,

Bachelor of Science in Computer Science and Engineering

Developing a Framework for Vehicle Detection, Tracking and Classification in Traffic Video Surveillance

Rumi Saha
ID : 1504115

April, 2021

Department of Computer Science & Engineering
Chittagong University of Engineering & Technology
Chittagong-4349, Bangladesh.

Developing a Framework for Vehicle Detection, Tracking and Classification in Traffic Video Surveillance

This thesis is submitted in partial fulfillment of the requirement for the degree of Bachelor of Science in Computer Science & Engineering.

Rumi Saha

ID : 1504115

Supervised by
Professor Dr. Mohammad Shamsul Arefin
Professor
Department of Computer Science & Engineering (CSE)
Chittagong University of Engineering & Technology (CUET)

Department of Computer Science & Engineering
Chittagong University of Engineering & Technology
Chittagong-4349, Bangladesh.

The thesis titled **"Developing a Framework for Vehicle Detection, Tracking and Classification in Traffic Video Surveillance"** submitted by Roll No. 1504115, Session 2019-2020 has been accepted as satisfactory in fulfillment of the requirement for the degree of Bachelor of Science in Computer Science & Engineering (CSE) as B.Sc. Engineering to be awarded by the Chittagong University of Engineering & Technology (CUET).

Board of Examiners

1. ______________________ Chairman
Professor Dr. Mohammad Shamsul Arefin (Supervisor)
Professor
Department of Computer Science & Engineering (CSE)
Chittagong University of Engineering & Technology (CUET)

2. ______________________ Member
Professor Dr. Asaduzzaman (Ex-officio)
Head of the Department
Department of Computer Science & Engineering (CSE)
Chittagong University of Engineering & Technology (CUET)

3. ______________________ Member
Professor Dr.Abu Mohammad Ashfak Habib (External)
Department of Computer Science & Engineering (CSE)
Chittagong University of Engineering & Technology (CUET)

Statement of Originality

It is hereby declared that the contents of this project is original and any part of it has not been submitted elsewhere for the award if any degree or diploma.

Signature of the Supervisor
Date:

Signature of the Candidate
Date:

Acknowledgment

First and foremost, I would like to thank God Almighty for giving me the strength to finish this work. The satisfaction that accompanies the successful completion of this thesis would be incomplete without the mention of people whose ceaseless cooperation made it possible, whose constant guidance and encouragement crown all efforts with success. I am grateful to my honorable project Supervisor Professor Dr. Mohammad Shamsul Arefin and Professor Dr. Asaduzzaman honorable Head of the Department of Computer Science and Engineering, Chittagong University of Engineering and Technology, for the guidance, inspiration and constructive suggestions which were helpful in the preparation of this project. I would also like to extend my gratitude to all of my teachers for their valuable guidance in every step of mine during the four years of learning stage. Finally I would like to thank my friends, senior, juniors and the staffs of the department for their valuable suggestion and assistance that has helped in successful completion of the project.

Abstract

Sustainable transport and traffic management encompasses numerous fields such technologies, including intelligent transportation systems and vehicle assistance systems. The most difficult of these three movements is intelligent transportation system (ITS) and smart vehicle-class vehicle classification.Vehicle detection and tracking has various civilian and military uses, such as in the field of Urban Traffic and Control. In this study, we present a methodology for vehicle detection and classification based on camera surveillance. Several template- and image-matching techniques were explored for verifying authenticity evaluated for different classifier and feature combinations. This system focused on improving the detection, tracking, and then classifying, of single vehicles using the Histogram of Oriented Gradients (HOG) and also presented a method for improving the classifier with a feature derivative that's essential to class identification and object on the linear Support Vector Machine (SVM) classifier. We also affirm the robustly categorize vehicles on shape and dimension feature extraction with the real-categorizer Addaboost classifier, which has moderate accuracy and low cost. At the end of the analysis, Kalman Filters were used to measure and minimize the number of missing vehicles. We ran our proposed system through a number of different videos, and found that it returned the most valuable results. The results show the algorithm's effectiveness.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

An overview of vehicle detection, vehicle classification, and vehicle tracking will be discussed in this chapter. The background and current state of the issue are presented in this chapter. In addition, this chapter also touches on the principles and motivations of the study.

1.1 Introduction

ITSR (Intelligent Transportation System Research) and vehicle classification have caught the attention of many academic and industrial specialists alike Intelligent Transportation System (ITS). In this system, we came up with a new method to effectively and efficiently and comprehensively track and classify vehicles In addition, we evaluated a number of already available detection and tracking systems. It's recommended that you use personal footage for scenarios of motorway driving. We've conducted experimental research to prove the value and feasibility of this strategy.

The advent of vehicles has created significant traffic problems for people who are trying to get around the city. Problems arose as a result of the increase in traffic. Vehicle traffic has increased, which has created a greater risk to people's health due to congestion. An increasingly widespread application of traffic monitoring is needed to deal with traffic flow.

Because of the great risk that systems carry, the area of automated security systems is of huge importance at this time. The vehicle traffic monitoring component not only provides data on vehicle movement and traffic, but also recognizes objects, classifies them, and identifies anomalies, while the human component recognizes people and brings context to vehicle and traffic behavior interactions.

As significant as a traffic camera systems go, a traffic surveillance system is a necessary part of a smart transportation strategy. Automated video surveillance these days is critical for traffic management and safety purposes. You can get by with a very small video camera for a short time.

When we want to analyze video algorithms that don't require much human input Vehicle detection and background modelling are the driving capabilities of video surveillance systems. With video sensors and high-speed hardware for video processing getting more and more accessible, interesting problems like vehicle tracking and security are opened up. The main advantage of video cameras is that they are relatively inexpensive. It is often difficult and impractical to manually examine a large quantity of data while little or no human intervention is necessary for video analysis, algorithms requiring a minimal amount of human input are ideal. There are three main reasons for using video surveillance systems: background modeling, vehicle classification, and vehicle detection. Video sensors and hardware are making it easier to deal with video understanding issues like tracking and classifying targets. For example, accidents related to traffic flow of traffic, traffic jams, congestion, and the consequent emissions of pollution. The future of managing our cities' traffic flow lies in extracting information from our roads, which means reducing traffic problems such as congestion, pollution, and accidents. As a result, traffic surveillance work is on Intelligent Transportation system (ITS) is getting greater attention. Many dangers result from the automated video surveillance systems.

The goal of this project is vehicle classification from high-quality video footage. The research is planned to concentrate on finding and tracking vehicles in traffic sequence. Basing vehicles on object dimensions. It is a critically important and difficult task to scan objects in motion film. In order for the system to obtain accurate time information about objects, it must monitor them as well as temporal information. The first step in processing a video sequence is finding objects that are moving. Several components may be employed, such as video monitoring, traffic management, and time controls, among others.

1.2 Background and Motivation

The Intelligent Transportation System Research Field is the crucially important research today, especially in vehicle detection and classification (ITS). Reducing risk is a priority. Self-driving cars will be prevalent in the near future .The topmost priority for these vehicles is to be able to identify one another. Without independent surveillance, the household and industrial security systems will be worthless. Advancement has been made in this area in the past, but it was truly amazing in terms of with support vector machine (SVM) and hierarchical operators, we were able to obtain a robust vehicle detection, vehicle class classification, and vehicle tracking. Once the state information is obtained, the Kalman filter is applied. The system uses a proper algorithm to add a greater number of trackers while reducing computational complexity, thus increasing the detection and classification and analysis time.

1.3 Objectives

Every activity of research has a specific objective to achieve. Here we've tried to increase vehicle recognition, monitoring and classification accuracy in this thesis. The primary aim of detection, monitoring and classification of vehicles from the high quality video sequence.

To achieve the following objectives, the thesis is performed:

1. Detection and tracking of multiple moving vehicles in a video sequence.

2.Also classification of vehicles into their types on the basis of object dimension

1.4 Scope and Limitation

It will work well for surveillance purposes, this model can successfully be used in driving sequences that are changed from day to day and vehicle to vehicle. Since a vehicle has been shown to have sensors to detect a vehicle as well as motion, this system will be able to track vehicles while they are moving. The aforementioned model shall fulfill many promises from the models that have been broken.

1.5 Application of our proposed method

The established video-based vehicle detection system was used for advanced warning of congestion and queues in work areas and on freeways during special events. The advanced warning system is a collection of video monitoring stations with video recording equipment and our video tracking system. Before work zones or special event sites, vehicle queue lengths, speed and counts were tracked.

1.6 Challenges and Research Area

The complexity of a vehicle tracking system increases in places where motion, illumination, and poor weather conditions make it difficult to see or impair it in a rural setting, street setting. The higher precision demand adds a great deal of difficulty to the requirements of the job.

Sustainable transport and traffic management encompasses numerous fields such technologies, including intelligent transportation systems and vehicle assistance systems. There has been a sharp rise in the importance placed on road traffic monitoring in the context of the Intelligent Transportation System (ITS) over the past few years. Through the use of surveillance of both vehicular and pedestrian traffic, as well as human activity, it is possible to obtain a better view of the general scene as well as various statistics and classify details about it as well as well as the differentiating incidents, including identifying abnormal behavior, and accidents.

1.7 Organizing of the Thesis

The structure of this thesis can be outlined as follows:

Chapter 1 describes some introductory description of Vehicle detection and classification System

Chapter 2 gives an overview about related works and theory related to our proposed System

Chapter 3 discusses about the full process of the system

Chapter 4 describes the implementation of the developed the system.

Chapter 5 describes the experimental results and evaluation of the system.

Chapter 6 presents the conclusions.

Chapter 2

Literature Work and Review

In this chapter, section 2.1 describes the characteristics of vehicle. Section 2.2 describes related work which analyzed vehicle detection, tracking and classification system Section 2.3 describes the Histogram of gradient and 2.4 describes about support vector machine. Kalman filter and Haar Like feature describe in 2.5 and 2.6. Finally, this chapter is concluded in section 2.7.

2.1 Vehicle

2.1.1 Characteristics of Vehicle

The most important characteristic of the vehicle is its form regardless of which country or band it is.In different screens, the appearances will vary by pixel and unit of measure. Vehicle performance, gearing lines, vehicle lines, vehicle-geometry lines, and road-geometry are all included in the vehicle specs.

Figure 2.1: Example of vehicle

2.2 Related work

In the literature, a number of detection and tracking methods have been mentioned. Detection and tracking of targets is key to further investigation of vehicle movements. The vehicle detection system uses digital image processing to make distinctions such as differentiating a vehicle and its surroundings possible. Culture never follows policies; policies are made to conform to people. In a study on multiple-person tracking [11] an effective approach is discussed. Feature extraction was performed using the PCA and neural networks, according to Matthews et al. [12]. Can be referred to as a real-time traffic monitoring system. This uses a feature-based method in order to detect vehicles on crowded streets, combined with object-detection. Instead of looking at the entire vehicle features, the handling components are tracked however, it is computationally expensive. Tracking and locating vehicles was introduced by Cheng [1] who described a method of adaptive background learning for it. The image/video segmentation algorithm employs unsupervised analysis and spatial/temporal tracking to analyze the traffic video sequence. This work is both difficult and unsupervised, in the absence of the user and know it prior to training. Their paper proposed the use of three detection levels: raw images, as well as vehicle classification: at these levels: region, object classification and vehicle levels. Large amounts of both kinds of object and error are inevitable in this kind of work[2]. As a matter of fact, the experiment's findings show, highways fail on the streets where you have video sequences with just a few routes of vision.

For vehicle classification, Peng et al. [3] proposed a new method utilizing data mining approaches. Here license plate location and background subtraction were used to accurately locate the vehicle front. Additionally, in type recognition eigen-value vectors rather than an SVM were used to derive vehicle image probabilities. The greatest shortcoming of this method is that all vehicle license plates cannot be found.

Sun et al. [4] proposed a vehicle crash risk system that employs multicycle hypothesis generation and appearance-based verification techniques such as K-SVM for both initial detection and classification. The drawback of this project is that parameters failed when confronted with conditions, resulting in the error of an unknown environment.

Jayasudha et al.[5] presented a data mining overview in relation to road analysis [5] Data mining eliminates the deficiencies of other techniques but presents their advantages as well. Here they use a different road safety database, plus they must be able to choose a data discriminator, which compares a given model or set to a different set of targets Chouhan et al.[6] proposed a method to extract image features including size, texture, and dominant color for image retrieval with data mining and image processing methods. GLCM feature is used to quantify the visual complexity of an image. They use the Euclidean distance if the shape and texture are similar. The above feature extraction process works better only because of the fact that it is not as efficient as this one. The proposal [7] presented by Chen et al. utilized PCA for both feature extraction and classifiers - one is called Eigenvehicle, the other is called PC-SVM, and they compared their results to those of three vehicle classifiers on real-world, self-organized data by incorporating these methods to determine which worked best. Use of an unsupervised SPCM method in this case means you have to find the vehicle must supervised learning has its limitations. On-vehicle detection by Y. Zakaria et al. [9] exploited an SVM linear and non-linear classifier HOG feature scheme to train the system with different settings. For the training dataset, they used the KITI (kitchen thermal tissue transformation) dataset.

2.3 Histogram of Gradient (HOG)

HOG represents the gradient strength of a specified areas of an image. This has features in common with other methods for extracting information about objects. This is one of the major principles behind the HOG. The concept behind the HOG is that a discrete object's visual and physical appearance can be well shown in orientation using gradient patterns. Here are the steps to take: The first step in the de-convolution is breaking up the image into a connected group of small regions. After which, at the start of the gradient's orientation's calculation, in each cell, pick the histogram of all of all of the pixel data. Thus all features can be synthesized to get a single histogram descriptor.

2.3.1 Concept of HOG

Unfortunately, the feature was far less sensitive to light variations, to direction, and it wasn't especially able to tell how large or small an object was Also, the gradient histogram can be used to specify the particular gradient orientation within the layer, so that the first starting layer options can be verified easily.

At pixel (x,y) it was determined: magnitude and gradient of the function at (x,y):

$$\begin{cases} G(x,y) = \sqrt{G_x(x,y)^2 + G_y(x,y)^2} \\ \alpha(x,y) = \arctan \frac{G_x(x,y)}{G_y(x,y)} \end{cases} \tag{2.3.1}$$

Whereas Gx(x,y) is horizontal gradient held steady, Gy(x,y) is vertical gradient and (x,y) (direction of gradient) was variable

2.3.2 Advantages of HOG

There are a few important advantages to using the HOG descriptor. Since it is on single pixel samples, it is invariant to both photometric and geometric transformations. Transformations, so only drastic alterations to the architecture would have to be applied to larger locations. Moreover, given that pedestrians maintain an upright position, strong photometric sampling and coarse orientation only affect small amounts of the 3D data, only strong photometric normalization is required. Humans, on the other hand, are better able to detect changes in imagery thanks to the HOG descriptor. As well, block normalization will allow changes in illumination and shading to be better preserved.

2.3.3 Overview

The orientation of a pixel is determined by the direction of the slope of the gradient. The data is then organized into a predefined intervals (Figure 2.2) and feature vectors are calculated.

We do this by partitioning the image into cells and regions for feature extraction, Listed below are the details:

1. Take the image to 32 pixels and use a Guassian filter to smooth it.

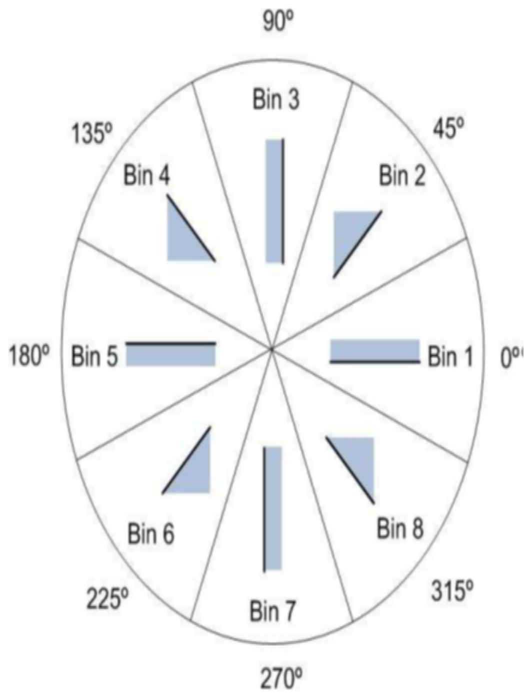

Figure 2.2: The 8 histogram bins for quantising the gradient orientations

2. Divide the final image into 16 sectors of 8x8 pixels in each

3.For each group of two-by-by-two group of adjacent cells, orthogonal cells, this method forms a block. The blocks in the diagram are all combined into nine groups.

4. 4. Horizontal and vertical gradients (dx and dy) for each pixel, I(x, y) in the block using the following equation:

dx = I(x + 1, y) I(x 1, y) (2.2)

dy = I(x, y + 1) I(x, y 1) (2.3)

5. Calculate the gradient orientation, $a_1 foreachpixel$:

$a_1(x, y) = 1/tan(dy/dx)$ (2.4)

6. Specify 8 histogram bins for orientation.

The feature vectors of each block are concatenated to form a HOG histogram with 8 values that covers the range from minimum to maximum intensity. This classifier's training vector is used by the classifier.

2.3.4 Variant of HOG Features

Several experimental designs are tested with HOG in the study. These approaches are similar in feature extraction, but have different feature values:

1. Number of angles of orientation

2. Number of bins of histogram

3. The pixel gradient calculation method Experiments determined the parameters which yielded the best classification results.

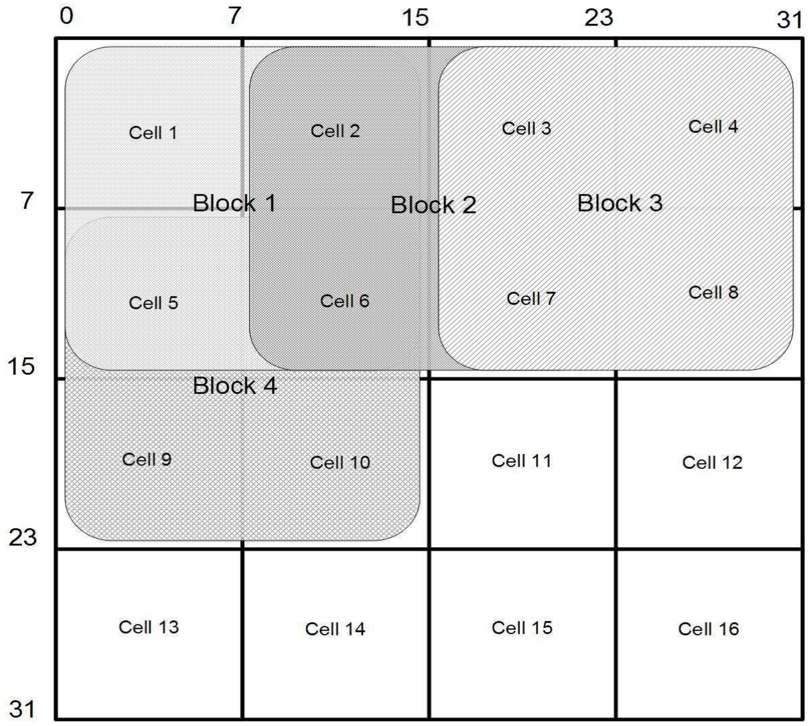

Figure 2.3: Dividing 32 x 32 in 16 cells and 9 blocks for extraction of HOG features (for clarity, only 4 blocks are drawn)

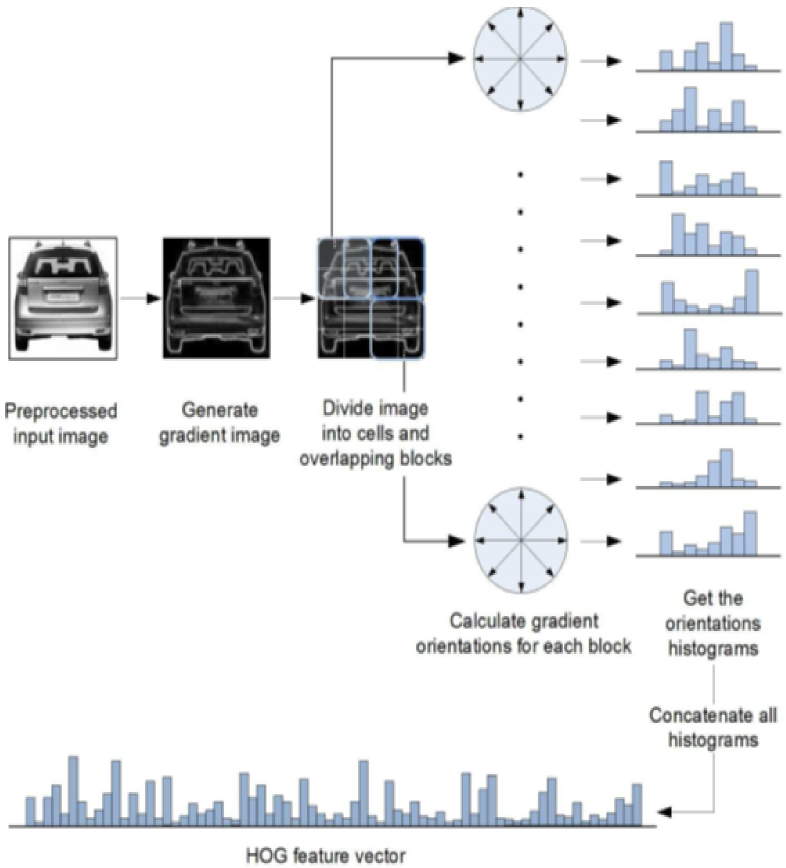

Figure 2.4: Displays the HOG function extraction summary

2.3.5 Histogram of Orientation Gradient Algorithm

Algorithm 1: HOG calculation

1: input I: a gradient orientation image

2: initialization H $\rightarrow$ 0
3 : for every positions (p, q) in the image do

4: i$\rightarrow$ $I(p, q)$
5 : k $\rightarrow$ including the small region (p, q)

6: for all offsets (x, y) to match neighbors do

7: if (p + x, q +y) contains of the image then

8: j$\rightarrow$ $I(p + x, q + y)$
9 : $H(k, I, j, x, y)H(k, I, j, x, y) + 1$

10: end if

11: end for

2.4 Support Vector Machine (SVM)

A supervised machine learning algorithm that is suitable for both classification and regression challenges is known as a support vector machine That is, however, most commonly in the task of classifying things. It plots each data point in n-dimensional space, with each feature having a specific value on a separate coordinate.

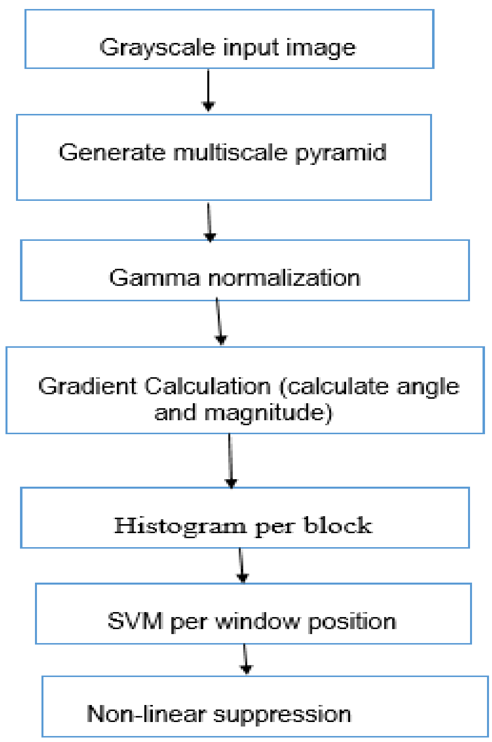

Figure 2.5: HOG feature extraction with classification procedure

2.4.1 Advantages of SVM

1. It looks good with wide white space around the image

2. In high-dimensional spaces, it has proven effective

3. It can be successful when there are more features to the model than samples.

4. This model uses a sub-sampled training function (called support vectors), so it is a form of data compression.

2.4.2 SVM Overview

The concept of decision planes is used in support vector machines. A decision plane partitions a collection of objects into separate class groups. The schematic diagram is found below. The objects in this scenario are either in the GREEN or the RED class. Using a dividing line, objects on the right side of the boundary are all GREEN, and those on the left are color-coded in RED. Should any new object (any white circle) land to the right be classed, that falls to the lower right, it should be GREEN (or classified as RED should it fall to the left of the separating line).

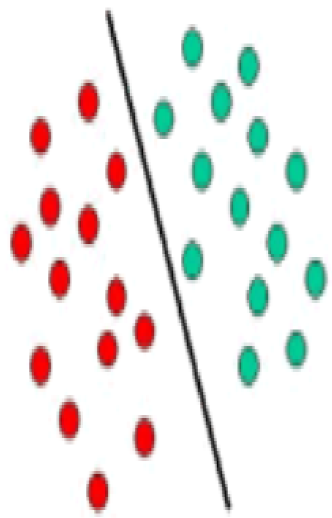

Figure 2.6: Decesion plane separator two classifier

In other words, a classifier that uses a single attribute (such as "color") to assign objects to their own group (i.e., Green and Red, in this case) Most of the time, a more sophisticated arrangement must be employed in order to make the optimal distinction (especially with new items), e.g., e. Exactly how new objects should be classified, e.g. (train cases).here is described in the picture. it is obvious to see that a more complicated curve is needed to achieve a complete separation of the GREEN and RED objects (which is more complex than a line). hyperplane classifiers are excellent at classifying things based on drawn lines Fusion Splits are very well to deal with problems of this type.

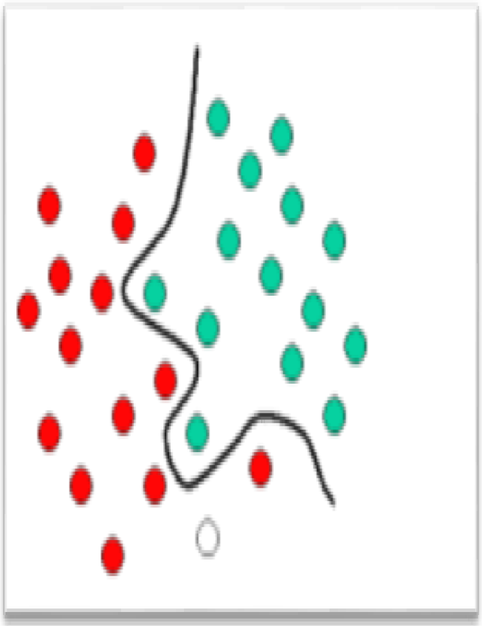

Figure 2.7: Complex structure of classifier

Below(Fig. 2.7) is a basic illustration of how Support Vector Machines work. functions are being used on their original position, e.g., their original objects (i.e., like to say, "mapped, but shifted") have new objects mapped, for example, rearranged as, and that is known as 'transformation functions'. There is an objective way of rearranging the objects is to determine (transformation). In this new environment, all the mapped objects are straight lines, so complex curves (on the left) can be replaced by simple, and thus it's a simpler task to find an optimal path to find a straight line that joins the GREEN and the RED sets.

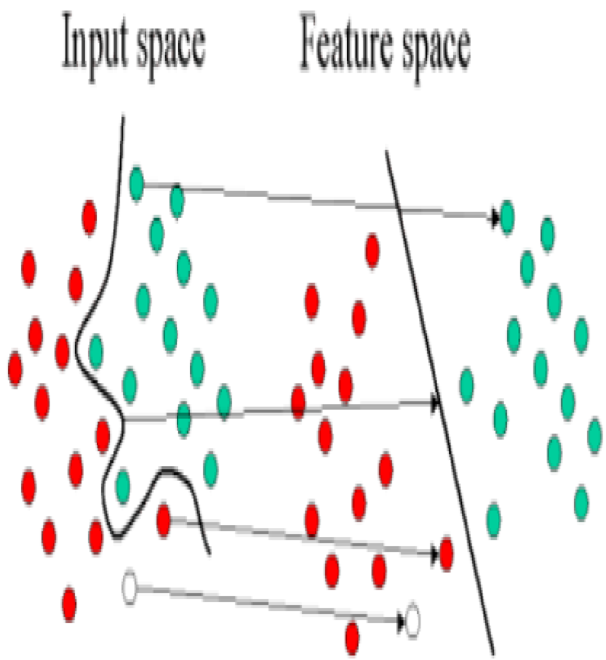

Figure 2.8: Solve complex structure by kernels

2.5 Kalman Filter

To deal with the moving objects, a Kalman filter is used. Also, in addition, it will enhance auto tracking, and lower frame rate losses. For a short time, the filter estimates and obtains feedback by measuring the noise from the cycle.

The Kalman filter is based on two equations: one that generates the state of the system, and another that keeps track of it. The state equation and the observed equation follow one another because the state of the State follows the process of being given.

$$K_k = p_{k;k\ -1}.H_k.(H_k.P_{k-1}H_k + R_k)^{-1} \tag{2.5.1}$$

It is the vector of observance of the moment k; it represents the state transition matrix from the moment k − 1 to k; it means the observer matrix of the moment k −1; it represents the observational matrix; it stands for system noise.

The Kalman filtering process consists of two steps: Prediction phase in which the next condition is predicted by reason of previous measurements and the update stage when the system state is estimated at the point of time.
The prediction and estimation calculated as:

$$\hat{x_k} = \phi\hat{x} - k - 1 \tag{2.5.2}$$

$$\hat{x_k} = \tilde{X_k} + K_k(Z_K - H\tilde{x_k}) \tag{2.5.3}$$

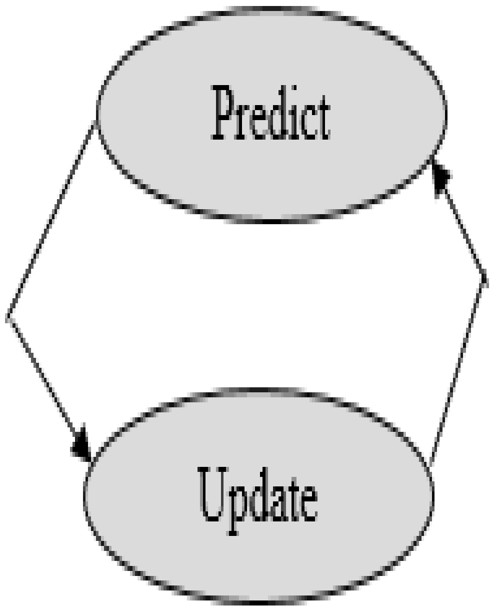

Figure 2.9: Kalman filter working procedure

2.5.1 Application of Kalman-filter

Kalman Filtering's applications are different in the real world. An example request would provide accurate, constantly modern information on the location and speed of an object, each of which includes errors and only a sequence of observations regarding its position. In a radar application where one is interested in tracking a target, information on location, speed and acceleration is being measured each time instantly, to provide a similar, more concrete example with a lot of noise degradation. In order to remove the effects of noise and obtain a good estimate of the position of the goal at the present time (filters, future times (predictions) or in the past, the Kalman filter exploited the dynamic of the target, which govern its time (interpolation or smoothing). Weather forecasting, language improvements etc. are other applications

2.5.2 Kalman filter Algorithm

Algorithm 2: Kalman filter Tracking

1: **If** (time==0) (

2: Consider a new track for any detected vehicle

3:)

4: **Else**

5: For (All tracks current)

6: Kalman-filter predicts new position and track size

7: Overlap new location found with all blobs

8: **If** (Overlap (track (a), blob (b))! =0)

9: Mark match-matrix [a] [b];

10: Analyze (match-matrix);

2.6 Haar like feature

Digital image attributes in object recognition are hair-like characteristics. They owe their name to the intuitive likeness of hair waves and were used in the first face detector in real time. In order to improve the recognition of objects within images, this was used to increase the dimensionality of the feature gathering. This has been popular because some of these features will improve the definition of the object. For instance, a hair-like tilt in two rectangles may indicate an edge at 45.

2.6.1 Haar like feature based cascade classifier

The "cascade" time period within the classifier call means that the resulting classifier consists of several less challenging classifiers (levels), sooner or later applied to a proximity until at any point the candidate is refused or the level exceeds all levels. The word "reinforced" method means that the classifiers are themselves complex at any point in the cascade, and that one of the four distinct stimulus strategies (weighted vote), can be constructed from easy classification systems (weighted voting). Discrete Addaboost currently backed, Real Addaboost, Gentle Addaboost and logitboost. The basic classifiers are decision-tree classifiers with at least 2 sheets. Haar like functions are entered into the basic classification devices and measured as described below. The modern set of rules uses hair-like characteristics:

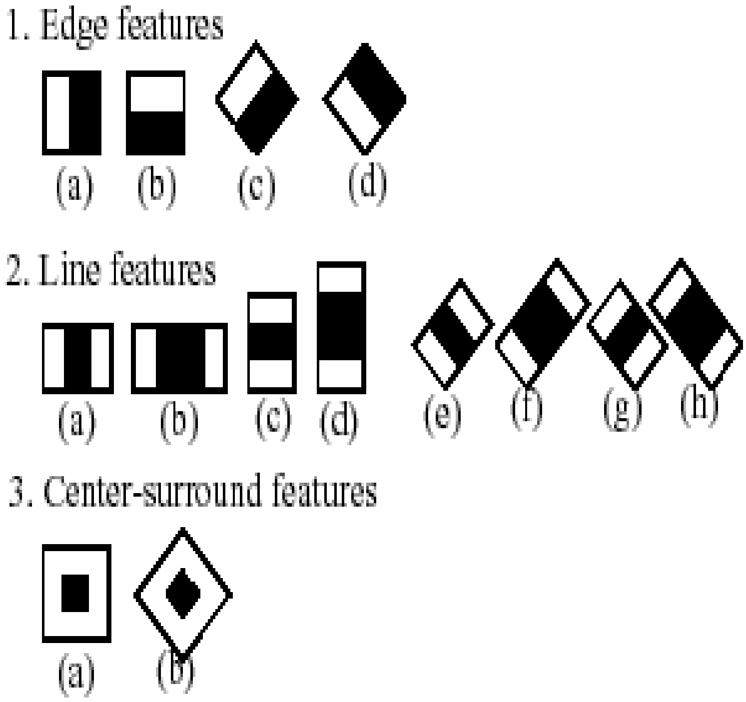

Figure 2.10: Different haar like features

2.7 Conclusion

This chapter is dedicated to basic concept of the Histogram of Oriented Gradient (HOG) and Support Vector Machine (SVM), Kalman Filter and Haar like feature.And also described the lietrature work related to our proposed system.

Chapter 3

System Architecture and Design

This chapter describes my research work's working principle. The flow chart diagram includes work steps to understand the method clearly.

3.1 Proposed Methodology

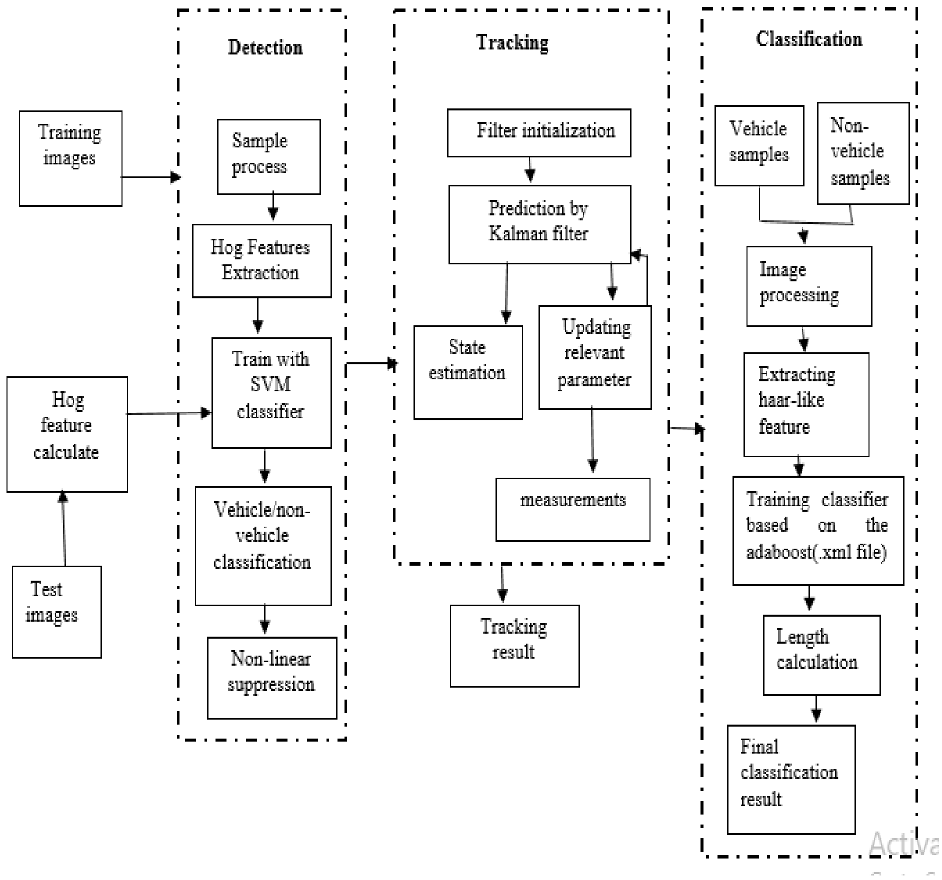

Figure 3.1: System Architecture of the proposed framework

As we discuss previous related tasks we will take steps to detect, track and classify vehicles. We note that the entire procedure is limited, which is why the Intelligent Transportation System cannot provide a very efficient framework. We proposed an efficient framework of that field in Figure 1, which could reduce the limits of the background work, to address all the limitations of that method. The whole vehicle detection and tracking procedure was shown with algorithm 3. A few basic modules are included in the automotive detection and tracking system architecture: initialization dataset, extraction of features, classification and tracking. The dataset initialization function is to create a database with positive and negative samples. The extraction module is important for the generation of positive samples. The module for the classification of separate and non-vehicle samples is for training. The tracking module detects all moving vehicles and reduces the number of objects that are missing.

3.2 Dataset

The algorithms presented in different scenes during the day, including roads, urban streets, narrow streets, etc., are collected for evaluation. In the first stage 7,325,samples including 3,425 vehicle (positive samples) and 3,900 non-vehicle samples have been collected for training and testing (negative samples). Vehicles include different types of vehicles, like cars, trucks, buses, and different colours, like red, white and blue. Furthermore, the car samples include both vehicles near and far from the camera. Traffic signs etc. at both stages. The example of vehicle and non-vehicle picture training is illustrated in figure 3.2 and figure 3.3. In 64 RGB images, the training samples were re-dimensioned.

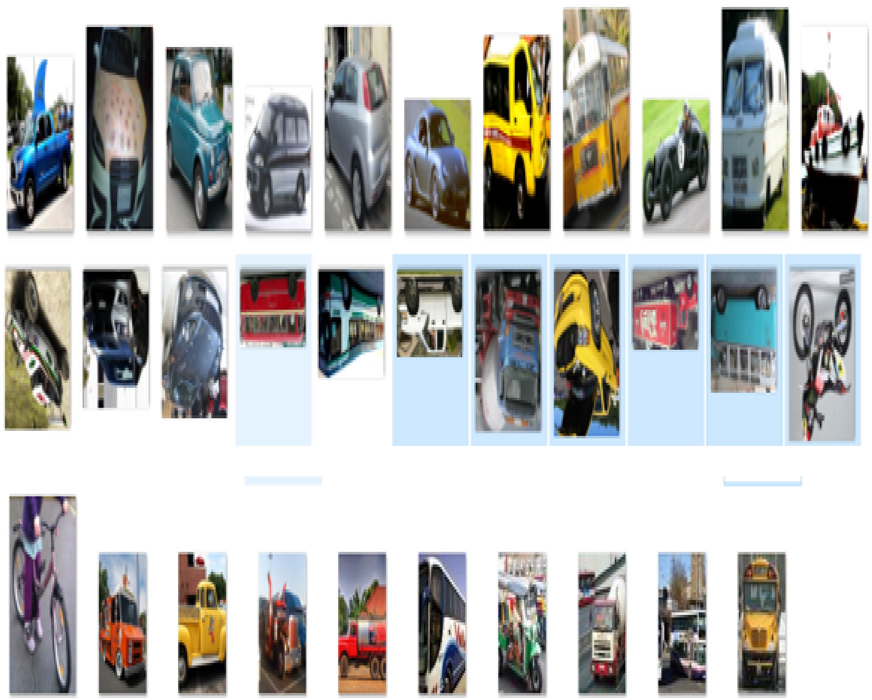

Figure 3.2: Samples for vehicle images

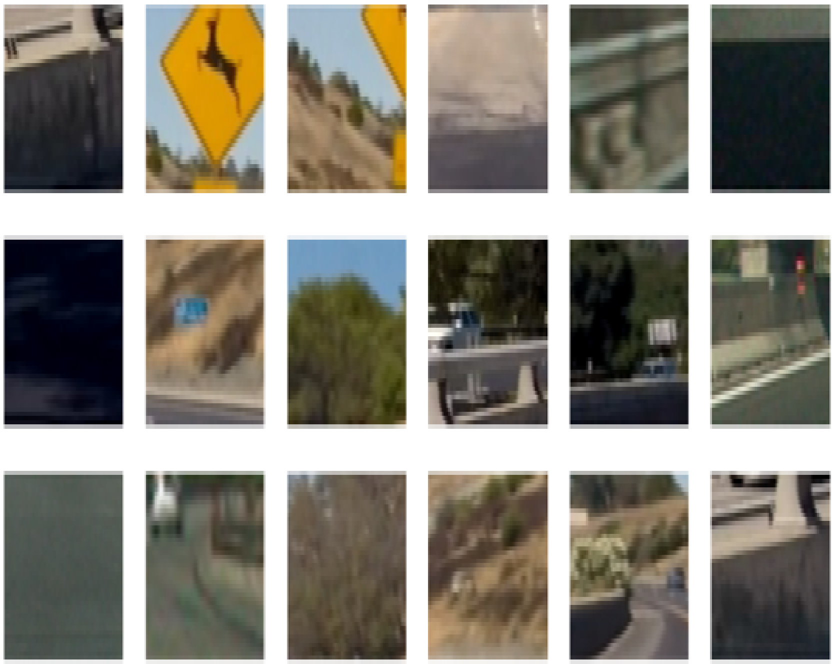

Figure 3.3: Samples for non-vehicle images

3.3 Sample Process

Before extracting the feature, the collected samples are then preprocessed. The sample has been modified to a size of 64 pixels per 64 pixels in height and 64 pixels in width. The height is selected based on the original sample, so that the sample retains the average aspect ratio, the width is determined. The next step is to calculate the different changes in the HOG function for a sample, so that each sample has the same vector length, the cell size changes for each sample size.

3.4 Histogram of Oriented Gradient (HOG) Feature Extraction

The principal components are the extraction of vehicle detection Histogram of Oriented Gradients (HOG). The original HOG calculation takes 5 steps. First of all, the image undergoes color normalization and gamma correction. The image is divided into a grid of the cell. The cells are grouped in larger blocks which make it possible for the cells to be part of more than one block. Example of division into 16x16 cells where each cell contains 256 pixels and 2x2 blocks, which indicate 2 cells for each block. The blocks in the figure have an overlap ratio of 50%, with half of the block cells shared with the next block. Cell sizing and block size are parameters to be determined by the user based on the image size and the quantity of information to collect.

In the object detection, the key element of HOG is extraction are follows:

The images are converted to color input in grayscale. The gamma correction procedure is used to standardize the (normalized) color space of the image input; the purpose was to adjust the image contrast, decrease the shadow of the local image and the impact of the lighting changes. It can, however, reduce noise interference. Measure the gradient mainly in order to gather contour details and reduce the interference with light. Cell steering gradient of the project. Every cell within a block is normalized; normalization has the light, the shade and the corners compressed, and after normalization the descriptor of the block is called the HOG descriptor. Pick up HOG features in the detector space from all blocks; this step is to gather and merge HOG features into final classification vectors to overlap detector blocks.

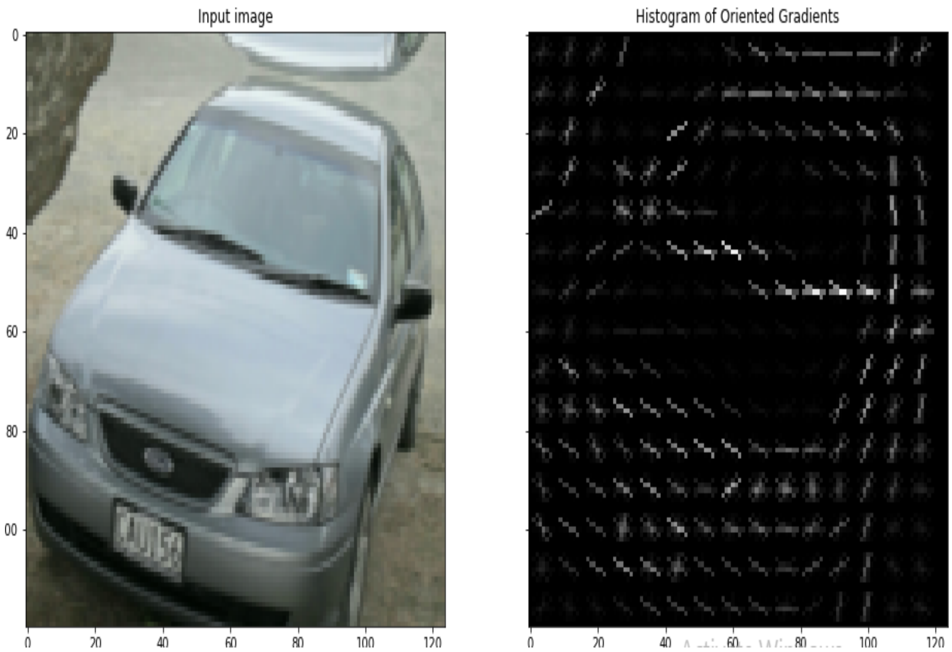

Figure 3.4: Vehicle image HOG feature

This figure shows extraction feature 3.4 3.5 HOG from the image of the vehicle and non-vehicle..

The HOG variants use the same technique for extracting the function but changes in the following parameters:

1. Number of guidance angles.
2. Number of histogram containers.
3. The method used to determine the pixel gradient.

We may also use color transformations to attach binned color and histograms to your HOG function vector optionally.

3.5 Support Vector Machine (SVM) classifier

Vector support machine (SVM) is a type of structural risk theory-based algorithm. If our results contain exactly two classes, we can use a Support Vector Machine (SVM) for binary classification. A SVM categorizes data in order to look for the best hyperplane in one class for

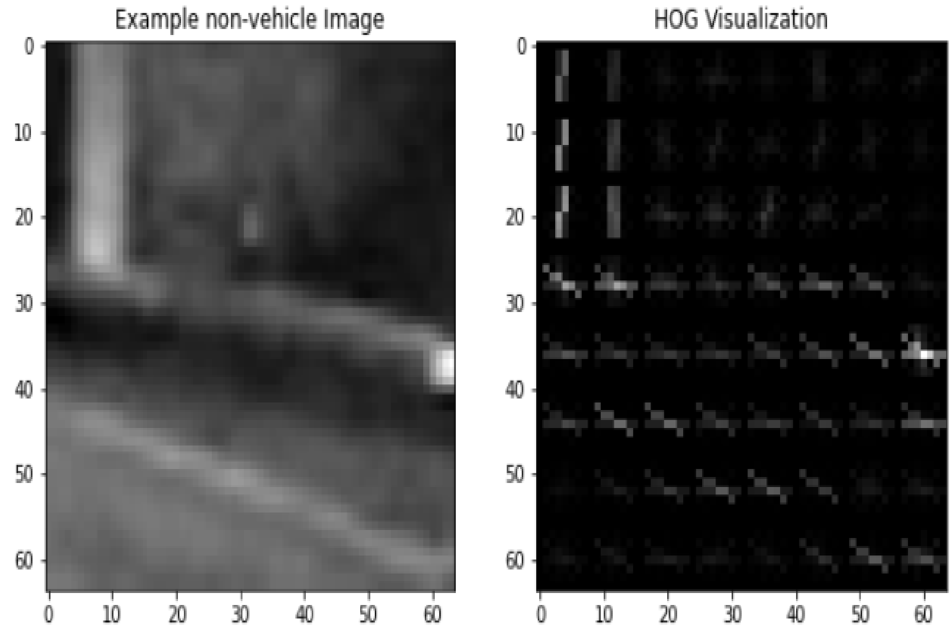

Figure 3.5: Non-Vehicle image Hog feature

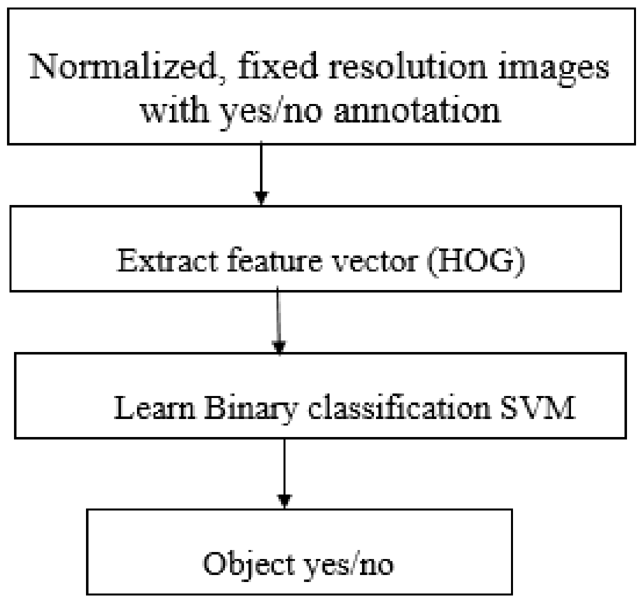

Figure 3.6: Train with the classifier

all data points from the other. The best hyperplane for an SVM is the one with the biggest gap.

With the highest interval, two classes are divided into two classes. SVM and non-linear classification are designed to recognize artifacts in a linear classification. A Gaussian kernel has been used as a non-linear SVM kernel. Perform a dense multi-scale scan with the classifier at each location of the test image and reveal preliminary object determinations. HOG features combined with the SVM classification were commonly used in image recognition, especially when object detection was highly efficient. The purpose is to differentiate between the two groups by a feature that is induced from the available examples as separate vehicles and non-vehicles. By analyzing the accuracy with the other classification, the accuracy of HOG-SVM was shown to be greater than the other classification .

3.6 Non-linear Suppression

A Non-linear deletion is carried out to minimize the overlapping of detected windows into one window, while a deletion process is carried out with a percentage overlap of greater than a certain threshold on each window according to the confidence of the classifier window with the highest confidence level.

3.7 Kalman Filter for Tracking

The next step is to track vehicles for objects moving, from picture to picture. We are interested to contrast an algorithm with a lower tracking complexity because of the relatively high cost of the detector. That's why we choose to use the Kalman Filter tracking algorithm to track our system objects. The system can help to increase the accuracy of auto tracking and decrease framework losses. Typically, the Kalman filter has two sections. The first is forecasting and updating.

***Prediction* :** The Kalman filters are used in the subsequent frame to predict the position. The speed of the vehicle is measured on the distance of the blobs. The current frames of the position of the vehicle are calculated using the speed, the current frame location and the time from the last frame. It depends on the state and serves to predict and error co varrience the state vector.

Calculating vehicle positions: We use a heuristic approach in which every patch of a vehicle is moved around to cover as many vehicle-related blobs as possible. This is the actual location of the vehicle.

***Estimation* :**Measurement in the picture co-ordinates system, as determined above. It updates the prediction parameters to reduce the error of the predicted and measured position of the vehicle.

The next step is to correct the prediction stage error by forecasting the location of the vehicle in the next frame. For the correction step function, we use the function. We now add two color constraints and the corresponding point size characteristics. The first condition is a blob contrast in two color frames in a row. The first condition is a blob contrast in two color frames in a row. In the frame n, the blob must be identical compared to the color of the blobs. In two frames, the second condition is about blob size. As the distance between the camera and the vehicle varies, the vehicle's size remains always different. However, from one frame to the next, the size variance ratio is minimal. In one consecutive frame, two blobs must be low. So, when two blobs match the colors and the difference in size between them is less than the threshold, and the predicted choice fulfills these terms, we consider that the prediction is true. In that case the criteria are met for each vehicle to be selected as an acceptable choice if two vehicles are similar to each other. .

3.8 Haar –like feature Extraction

Many methods of machine learning benefit from being calculated more effectively. Haar Classification is a decision-based technology in the course of the training process which creates a statistically improved rejection cascade. First, a grader train with a few hundred positive samples of the same size (namely the cascade of boosted graders with hair-like characteristics) is sized, and negative samples can also size the random images as positive pictures. Listed these pictures in a text file and created a vector after resizing (using OpenCV creates sample). Then the samples begin training with a cascade classification and built-in classification (.xml file). Through the training of the cascade-boosted classifier, we can extract hair like function.

3.9 Load Classifier

Training in a standard classifier is the main task of the work done in our last section. We load the cascading classifier (.xml file) and train you to give positive pictures in accordance with our vehicle types..

3.10 Length Calculation

The most important training task in a standard classification is the work done in our previous field. We now load our.xml file for positive photos of your vehicle type, Cascade Classifier.. The form based extraction of the object is a matter of great necessity in its dimensions, such as height, width, length etc. By using the following formula the amount and width of every vehicle detected are measured:

$$D = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{3.10.1}$$

3.11 Classification

Recently there has been much research interest in the classification of vehicle types. It is based on the shape or appearance of the vehicle in two principal directions. The classification is achieved by classifying cars by three groups, namely, large, medium and small, according to the size of vehicles. Because it is easy to find the vector length, the length was taken as parameter for the classification of vehicles by the given size. When any vehicle object comes to our territory of interest, it is calculated by means of the length calculation equation that the vehicle has been classified for this vehicle and then the length requirements have been met.

3.12 Algorithm for Detection and Tracking

Algorithm 3: Vehicle Detection and Tracking

Vehicle Type	Vehicles Name
Large	Bus, Truck
Medium	Car, Van
Small	Motorbike, Bicycle

Figure 3.7: vehicle type classification

Input: The video sequence is being monitored by the target specified
Output: Tracking result
for t→ 1 to N do
if 5 seconds has been detected, compared with the last detection then
HOG and SVM to detect the vehicle;
If objects of concern are found (ROI) Then
the Kalman filter tracking algorithm tracks an entity in the bounding box
if 10 seconds has been detected
Comparing with the last detection then Normal SVM and HOG to judge whether Object of a vehicle in boundary box still
else
Return to tracking the kalman filter tracking
end
else
Return to the last step for detection
end
else
Return to the last step for detection
end

Chapter 4

Software Development

4.1 Introduction

A developed vehicle detection-, tracking-und classification system is the primary goal of this thesis. Some picture processing concept is required to develop this technology. The software from Python is very useful for image processing tasks. The entire program is structured with python functions, because it is easy to handle and has a large number of features. On a machine with the window 10, 2.50 Core i5-4100 with a 4GB RAM system the developed vehicle detection, tracking and classification system has been implemented. In python 3.6.7 the system was developed (version).

4.2 Tools used for development

The tools are used for our system completion

1• Towards the OS

2.The software for Python

3• The digital camera or the camera based.

Image processing is therefore the way to analyze and manipulate a digital image primarily in order to increase its quality or to extract data from it that can be used.

4.3 Python image manipulation tools

Image Processing is therefore the method for analyzing and handling a digital image mainly designed to improve its quality or to extract data from it which can then be used. Illustration and feature extraction, image restoration, and image recognition are popular tasks for picture-processing, such as croppings, flip-outs, spinning and so on. Python is an acceptable alternative to such image processing tasks. This is because of its increased popularity as a science programming languages and the fact that several State Image Processing Tools are available in its ecosystem.

4.3.1 Cloud Server

In AI research, Google is rather aggressive. Google has developed over many years a TensorFlow IA framework and a Colaboratory development tool. TensorFlow is currently open and

Google has freed Colaboratory for public use. It is open source. Google Colab is known as the collaborator. The following can be done with Google Colab as a programmer:

A. Python Write and Run Code

B. Document your mathematical equations supporting code

C. Establish/Upload/Share notebooks

D. Google Drive import/save notebooks

E. External datasets from Kaggle are imported, for example

F. PyTorch, TensorFlow, Keras, OpenCV integrate

G. Free Cloud Free GPU service

4.3.2 Scikit Image

Scikit-image is an open source Python package that covers numerical arrays. It implements algorithms and utilities for research, education and business applications. It is also a fairly simple and straightforward library for those who are new to the Python ecosystem.

4.3.3 Numpy

Numpy is one of the main libraries for Python programming and supports arrays. An image is essentially a standard Numpy array that contains pixels of data points. We can then modify the image's pixel values through simple NumPy operations such as slicing, masking and fantastic indexing. You can load and view the image with skimage by matplotlib.

4.3.4 Scipy

The package is currently included with functions of linear and non-linear filtering, binary morphology, interpolation of the B-spline and entity measurement. Scipy is another of python's principal scientific modules, like Numpy, which can be used to manipulate and process images in fundamental tasks. Special features of the scipy and image submodule are NumPy's n-dimensional arrays.

4.3.5 PIL

PIL (Python Imaging Library) is a free Python language programming library which supports opening, manipulating and saving of multiple image file formats. However, with its last release in 2009, its growth remained stagnant. Fortunately it works for all major operating systems and supports Python 3, which is an actively developed PIL fork that's easier to install.

4.3.6 Opencv-python

OpenCV is one of the most popular libraries for computer vision applications (Open Source Computer Vision Library). OpenCV is the python API for OpenCV. OpenCV-Python is not only fast because the context includes code written to C/C++, but also easy to code and implement (due to the Python wrapper in foreground).

4.4 Image Formats Supported by PYTHON

The following image format are supported by PYTHON

A. BMP-Windows Bitmap

B. JPEG-Joint Photographic Experts Group

C. PNG-Portable Network Graphics

D. GIF-Graphics Interchange Format

E. PNM-Portable Any Map

F. .pcx -PC Paintbrush Bitmap Graphic

G. .ras - Sun raster graphic

The only saving format that is supported for this is PNG, other than PIL, which can be written to any format that is supported.

4.5 Conclusion

Here you will find the software used to implement the method and relate software functions. The results and analysis part of the system was described in the following chapter. We will try to test and analyze the approach for various samples.

Chapter 5

Experimental Result and Analysis

This section presents the implementation and performance assessment of our development framework with the analysis of our experimental result. First, we analyze our methods and why they have been used and graphics are shown. This chapter includes an input-output sample and a comparative analysis.

5.1 Implementation

Fig. 5.1 only showed the results if there is a false detection position. But after the tracking algorithm in Fig. 5.2, 5.3 and 5.4 is applied and multiple trucks can be traceable, the false positive rate may be reduced. In this bounding red color are the detection boxes, blue boxes are the tracking boxes.

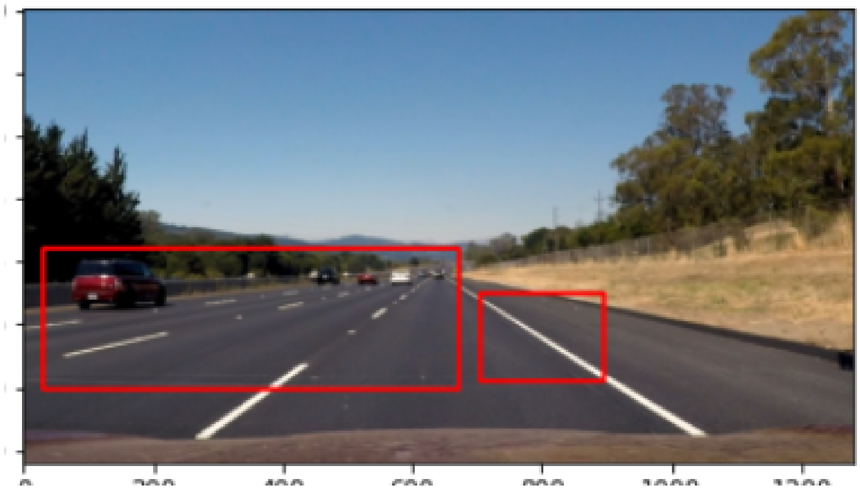

Figure 5.1: Results only for detection

Figure 5.2: Tracking urban road with high light

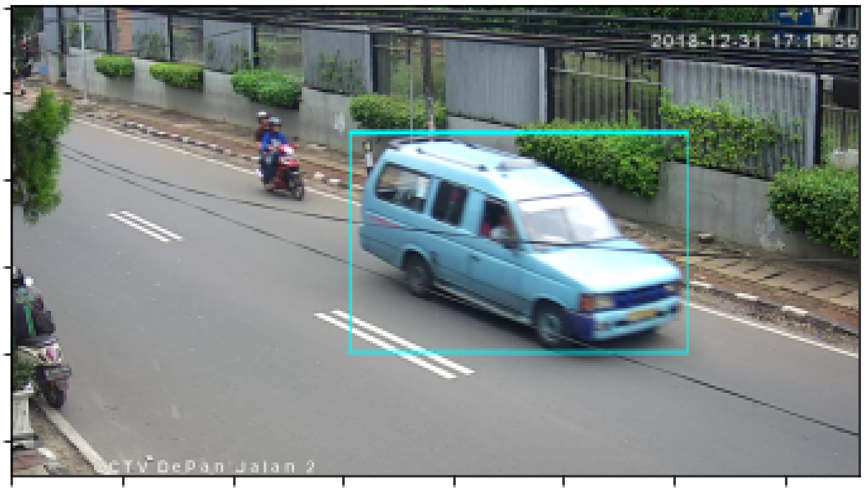

Figure 5.3: Tracking in rural congested road

Figure 5.4: Tracking with low light.

The Fig. 5 Example of vehicles detection and tracking of various scenes such as (5.2) high-light urban road, (5.3) rural congested road, (5.4) low-light vehicle.

Tracking vehicles in the city is much harder, as objects are close by each other and tress or any other background may shade the road as well as the vehicles. In the scenario where our suggested system works very successfully, some tracking results have shown in Fig. 5.(2-4). The control of the kalman-Filter based tracking system is carried out by using a different time/scenes. We can see that both videos and images have a good tracking system. It can track moving objects quickly as well. A new tracking object will be detected when a vehicle enters the scene, a new number and a tracking window for the new vehicle will be configured .

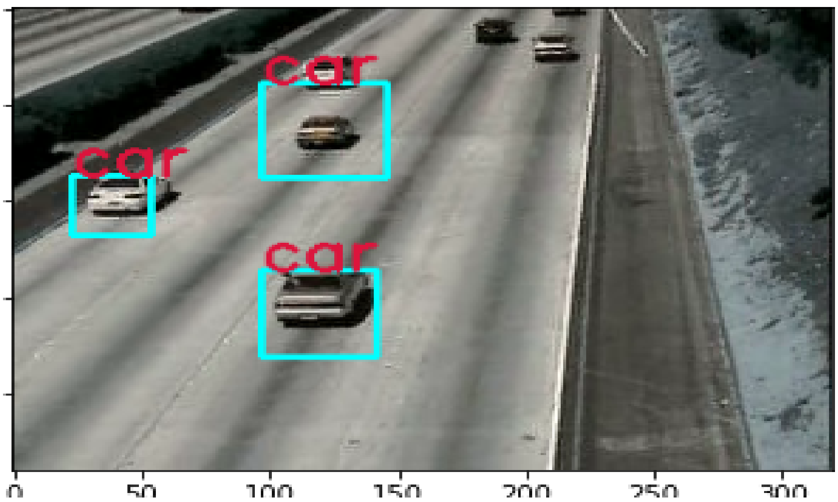

Figure 5.5: Car type of vehicle

On the highway image sequences we tested the system. Most vehicles can be monitored and classified successfully through the system. A variety of videos and pictures have been chosen to

Figure 5.6: Bus type of vehicle

test various vehicle types, including three types of cars, buses and motorcycles. By connecting algorithms based on car extract characteristics or dimensions to haar-cascade-based classification, we loop them through and draw a rectangle around each frame of the videos. After that, the framework that can automatically extract and categorize the vehicles into various traffic video scenes has been integrated..

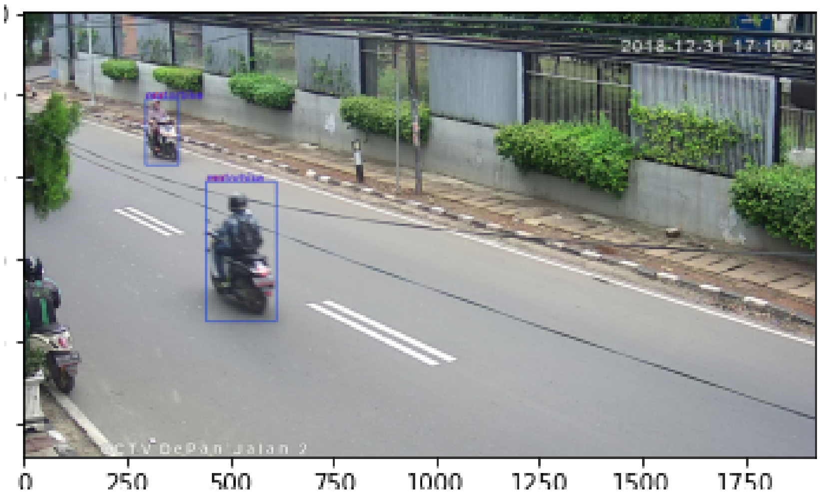

Figure 5.7: Motorbike types of vehicle

Fig. 5.5, 5.6 and 5.7 Example of vehicle classification in various types, such as 5.7 vehicle motorbike type, 5.5 vehicle car type, 5.6 vehicle type bus.

5.2 Performance Evaluation

We precisely break the training data train and test it with 20% of vehicle images and 20% of sub-images for vehicle training to assess its efficiency. to assess the proposed solution. For testing as a data set we used a fixed set of 231 subframes of vehicles and non-vehicles.

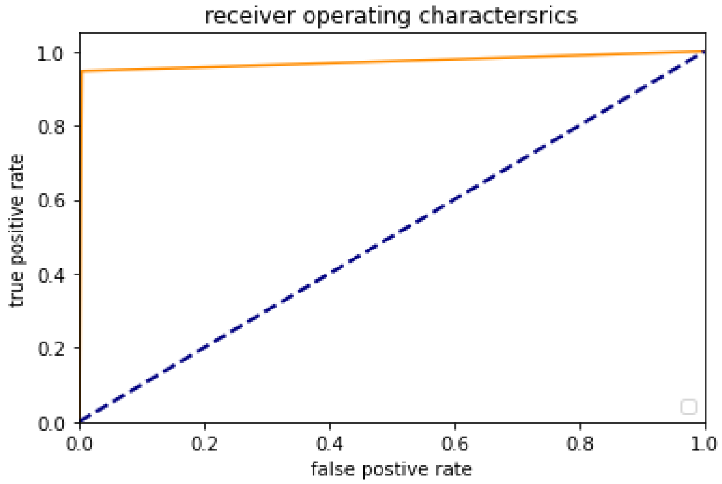

Figure 5.8: Receiver operating characteristics curve

The ROC curve is used to assess our model and check how good or poor it is. It helps to view the performance of our classification system and is confined to the measurement of binary classification issues. The true positive rate on the X axis is mapped against the real negative value on the Y-axis by the probably curve. False positive rates show the Y-axis at different values which can distinguish between positives and negatives only correctly. The curve shows that sensitivity is highest and precision is lowest at the point where both false positive and true positive rates converge. This means all positive class points are classified correctly and all negative class points are wrongly classified. All pins above blue lines are indicative of a condition in which the proportion in the positive class of properly classified points exceeds the proportion in the negative class of erroneously classified points.

The number of points that are detected and tracked properly by the proposed method is correct correspondence. And the total number of match points is the term for correspondence.

$$Accuracy = \frac{Number - of - corrected}{Number - of - correspondence} * 100\% \tag{5.2.1}$$

5.2.1 Performance of Kalman Filter Tracking

Parameter	Video1	Video2
Height of frame	310	7250
Width of frame	310	1270
Length of video	13s	25s
Number of frames	253	915
Execution Time	2.4minutes	5.9 minutes

Table 5.1: Vehicle detection and tracking execution time analysis

The videos we tested in the traffic sequence for the TABLE 5.1 results showed information such as the height and width of these videos, number of frames, total working time. From this result, we can understand the accuracy and effectiveness of our proposed system

As shown in TABLE 5.2, we use traffic sequences of video to check the Kalman Filter-based tracking system, with the tracking results as shown above. We see a good tracking result for

Test Videos	Tracking Vehicles	Actual vehicle	Accuracy
video1	62	68	91%
video2	35	36	97%
video3	5	5	100%
video4	17	17	100%
video5	22	25	88%
video6	12	12	100%
Average accuracy			96.00%

Table 5.2: Performance Analysis for videos

the system and can track moving objects like moving vehicles as well. When a new vehicle enters the scene, it shows a new Tracking Object, distributes a new number and initializes a Tracking in the vehicle.

This has a strong detection and tracking effect for moving destinations such as vehicles quickly. This detects and tracks vehicles that join the scenario randomly. As we assume that good tracking performance depends on good monitoring, we detect more computer complexity. The method is also a mainstream of the current trend of research called detection by tracking.
.

5.2.2 Comparison with other Existing Method

Method	Accuracy	False rate
PCA + NN	85%	7.5%
Wavelet+ Gabor + NN	89%	6.2%
Gabor + SVM	94.5%	5.4%
HOG + SVM + Mean - shift	94%	3.7%
HOG + PCA + SVM + EKF	95%	3.5%
Our proposed method	95.25%	3.7%

Table 5.3: Comparison with some existing method with our proposed method.

The comparison of other vehicle detection and tracking methods existing with our proposed methods has been demonstrated in TABLE 5.3. From where we see that wavelet feature extraction [17] has achieved greater accuracy than another method with another dimension reduction feature, but the false rate also increases. Other existing HOG [16] methods are less precise than those of other tracking systems with SVM and Mean Shift Tracker, and this method has been unsubscribed when the weather is bad and traffic scenes congested. However, the use of Kalman filter trackers with a mean shift tracker increases its precision and decreases the false rate.Due to the exactness, we can see that the mean detection and tracking accuracy is 95.25%. The wrong rate of our method being proposed is also 3.4% lower. As the amount of video increases, it is difficult to track

5.2.3 Performance of classification

The right grading rate and 15 fps frame rate were achieved by 90 percent. Classification accuracy is also an important part of the classification of the machine learning classification. The number of vehicles, divided into the total number of vehicles, is specified correctly. As the success of the project will depend on a correct line for the view of the camera, the camera had to be placed directly on top of the road flow over an overhead bridge to reduce vehicle occlusion. Certain results of our system can be seen in Fig. 5(5-7).

The main reason for the classification errors was the slight divide between vehicle classes. As we only use the scale as our metric, it is not possible to properly categorize all vehicles. We need to explore more features to further improve our classification rate success. On a day different time our information was collected. We intend to further test the system in the more complex scenes and in a wider variety of conditions of lighting and weather. We are measuring the number and classification rate for each class and, as shown in 5.4, the final result for all vehicle types.

Vehicle Types	Number of vehicles	Number of classified vehicles	Success rate
Bus	7	6	85.7%
Car	6	6	100%
Motorbike	9	8	88.8%
Total	22	20	90.90%

Table 5.4: Performance analysis for classification result vehicle

5.3 Conclusion

Here, our detection, tracking and classification system performance was determined. We used the image, video testing and compares our method proposed with existing systems in our experiment. We have evaluated the precision of our proposed framework by using the exacting equation.

Chapter 6

Conclusion

6.1 Conclusion and Future Work

Robust and accurate vehicle detection, tracking and classification of frames collected by a moving vehicle is a significant issue for autonomous self-guided vehicle applications. This paper provides a summary study of the techniques proposed that were used in traffic video and images with a very high accuracy of tracking and classification rate. In this article we have done an extensive analysis of the latest literature on computer vision technology used in the surveillance and monitoring of traffic on a video basis.

The idea of using histograms of gradients to extract characteristics from various scales and orientation is essential to our approach. The filter Kalman tracks several video objects. A Haar like feature function with a cascade classifier which categorizes vehicle on the basis of the threshold value has been classified. The aim of the presented approach is to reduce the time for calculation. Experimental results show that the system is precise and efficient and that the system can achieve a high performance rate for vehicles.

Bibliography

[1]. C. Zhang, S. Chen, M. Shy, S. Peeta, "Adaptive Background Learning for Vehicle Detection and Spatio- Temporal Tracking" 0-7803-8185-8/03/17.000$IEEE, pp.797 - 801, 2003.$

[2]. S. Gupte, O. Masud, R. F. K. Martin, "Detection and Classification of vehicles" TRANSACTION ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 3, NO. 1, IEEE, pp. 37-47, 2002.

[3]. Y. Peng, J. S. Jin, S. Luo, M. Xu, Sherlock Au, Z. Zhang, and Yue Cui, "Vehicle Type Classification Using Data Mining Techniques" J.S. Jin et al., The Era of Interactive Media, DOI 10.1007/978-1-4614-3501-$3_27, Springer Science + Business Media, LLC, pp.325 - 335, 2013.$

[4]. Z. Sun, R. Miller, G. Bebis and D. DiMeo, "A Real-time Precrash Vehicle Detection System", Procceding on the sixth workshop on Application Computer Vison(WACV'02)IEEE , pp. 253-259 , 2002.

[5]. K. Jayasudha Dr. C. Chandrasekar, "AN OVERVIEW OF DATA MINING IN ROAD TRAFFIC AND ACCIDENT ANALYSIS" Journal of Computer Applications, Vol – II, No.4, pp. 32-37, 2009.

[6]. P. Chouhan, M. Tiwari, "Image Retrieval Using Data Mining and Image Processing Techniques" DOI 10.17148/IJIREEICE.31212, pp. 53-58, 2015.

[7]. C. Zhang, X. Chen, W. Chen, "A PCA-based Vehicle Classification Framework" Proceedings of the 22nd International Conference on Data Engineering Workshops (ICDEW'06) 0-7695-2571-7/06 20.00$IEEE, pp.451 - 458, 2006.$

[8]. Z. Sun, G. Bebis and R. Mille, "ON-ROAD VEHICLE DETECTION USING GABOR FILTERS AND SUPPORT VECTOR MACHINES" 7803-7S03-3/02/17.00 02002 IEEE, pp. 1019-1022, 2002 .

[9]. Y. Zakaria, H. E. Abd El Munim, M Ghoneima, S. Hammad , "Modified HOG based on-road Vehicle Detection Method" International Journal of Pure and Applied Mathematics, Volume 118 No. 18, pp. 3277-3285, 2018

[10]. Q. Tian, L. Zhang and Y. Wei, "Vehicle Detection and Tracking at Night in Video Surveillance ", iJOE – Volume 9, Special Issue 6: "AIAIP2012", July 2013, pp.60-64.

[11]. H. Tao, H.S. Sawhney and R. Kumar, (1999) "A sampling algorithm for tracking multiple objects", in Vision Algorithms.

[12]. N. Matlhews, P. An, D. Chamley and C. Hank, "Vehicle Detection and Recognition in Greyscale imagery," Cvnlrol Engineering Pmcrice vol. 4, pp.473479, 1996

[13]. H. Cho, P. E. Rybski and W. Zhang, "Vision-based Bicycle Detection and Tracking using a Deformable Part Model and an EKF Algorithm"

[14]. J. Kong, M. Qiu and K. Zhang, "Authoring Multimedia Documents Through Grammatical Specification," IEEE ICME 2003, Baltimore, USA, pp.629-632, 2003.

[15]. M. Qiu, G. Song, J. Kong and K. Zhang, "Spatial Graph Grammars for Web Information Transformation," IEEE Symposium on Visual/Multimedia Languages (VL), Auckland, New Zealand, pp.84-91,2003.

[16]. W. Xu, M. Qiu, Z. Chen, "Intelligent Vehicle Detection and Tracking for Highway Driving," IEEE International Conference on Multimedia and Expo Workshops, DOI 10.1109/ICMEW.2012.19, pp.67-72, 2012.

[17]. Mrs N.G. Chitaliya, Prof. A.I. Trivedi, “Feature-Extraction using Wavelet-PCA and Neural Network for Application of Object Classification and Face Recognition,” 2010 Second International Conference on Computer Engineering and Applications, IEEE DOI 10.1109/ICCEA.2010.104,pp.51
514,2010.