computer Vision and Imaging in Intelligent Transportation Systems," *IEEE Press: Hoboken, NJ, USA,* pp. Chapter 5, pp. 101–130., 2017.

[97] M. Al-qaness, A. Abbasi, H. Fan, R. Ibrahim, S. Alsamhi and A. Hawbani, "An improved YOLO-based road traffic monitoring system," pp. 103, 211–230. , 2021.

[98] T. Xu, Z. Zhang, X. Wu, L. Qi and Y. Han, "Recognition of lane-changing behaviour with machine learning methods at freeway off-ramps," *Phys. Stat. ech. ppl,* p. 567, 2021.

[99] D. Rosenbaum, F. Kurz, U. Thomas, S. Suri and P. Reinartz, *Towards automatic near real-time traffic monitoring with an airborne wide angle camera system,* Eur. Transp. Res. Rev. 2009, 1, 11–21. , 2009.

[100] M. Fachrie, "A Simple Vehicle Counting System Using Deep Learning with YOLOv3 Model," *Journal RESTI (RekayasaSistem Dan TeknologiInformasi)*, p. 462–468, 2020.

[101] H. Song, H. Liang, H. Li, Z. Dai and X. Yun, *Vision-based vehicle detection and counting system using deep learning in highway scenes,* Eur. Transp. Res. Rev. 2019, 2019.

[102] S. Awang, N. Azmi and M. Rahman, "Vehicle Type Classification Using an Enhanced Sparse-Filtered Convolutional Neural Network With Layer-Skipping Strategy," 2020.

[103] W. Sun, X. Zhang, X. He, Y. in and X. Zhang, "A Two-Stage Vehicle Type Recognition Method Combining the Most Effective Gabor Features," in *CMC-Comput. ater. Contin.* , 2020.

[104] Q. Kang, H. Zhao, D. Yang, H. Ahmed and M. ., "Lightweight convolutional neural network for vehicle recognition in thermal infrared images," *Infrared Phys. Technol. ,* p. 104, 2020.

[105] C. Zhang and B. Du, *Image-Based Approach for Parking-Spot Detection with Occlusion Handling,* J. Transp. Eng. Part Syst. 2020,146., 2020.

[106] P. T ˘tulea, F. C˘lin, R. Brad, L. Brâncovean and M. Greavu, *An Image Feature-Based Method for Parking Lot Occupancy.,* Future Internet 2019, 11, 169. , 2019.

[107] Wei, Jian, Jianhua He, Yi Zhou, Kai Chen and Zuoyin T, "Enhanced object detection with deep convolutional neural networks for advanced driving assistance," in *IEEE transactions on intelligent transportation systems 21,* , 2019.

[108] Srigrarom, Sutthiphong and Kim Hoe Chew, "Hybrid motion-based object detection for detecting and tracking of small and fast moving drones.," in *International Conference on Unmanned Aircraft Systems (ICUAS),* 2020.

[109] Q. L. B. a. J. S. Fan, "A closer look at Faster R-CNN for vehicle detection," in *IEEE intelligent vehicles symposium (IV)*, 2016.

[110] Mao, Qi-Chao, H.-M. Sun, Ling-QunZuo and R.-S. Jia, "Finding every car: a traffic surveillance multi-scale vehicle object detection method," *Applied Intelligence 50, no. 10 ,* pp. 3125-3136., 2020.

[111] D. K. Yadav and K. Singh, "Adaptive Background Modeling Technique for Moving Object Detection in Video under Dynamic Environment," *International Journal of Spatio-Temporal Data Science, Inderscience,* pp. vol. 1, no. 1, pp. 4-21, 2019.

[112] L. J, S. Z, G. J and Y. Y, "Real-time detection of steel strip surface defects based on improved YOLO detection network," in *IFAC-Pap 2018,* 2018.

[113] D. Lee, "Effective Gaussian mixture learning for video background

subtraction," in *IEEE Trans. Pattern Anal. Mach. Intell. 27 (5) (2005)* , 2005.

[114] D. Chaudhari and S. Porwal, *Frequency and spatial domains adaptive-based enhancement technique for thermal infrared images,* 2014.

[115] Z. Dordevic, S. Graovac and S. Mitrovic, "Suboptimal threshold estimation for detection of point-like objects in radar images," *Eurasip . Image Video Process., Springer Open Access .,* p. 1–12., 2015.

[116] A. Akula, N. Khanna, R. Ghosh, S. Kumar, A. Das and H. Sardana, *Adaptive contour based statistical background subtraction method for moving target detection in infrared video sequences, .,* Infrared Phys. Technol. 63 (2013) 103–109., 2013.

[117] Dong, X. Huang, Y. Zheng, S. Bai and W. Xu, "A novel infrared small moving target detection method based on tracking interest points under complicated background, .," *Infrared Phys. Technol.,* p. 36–42., 2014.

[118] Tao, Huanjie and X. Lu, *Automatic smoky vehicle detection from traffic surveillance video based on vehicle rear detection and multi-feature fusion,* 2019.

[119] Lu, Shengyu, B. Wang, H. Wang, L. Chen, M. Linjian and X. Zhang, "A real-time object detection algorithm for video.," in *Computers & Electrical Engineering 77* , 2019.

[120] Hong, Feng, C.-H. Lu, C. Liu, R.-R. Liu and J. Wei, "A traffic surveillance multi-scale vehicle detection object method base on encoder-decoder," in *IEEE Access 8* , 2020.

[121] Jain, Naman, ShreeshaYerragolla and T. Guha, "Performance Analysis of Object Detection and Tracking Algorithms for Traffic Surveillance Applications using Neural Networks," in *Third International conference on I-SMAC* , 2019.

[122] "Bounding Box Prediction from Scratch using PyTorch," [Online]. Available: https://towardsdatascience.com/bounding-box-prediction-from-scratch-using-pytorch-a8525da51ddc.

[123] S. Paris, P. Kornprobst, J. Tumblin and F. Durand, *A Gentle Introduction to Bilateral Filtering and its Applications,* https://people.csail.mit.edu/sparis/bf_course/course_notes.pdt.

[124] T. Bouwmans, A. Sobral, S. Javed and S. Jung, "Decomposition into Low-rank plus Additive Matrices for Background/Foreground Separation: A Review for a Comparative Evaluation with a Large-Scale Dataset," pp. Computer Science Review, vol. 23, pp.

[125] T. Yang, X. Wang, B. Yao, J. Li, Y. Zhang, Z. He and W. Duan, "Small Moving Vehicle Detection in a Satellite Video of an Urban Area", Sensors

(Basel),," pp. vol. 16, issue 9, 1528, pp. 1-15, Sept. 2016.

[126] S. Paris and F. Durand, "A Fast Approximation of the Bilateral Filter Using a Signal Processing Approach," *Computer Vision - ECCV, Lecture Notes in Computer Science, Springer, Berlin, Heidelberg,* pp. vol 3954, pp 568-580, 2006.

[127] R. C. Gonzales and R. E. Woods, *Digital image processing,* 4th Ed., Pearson Education, pp. l - 1026, 2018.

[128] J. M. McHugh, J. Konrad, V. Saligrama and P. M. Jodoin, "Foreground-adaptive background subtraction," in *IEEE Signal Processing Letters*, 2009.

[129] G. R, D. J, D. T and M. J, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2014.

[130] J. Choi, H. J. Chang, Y. J. Yoo and J. Y. Choi, "Robust moving object detection against fast illumination change," *Computer Vision and Image Understanding,* pp. pp. 179-193, 2012.

[131] K.-J. P.-K. K. Y.-S. C. a. D.-H. C. Kim, "Multi-scale detector for accurate vehicle detection in traffic surveillance data," in *IEEE Access 7 (2019)*, 2019.

[132] X. Ding and R. Yang, "Vehicle and Parking Space Detection Based on Improved YOLO Network Model," in *J. Phys. Conf. Ser. 2019*, 2019.

[133] B. Cai, R. Alvarez, M. Sit, F. Duarte and C. Ratti, "Deep Learning-Based Video System for Accurate and Real-Time Parking Measurement.," in *IEEE* , 2019.

[134] W. Sun, X. Zhang, S. Shi and X. He, *Vehicle classification approach ased on the combined texture and shape features with a compressive DL,* IET Intell. Transp. Syst. 1069–1077. , 2019.

[135] Z. Iftikhar, P. Dissanayake and P. Vial, "Computer Vision Based Traffic Monitoring System for Multi-track Freeways," *Intelligent Computing ethodologies; Springer International Publishing: Cham,* 2014.

[136] S. Alghyaline, N. El-Omari, R. Al-Khatib and H. R.-V. Al-Kharbshh, *An Efficient Real-Time Vehicle Counting Approach,* J. Theor. ppl. Inf. Technol. 2019, 97, 2062–2075., 2019.

9 Publications

Journal Papers (Scopus):

1. Sneha Mishra, Dileep Kumar Yadav., "Intelligent technique for moving object detection from problematic video captured through camera sensor", Recent Advances in Electrical & Electronic Engineering, pp. 1-12, May, 2023. (DOI: 10.2174/2215083810666230510113140) (Scopus)

(https://www.eurekaselect.com/article/131682?utm_source=TrendMD&utm_medium=cpc&utm_campaign=RecentAdvances_in_Electrical_%2526_Electronic_Engineering_TrendMD_0)

2. Sneha Mishra, Dileep Kumar Yadav., "Vehicle Detection in High Density Traffic Surveillance Data Using YOLO.v5", Recent Advances in Electrical & Electronic Engineering, pp. 1-14, May-2023. DOI: 10.2174/2352096516666230428103829. (Scopus)

(https://www.eurekaselect.com/article/131304?utm_source=TrendMD&utm_medium=cpc&utm_campaign=Recent_Advances_in_Electrical_%2526_Electronic_Engineering_TrendMD_0)

3. Sneha Mishra, Dileep Kumar Yadav "Detection of Moving Vehicle in Foggy Environment using Google's Firebase Platform", Turkish Online Journal of Qualitative Inquiry, Volume 12, Issue 6, July 2021. (ISSN: 9892-9921). (https://www.tojqi.net/index.php/journal/article/view/3518)

Conferences:

1. Sneha Mishra, Dileep Kumar Yadav, "Moving Object Detection in Video Captured by Static Camera: A Survey", International Conference on Networks and Cryptology. (NetCrypt 2020), JNU, New Delhi, 2020. (https://www.appleacademicpress.com/advanced-computer-science-applications-recent-trends-in-ai-machine-learning-and-network-security/9781774912393)

2. Sneha Mishra, Dileep Kumar Yadav, "A Rough Set Theory Based Edge Detection for Moving Object Detection", 2nd Asian Conference on Innovation in Technology (ASAINCON 2022), IEEE, PCCOER, Pune, India. (https://ieeexplore.ieee.org/document/9909108) (Scopus)

Book Chapters:

1. Sneha Mishra, Dileep Kumar Yadav, "Object Detection under Hazzy Environment for Real Time Application", CRC Press, Taylor and Francis, January-2021.

(Published, available online: https://www.routledge.com/Artificial-Intelligence-Technologies-Applications-and-Challenges/Sharma-Garg/p/book/9780367690809)

2. Sneha Mishra, Priya Porwal, Dileep Kumar Yadav, "Data Science and AI in e-Government and Society", Industry 4.0: Research Trends,

Challenges and Future of AI in Data Science, Under Book Series : Demystifying Technologies for Computational Excellence Moving Towards Society 5.0, CRC Press, Taylor and Francis, July-2020.

(Published, available online: https://www.amazon.in/Industry-4-0-Data-Science-Computational/dp/0367562855/ref=tmm_hrd_swatch_0?_encoding=UTF8&qid=1625814437&sr=8-1)