EE514/CS353: Machine Learning Project Report

Session: Spring 2022

Muhammad Haris 2021-03-0008

Syed Arham Mukhtar 2021-03-0009

Department of Electrical Engineering

Syed Babar Ali School of Science and Engineering

Lahore University of Management Sciences

Pakistan

Abstract

The aim of this paper is to develop a classifier that detects vehicles in a video stream. The paper will be applying different feature extracting methods including HOG, Colour Histogram and Spatial Binning to extract objects from the image/video before feeding it to the classifier. The classifier will be trained using different machine learning algorithms such as Support Vector Machines, Naïve Bayes, Logistic Regression and Neural Network. The purpose of training different classifiers is to evaluate the performance of different machine learning algorithms under some defined metric. A lot of work in vehicle detection has already been done in the literature; however, this paper will serve as a survey, providing a holistic comparison of different classifiers under different feature extracting techniques.

Contents

List of Figures

Vehicle Detection and Classification from Images

Chapter 1

1.1 Introduction

According to the World Health Organization (WHO), approximately 1.3 million people die each year due to road accidents. Governments and authorities have tried different methods to solve traffic problems, but no approach has survived the test of time. However, in recent years, proposed technological solutions to solve traffic related issues have gained significant attention due to the overwhelming success the technology has received in different fields. A self driving car is once such solution.

Vehicle object detection is an inherent part in such a solution. The system working in the self driving car must accurately detect vehicles for it to function successfully. Thus, vehicle detection has become a hot topic of research among researchers, and day after day, one sees new solutions and approaches in the field of object detection. This paper is also a small effort in this regard.

Like already mentioned, there are new technological approaches being developed day in day out, but machine learning and deep learning solutions have achieved significant success. In machine learning, different algorithms have been proposed to train classifiers to accurately detect and track moving vehicles. Support Vector Machines, Naive Bays, Neural Networks and

Logistic Regressions are just some of the methods which have gained significant success. This paper will look into these techniques in detail and will provide a comparative view in terms of some predefined goodness measures.

In supervised machine learning, a model is trained on labeled data which would then be evaluated on test data points before being used for real time applications. In the vehicle detection case, the data points would be labeled images containing different vehicles. However, one cannot just simply provide images to the classifier as it would take ages to process the images. Thus, this paper will use computer vision techniques to extract features from the data (images). The feature extraction techniques would help the system detect objects in the images. During the training phase, the detected objects would be labeled as vehicles; however, after the training phase, the classifier would need to predict the detected objects as vehicles (or not). HOG, Colour Histogram and Spatial binning are just some of the feature extraction techniques that this paper will be using to extract features from the images.

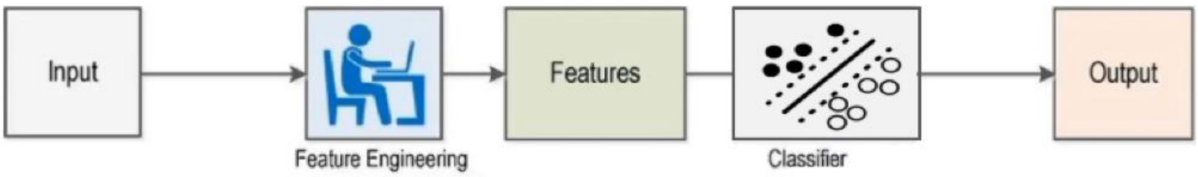


Fig1 : Project description

1.1.2 Methods used for vehicle detection

1.1.2.1 Support Vector Machines (SVM)

SVM is a linear model used for classification and regression problems. It constructs a hyper plane as maximum margin classifier which divides the data into two classes. Following is the mathematical formulation of the SVM.

Consider d-dimensional n data points ($x = [x^1, x^2 \dots x^d]$) with binary label $y \in \{-1,1\}$:

For training data $D = \{(\mathbf{x_1}, y_1), (\mathbf{x_2}, y_2), \ldots, (\mathbf{x_n}, y_n)\} \subseteq \mathcal{X}^d \times \mathcal{Y}$

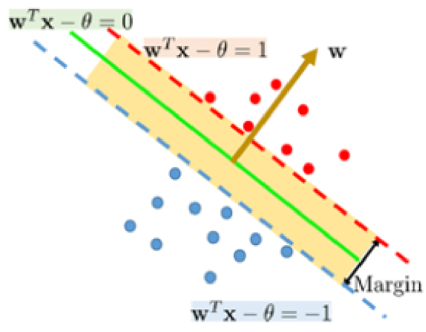
$\mathbf{w}^T\mathbf{x}_i - \theta \geq 1$ if $y_i = 1$ (Plus plane)

$\mathbf{w}^T\mathbf{x}_i - \theta \leq -1$ if $y_i = -1$ (Minus plane)

$\Rightarrow y_i(\mathbf{w}^T\mathbf{x}_i - \theta) \geq 1$

This can be formulated as a following optimization problem.

$$\underset{\mathbf{w},\theta}{\text{minimize}} \quad \|\mathbf{w}\|^2 = \mathbf{w}^T\mathbf{w}$$
$$\text{subject to} \quad y_i(\mathbf{w}^T\mathbf{x}_i - \theta) \geq 1 \quad i = 1, 2, \ldots, n$$



1.1.2.2 Naive Bayes

Naive Bayes classification uses Bayes' theorem to estimate (classify) the label(y) from the given data(x). It is called naive because it is based on a very strong assumption that features are independent of each other. Nevertheless, the classifier based on Naive Bayes performs extremely accurately for some classification problems. If x= [x^1, x^2…] represents a feature vector with class label y, then estimation of y given x is:

$$P(y \mid \mathbf{x}) = \prod_{i=1}^{d} P(x^{(i)} \mid y)\, P(y)$$

1.1.2.2 Logistic Regression

Logistic regression is a discriminative classifier which estimates y given x directly from the data. The name regression might be misleading as this algorithm does not fit a linear model in the feature space to predict a class; rather, it computes the probability of instance being in that class.

1.1.2.3 Neural Networks

Neural network in machine learning is a mathematic modeling of the biological neural system. Given a d-dimensional feature space (as described above) with binary classification, a neural network can be described in its simple form by the following depiction:

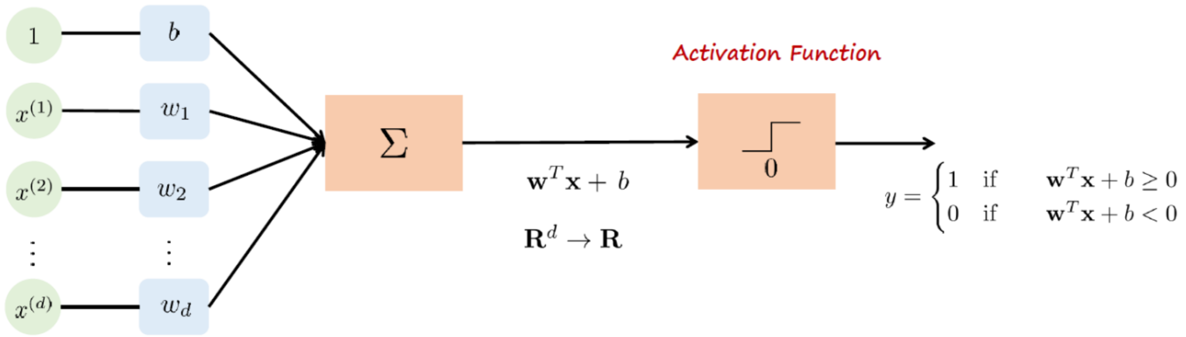


Fig 1a: Perceptron model

1.1.2.4 Sliding Window Technique

To detect a car in a test video frame, one has to start by picking a sliding window which is then fed as an input to the trained classifier. A positive classifier output means that the desired object is detected in a particular window. The process is repeated by sliding a window over every part of the video frame. For each window, the classifier outputs whether it has a vehicle or not. However, there is one caveat. One has to run sliding window several times while trying different window size each time, hoping to fit a window size over the object so that the classifier could detect it. This makes the computational cost of the technique really high. Thus, the researchers have come up with a better alternative approach: YOLO. YOLO uses neural networks to provide real-time object detection which makes the algorithm really fast.

1.2Mathematical Formulation

The basic mathematical modeling is done above.

1.3Data Processing

Data will be processed as per the needs. For HOG, the images will need to be converted to grayscale. For spatial binning, flattening and reshaping will have to be done. Moreover, to make things fast, resizing of images might be required. These details will be edited later once the team is done with the preprocessing part.

1.4Feature Engineering

1.4.1 Feature Extraction Techniques

1.4.1.1 Histogram of Oriented Gradients (HOG)

One cannot just simply use image intensities for classification because such a classifier will not function well. Thus, the research community has developed the concept of oriented gradients to extract features from the images upon which classification could be achieved. The HOG is a famous computer vision and image processing technique used for the purpose of object detection through feature extraction. The HOG technique counts the occurrences of gradient orientations in localized regions of an image. For each of these localized regions, the method would generate a histogram separately based on the gradients and orientations of the pixel values. The resultant HOG feature vector is a combination of all pixel level histograms which is then used with a classifier to classify images.

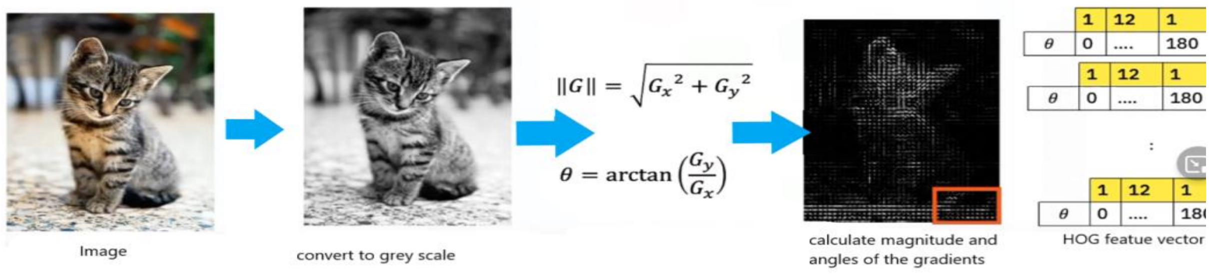


Fig 2: HOG

Below is the HOG feature visualization of the images belonging to the project dataset.

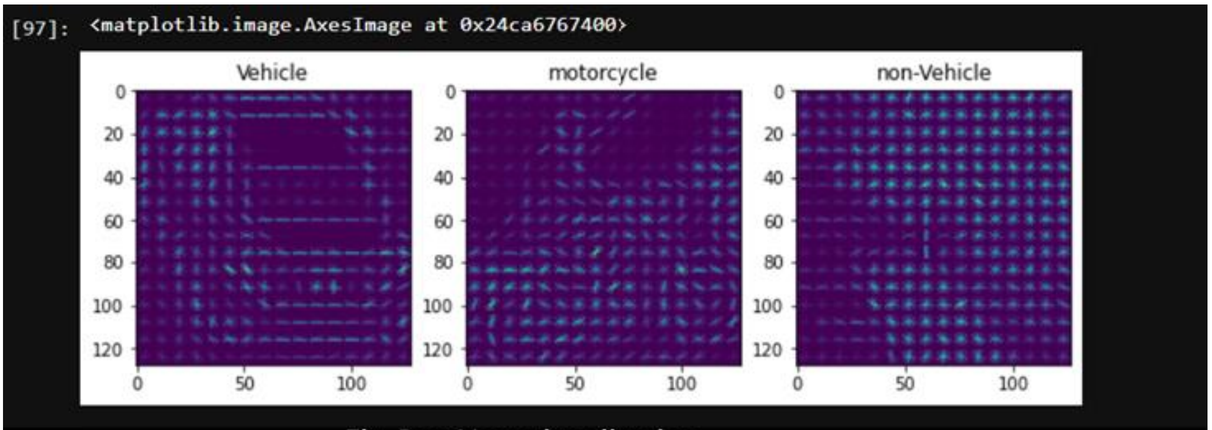


Fig 2a : Hog visualisation

1.4.1.2 Colour Histograms

Another simple but effective technique is to extract features through various colour schemes in an image. This can be achieved by plotting colour histograms. A color histogram is a depiction of the distribution of colours in an image.

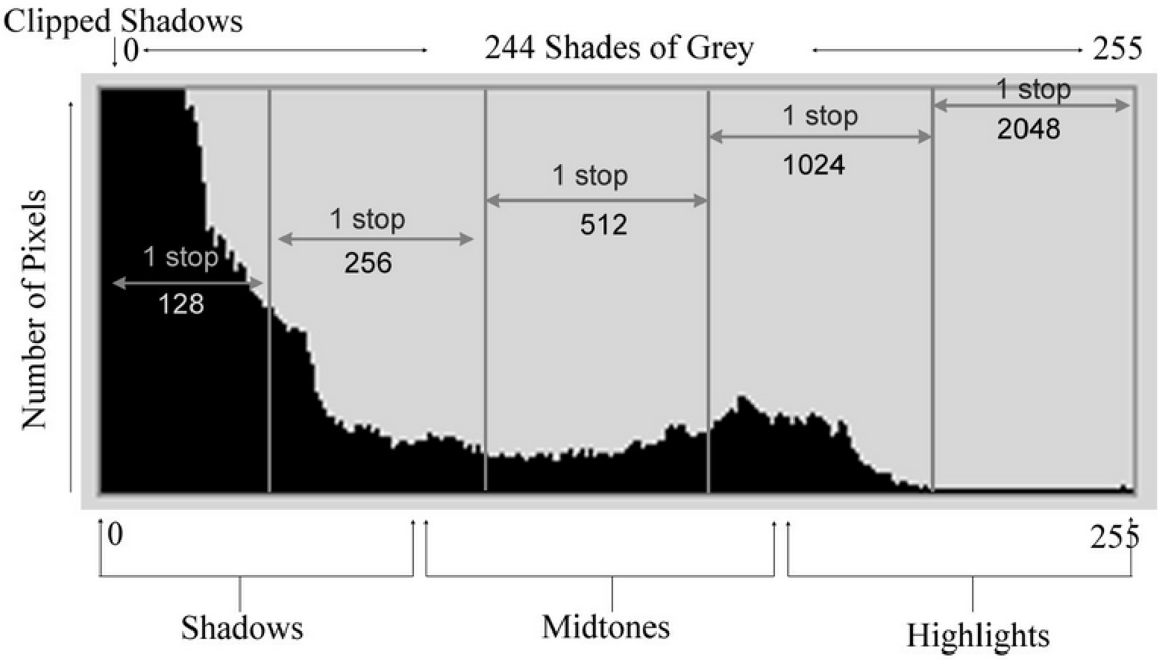


Fig 3: A colour histogram depicting light and dark shades in a particular image

1.4.1.3 Spatial Binning

In image processing, combining a cluster of pixels into a single pixel is called binning. For example, by binning 2x2 image, an array of four pixels becomes a one large pixel. A brief intuition behind spatial binning is summarized in the image below.

Fig 4: Intuition behind binning

1.5Implementation

The implementation will consist of six main stages. These are: loading the dataset, preprocessing, feature extraction, designing the classifier, training the classifier, evaluating the trained model.

After the model is trained, sliding window technique is used to allow the classifier to detect vehicles in a video frame. Sliding window technique permits the detection of localized objects by allowing the subset of an image data to be used in classification. This is extended to a video stream by allowing sliding window to be run on each frame to help detect the localized vehicle.

Following are some of the examples of a localized detection done on each frame of a video stream.

Fig 5: Sliding Window

1.6Performance Evaluation

Following table depicts the evaluation of different classifiers.

Classifier	Accuracy (%)
SVM	98.9538
Logitic regression	99.1385
Neural Network	98.8308
Naive Bayes	95.5385

Fig 6: Evaluation

1.7Conclusion

Vehicle detection is talk of the town as the world is moving towards self driving cars. Moreover, vehicle detection is an important step in machine learning based solutions aimed towards improving traffic surveillance. A lot of research has already been done in vehicle detection; however, this paper has provided a holistic approach to different vehicle detection techniques.