

# Vehicle Classification using Density based Multi-feature Approach in Support Vector Machine Classifier

Pradeep Kumar Mishra  
Ph.D Research Scholar Vision Image Processing Lab  
Department of Electrical Engineering, Indian Institute of Technology Bombay  
Mumbai 400076  
pkmishraiitb@ee.iitb.ac.in

Bi-lab Banerjee  
Ph.D Research Scholar Image Processing Lab  
Department of Centre of Studies Resources Engineering Electrical Engineering,  
Indian Institute of Technology Bombay  
Mumbai 400076  
bi-lab.banerjee@iitb.ac.in

## ABSTRACT

A density based multi-feature combination approach to classify vehicles using support vector machine (SVM) classifier is proposed in this paper. Features like Haar, Gradient, RGB and Pyramidal histogram of oriented gradients are extracted from vehicle images and a kernel density approximation is performed. The results are compared with another set of multi-feature combination and effects of using kernel density estimators and various kernels like Gaussian, Polynomial and Histogram intersection in SVM are also studied. It is found that the above mentioned feature combination produces good results in comparison and effects of scaling, orientation and shadowing were properly compensated. A sample set of 9360 images were used with half of them used for training the classifier and the other half for testing. The vehicles were clustered into four categories, namely 2-wheeler, 3-wheeler, light motor vehicle (LMV) and heavy motor vehicles (HMV) and an accuracy greater than 90% has been observed. ■

## 1. INTRODUCTION

The use of CCTV cameras to aid in traffic surveillance, control and management has seen a significant increase over the years. The analysis and statistics of traffic data from CCTV feed helps immensely in handling lane closures, unexpected traffic diversions, preventing traffic jams, providing solutions to accident prone areas and the like. But utilizing manual labor to analyze huge volumes of data would prove unreliable and highly tedious, because of which automated traffic control systems were introduced. These automated systems could run with minimal human assistance over a long period of time and there are a multitude of different algorithms used for video processing to detect and classify vehicles, estimate their speeds etc. The problem of classifying vehicles in a video is quite exclusive to a particular camera setup and region since it involves factors like environmental conditions, position of the camera, angle of approach of the vehicles with respect to the camera, scaling and skewing factors, diversity in make and model of vehicles belonging to same class etc. The design of a generic and reliable vehicle classification algorithm

is considered difficult but there are various types of engineering solutions for these problems.

Commonly, vehicle classification algorithms differentiate vehicles based on their dimensions like axle width, wheel base, length of the vehicle and such, but these features can overlap among different class of vehicles. The vehicle image feature based approach is a more concrete way of clustering similar vehicles than a dimension based approach. In this paper present a multi-feature classification algorithm which uses a non-linear kernel based SVM classifier for a multi-class vehicle classification system. Our technique makes use of an amalgamation of Haar features, PHOG features, gradient features and RGB features with the addition of kernel density approximation on images of vehicles taken head-on from a camera. The classes are divided into four categories namely 2-wheeler, 3-wheeler, light motor vehicle (LMV), heavy motor vehicle (HMV).

The paper is divided into sections with section 2 describing the previous work, section 3 detailing the algorithm's framework and feature sets, section 4 describes the results. The paper is concluded in section 5.

## 2. LITERATURE REVIEW

A wide variety of feature sets and techniques have been used by many for classification purposes especially for vehicles. Kafai *et al.* [12] presents a dynamic Bayesian network structure for vehicle classification and has obtained results superior to those achieved using k-nearest neighbor algorithm, LDA and SVM. This method requires rear view image of a vehicle including the license plate. Buch *et al.* [3] presents a 3D spatial modeling technique using motion silhouettes which is called 3D HOG. He has shown good results using 3D modeling data in comparison to ordinary histograms and FFT. He has used 3D model data which is not generic even in the vehicle shapes commonly found among same vehicle class in countries like India. Rashid *et al.* [17] presents a shape feature based classifier which is less prone to environmental factors. He uses fuzzy c-means algorithm for clustering and k-nearest neighbor classifier, but shape features have a tendency to overlap and thus would prove inaccurate. Li

*et al.* [15] has used moment invariants of vehicle regions as feature set along with BP neural network for classifying vehicle into three basic classes based on size, namely small, medium and large. A visual vehicle classification using HOG for different orientations of vehicles is proposed by Rybski *et al.* [19] which can be integrated with LIDAR detector. This method requires various orientation images of the same vehicle for effective classification which might lead to a huge and redundant training data set. Meselodi *et al.* [17] presents a novel way of classifying bicycles and motorcycles by extracting the features from the wheel region of the vehicles and using a SVM classifier, but such specific features are not an efficient way to classify vehicles.

### 3. CLASSIFICATION FRAMEWORK

The proposed algorithm for classification consists of multiple feature extraction, application of kernel density approximation over the feature set to extract the probability vectors and classifying them using a SVM classifier. The steps involved are depicted in figure 1.

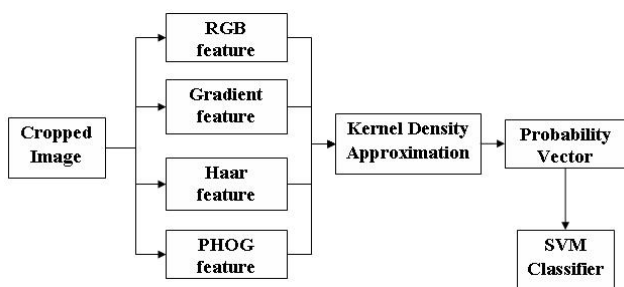


Fig. 1. Flow chart of vehicle classification mechanism

The following 4 feature used in our algorithms - RGB color, gradient, PHOG and Haar features to lessen the disadvantages of a pixel wise color modeling.

- (1) Haar feature - Haar-like features are popular image features used in object recognition by Sun *et al.* [20], Bai *et al.* [1], Haselhoff *et al.* [8] which are similar to Haar wavelets used in the first real-time face detector by Viola *et al.* [22]. One other contribution was to use summed area which they called integral images. Integral images can be defined as two-dimensional lookup tables in the form of a matrix with the same size as that of the original image. Each Haar-like feature may need more than four lookups, depending on how it was defined. Viola and Jones's 2-rectangle features need six lookups, 3-rectangle features need eight lookups, and 4-rectangle features need nine lookups.
- (2) Pyramidal Histogram of Oriented Gradients(PHOG) - Estimating the shape of an object is often described by the edge orientations of the object. Therefore calculate gradients of edges and put them in groups get PHOG which is used for multi class object classification as demonstrated by Bai *et al.* [2] and Gehler *et al.* [5]. Then in each interval a 'bin' and put each pixel in a bin such that gradient of edge at that pixel lies in the interval denoted by that bin.
- (3) RGB color feature - RGB color components are generally used to classify objects which have significant differences in their pixel colors. There are various color spaces like RGB, YCrCb, HSV etc and have chosen RGB due to its simplicity. Ghazalli [13] uses RGB color feature with Fuzzy logic classifier to classify palm fruits for oil extraction according to their ripeness. Gani [21] also uses RGB color feature with Fuzzy logic classifier for classifying pineapples for export according to their maturity level. Campbell [4] has used

RGB color data along with LIDAR data for developing an automated urban area classification system.

- (4) Gradient feature - Gradient based features has been chosen for their robustness to illumination based changes and used both direction and magnitude features. Javed *et al.* [10], Liu *et al.* [16], Mudigonda *et al.* [18], Huang *et al.* [9], Gavrilescu *et al.* [14] have all used gradient features for classification.

### 3.1 Kernel Density Estimation

Kernel density estimation (KDE) is one of a smoothing techniques and is also a non-parametric way to estimate the probability density function of a random variable. For  $(x_1, x_2, \dots, x_n)$  which is an independent and identically distributed random variable drawn from an unknown distribution with an unknown density  $f$ , estimated shape of this function  $f$  can be computed using its kernel density estimator which is given by

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - x_i) = \frac{1}{nh} \sum_{i=1}^n K_h \left( \frac{x - x_i}{h} \right) \quad (1)$$

where  $K(\cdot)$  is the kernel. The kernel is symmetric but it can either be a positive or negative function that integrates to one with  $h > 0$  as the bandwidth also known as smoothing parameter. A kernel  $K$  is called a scaled kernel if it has a subscript  $h$  and is defined as  $K_h(x) = \frac{1}{h} K\left(\frac{x}{h}\right)$ . The bandwidth  $h$  should be assigned a value as small as the data would allow. A range of kernel functions are commonly used for KDE namely uniform, triangular, biweight, triweight, Epanechnikov. In our experiment triangular kernel is use.

### 3.2 Kernel density approximation

Kernel density approximation (KDA) is a non-parametric multimodal density function based approximation utilizing a mixture of Gaussians which was proposed by Han *et al.* [7], [6]. Jin *et al.* [11] also has used similar approach for non-parametric clustering. A Gaussian distribution or a mixture of Gaussians is predominantly used in computer vision problems because of their simplicity and efficiency and has very good estimation and prediction statistics. So utilized this technique to extract the probability vectors and to enhance the accuracy of our classifier and used the technique described by Han *et al.* [7]. For construct a one dimensional density function for each feature at each pixel by kernel density estimation based on Gaussian kernel as follows:

$$\hat{f}_F(x) = \frac{1}{\sqrt{2\pi}} \sum_{i=1}^n \frac{K_{F,i}}{\sigma_{F,i}} \exp\left(-\frac{(x - x_{F,i})^2}{2\sigma_{F,i}^2}\right) \quad (2)$$

where  $\sigma_{F,i}$  and  $\kappa_{F,i}$  are the bandwidth and weight of  $i$ th kernel respectively. KDA finds the local maxima in the density function (equation 2) and a mode-based representation of density is obtained by estimating all the parameters for a Gaussian mixture. The simplified density function using KDA is

$$\tilde{f}_F(x) = \frac{1}{\sqrt{2\pi}} \sum_{i=1}^{m_F} \frac{\tilde{\kappa}_{F,i}}{\tilde{\sigma}_{F,i}} \exp\left(-\frac{(x - x_{\tilde{F},i})^2}{2\tilde{\sigma}_{F,i}^2}\right) \quad (3)$$

where  $x_{\tilde{F},i}$  is computed by the mean-shift mode finding procedure, and  $\tilde{\sigma}_{F,i}$  and  $\tilde{\kappa}_{F,i}$  are the estimated standard deviation and gaussian weight of each mode. The end result of KDA on the feature set are probability vectors which are then used by a linear SVM classifier.



Fig. 2. Vehicle Classes - two wheeler, three wheeler, LMV and HMV.

#### 4. EXPERIMENTS

In our tested algorithm with images of vehicles of different sizes and shapes and considered two wheelers, three wheelers, light motor vehicles (including light cargo vehicles), and heavy motor vehicles which are shown in figure 2. The images have been taken for different orientations and with camera calibration. After the vehicle's features are extracted, the classification step comes into the picture. The tools used Support Vector Machine classifier with the Matlab LIBSVM implementation and changed the training set size from 40% of the total images available to 70% with a step increase of 10%. After training image applied the classification step for unknown samples and present the performance of our classification algorithm along with precision-recall (P-R) graphs for various test cases. (true positive),

Figures 3, 4 and 5 depict the P-R graphs for our feature combination in low, medium and high density traffic scenes, where there is a variable amount of occlusion and probability of misclassification varies. The traffic density is classified as low, medium and high depending on the number of vehicles passing in a ten minute period. For low density the number of vehicles passed in a ten minute period is anywhere between 0 to 600, while it is 600 to 900 for medium density and 900 to 1300 for high density traffic. These scenarios are shown in figures 6, 7 and 8. It can be noted the feature combination we have selected shows better results in comparison is because each of the individual features are good for classifying in their own aspect and the fact that they represent different features of the same object adds to the improved accuracy. For instance Haar features are computationally cheap and have been widely used for object recognition and the like because of their use of overlapping integral images. RGB color features are used because vehicles have distinct color in majority of the samples and they act as a good discriminative feature. Gradient features are used because of their ability to resist illumination changes which have pronounced effect on outdoor scenarios. PHOG features provide histogram based edge descriptors which work perfectly for vehicle classification since the vehicles have definite edges. All these features combined make an excellent feature set for classification.

Figures 9, 10, 11, 12 and 13 portray the effects the feature combination and individual features have on the SVM classifier. Since the individual features have more of generalized descriptors to differentiate pixels, their accuracy based performance through P-R study with linear SVM is not good. But the feature combination with its multitude of discriminative features describing each pixel provides good density for SVM to provide satisfactory results. It can be inferred from these figures that the multi-feature combination of Haar, PHOG, Gradient and RGB features show reliable and more accurate results than the individual feature sets. Figure 14 is a comparison between various feature combinations including KDE, KDA, SIFT and shape features. It can be noted that the PHOG, Haar, gradient and RGB feature combination

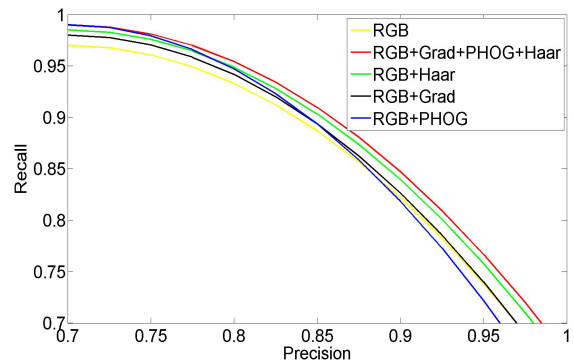


Fig. 3. P-R curves for our algorithm with different feature sets in low density traffic.

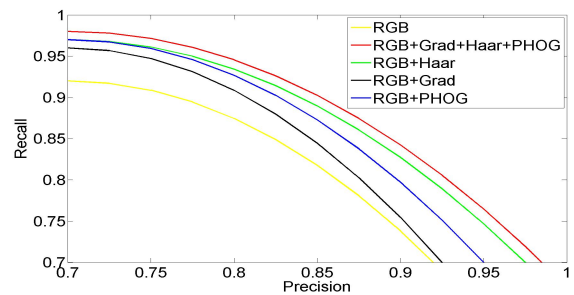


Fig. 4. P-R curves for our algorithm with different feature sets in medium density traffic.

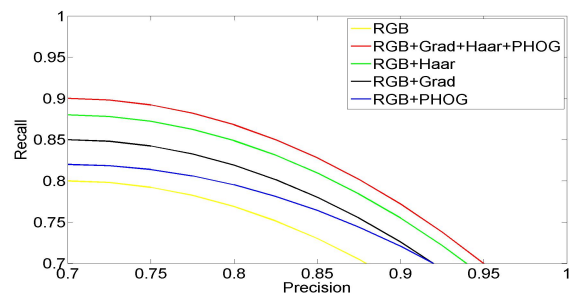


Fig. 5. P-R curves for our algorithm with different feature sets in high density traffic.

along with kernel density approximation produces the best results among the lot.

The accuracy of our proposed algorithm with RGB, Grad, Haar and PHOG feature combination is depicted in tables 1 to 6 with various comparisons for a given train-test set pair of calibrated data. In our experiment 9360 images are used and considered 50% of the dataset as the training set and rest as the test set. The tables 1 and 2 depict the difference in accuracy between two feature combinations PHOG, Haar, grad, RGB and PHOG, Haar, SIFT, shape feature without using any kernel density functions. Tables 3, 4, 5, 6, 7, 8 are the results obtained from the classifier using traditional Gaussian, polynomial and histogram intersection kernels for the two feature combinations. Table 9 and 10 show the results obtained when combining the feature sets with kernel density estimation. The tables 11 and 12 show the accuracies achieved with the feature combinations when paired with kernel density approximation.



Fig. 6. Low density traffic.



Fig. 7. Medium density traffic.

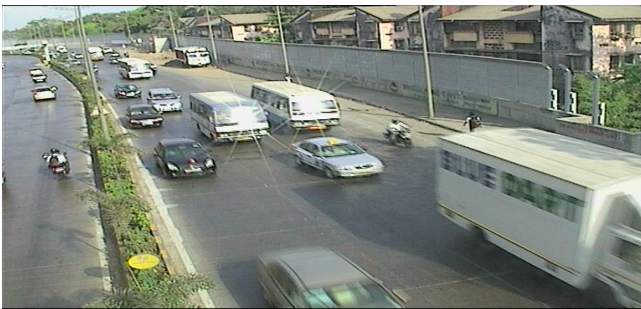


Fig. 8. High density traffic.

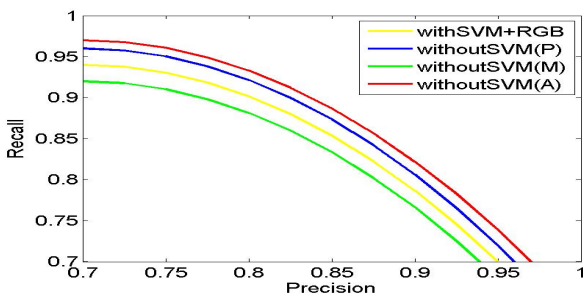


Fig. 9. P-R curves for SVM with RGB feature.

Table 1. Confusion matrix for SVM with sift+shape+haar+phog

	HMV	LMV	3-Wheeler	2 -Wheeler
HMV	248	146	15	17
LMV	186	2430	106	36
3-Wheeler	78	215	490	140
2-Wheeler	19	50	53	451

Accuracy = 77.3%

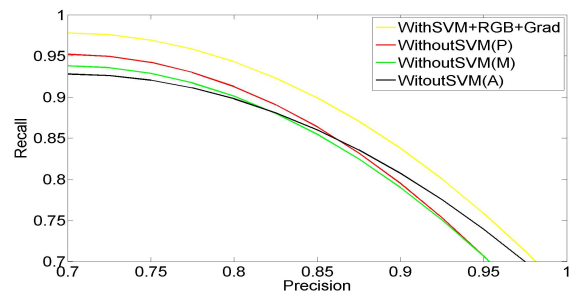


Fig. 10. P-R curves for SVM using RGB+Grad feature.

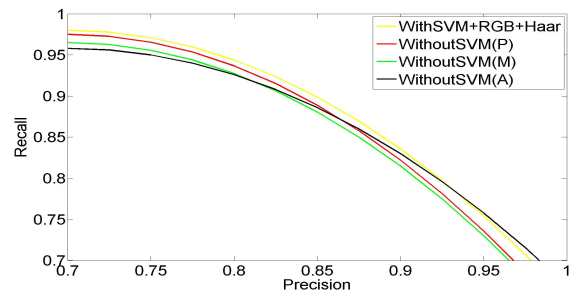


Fig. 11. P-R curves for SVM using RGB+Haar feature.

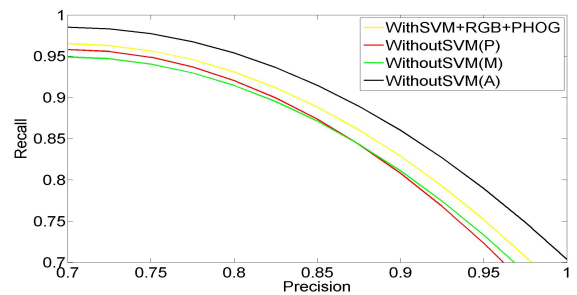


Fig. 12. P-R curves for SVM using RGB+PHOG feature.

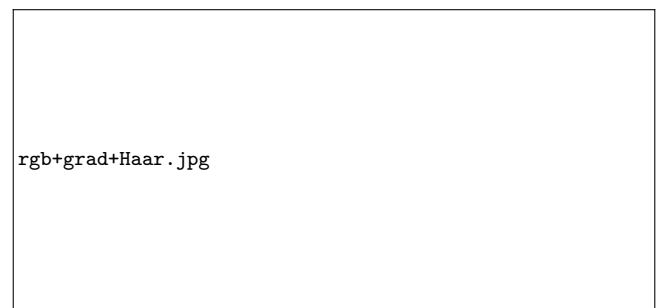


Fig. 13. P-R curves for SVM using RGB+Grad +Haar+PHOG feature.

Table 2. Confusion matrix for SVM with rgb+grad+haar+phog

	HMV	LMV	3-Wheeler	2 -Wheeler
HMV	378	29	1	10
LMV	74	2629	42	13
3-Wheeler	156	68	585	122
2-Wheeler	24	10	9	530

Accuracy = 88.1%

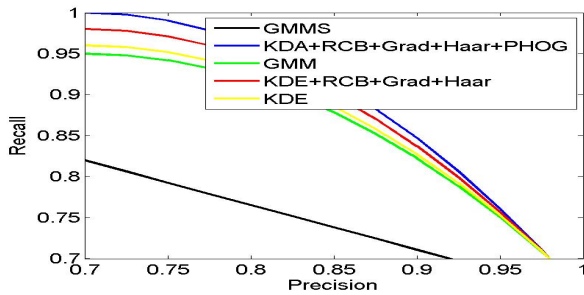


Fig. 14. P-R curves for different density functions.

**Table 3. Confusion matrix for SVM with sift+shape+haar+phog combination and Gaussian RBF kernel**

	HMV	LMV	3-Wheeler	2 -Wheeler
HMV	292	106	4	16
LMV	36	2639	56	27
3-Wheeler	78	185	540	128
2-Wheeler	11	46	3	513

Accuracy = 85.1%

**Table 12. Confusion matrix for KDA with rgb+grad+haar+phog**

	HMV	LMV	3-Wheeler	2 -Wheeler
HMV	404	13	1	0
LMV	57	2698	3	0
3-Wheeler	0	10	908	13
2-Wheeler	0	0	25	548

Accuracy = 97.4%

**Table 5. Confusion matrix for SVM with sift+shape+haar+phog combination and polynomial kernel**

	HMV	LMV	3-Wheeler	2 -Wheeler
HMV	399	13	0	6
LMV	82	2589	68	19
3-Wheeler	48	71	534	278
2-Wheeler	26	8	1	538

Accuracy = 86.7%

**Table 6. Confusion matrix for SVM with rgb+grad+haar+phog combination and polynomial kernel**

	HMV	LMV	3-Wheeler	2 -Wheeler
HMV	362	41	15	0
LMV	115	2580	43	15
3-Wheeler	0	42	840	49
2-Wheeler	0	15	43	515

Accuracy = 91.8%

## 5. CONCLUSIONS

In our proposed algorithm a multi-feature density based algorithm for vehicle classification, where KDA is used with multiple features such as RGB, Grad, Haar and PHOG. For classification, an SVM using the probability vectors from the given feature set is employed. Our classification algorithm shows better performance in combination with KDA than with KDE and just the features alone.

## 6. REFERENCES

- [1] H. Bai, J. Wu, and C. Liu. Motion and haar-like features based vehicle detection. In *Multi-Media Modelling Conference Proceedings, 2006 12th International*, pages 4–pp. IEEE, 2006.
- [2] Y. Bai, L. Guo, L. Jin, and Q. Huang. A novel feature extraction method using pyramid histogram of orientation gradients for smile recognition. In *Image Processing (ICIP), 2009 16th IEEE International Conference on*, pages 3305–3308. IEEE, 2009.
- [3] N. Buch, J. Orwell, and S. Velastin. 3d extended histogram of oriented gradients (3dhog) for classification of road users in urban scenes. 2009.
- [4] N. W. Campbell, B. T. Thomas, and T. Troscianko. Automatic segmentation and classification of outdoor images using neural networks. *International Journal of Neural Systems*, 8(01):137–144, 1997.
- [5] P. Gehler and S. Nowozin. On feature combination for multiclass object classification. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 221–228. IEEE, 2009.
- [6] B. Han, D. Comaniciu, and L. Davis. Sequential kernel density approximation through mode propagation. In *Proceeding of European Conference on Computer Vision*, 2004.
- [7] B. Han, D. Comaniciu, Y. Zhu, and L. Davis. Sequential kernel density approximation and its application to real-time visual tracking. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(7):1186–1197, 2008.
- [8] A. Haselhoff and A. Kummert. A vehicle detection system based on haar and triangle features. In *Intelligent Vehicles Symposium, 2009 IEEE*, pages 261–266. IEEE, 2009.
- [9] L.-L. Huang, A. Shimizu, Y. Hagihara, and H. Kobatake. Gradient feature extraction for classification-based face detection. *Pattern Recognition*, 36(11):2501–2511, 2003.
- [10] O. Javed, K. Shafique, and M. Shah. A hierarchical approach to robust background subtraction using color and gradient information. In *Motion and Video Computing, 2002. Proceedings. Workshop on*, pages 22–27. IEEE, 2002.
- [11] Y.-x. Jin, K. Zhang, J. T. Kwok, and H.-c. Zhou. Fast and accurate kernel density approximation using a divide-and-conquer approach. *Journal of Zhejiang University SCIENCE C*, 11(9):677–689, 2010.
- [12] M. Kafai and B. Bhanu. Dynamic bayesian networks for vehicle classification in video. *Industrial Informatics, IEEE Transactions on*, 8(1):100–109, 2012.
- [13] G. Khairun Nor Aimi. *Palm oil classification using RGB and fuzzy*. PhD thesis, Universiti Malaysia Pahang, 2010.
- [14] A. Klaser and M. Marszalek. A spatio-temporal descriptor based on 3d-gradients. 2008.

- [15] X. Li, H. Fu, and J. Xu. Automatic vehicle classification based on video with bp neural networks. In *Wireless Communications, Networking and Mobile Computing, 2008. WiCOM'08. 4th International Conference on*, pages 1–3. IEEE, 2008.
- [16] X. Liu and T. Yu. Gradient feature selection for online boosting. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–8. IEEE, 2007.
- [17] S. Messelodi, C. M. Modena, and G. Cattoni. Vision-based bicycle/motorcycle classification. *Pattern recognition letters*, 28(13):1719–1726, 2007.
- [18] N. R. Mudigonda, R. Rangayyan, and J. E. L. Desautels. Gradient and texture analysis for the classification of mammographic masses. *Medical Imaging, IEEE Transactions on*, 19(10):1032–1043, 2000.
- [19] P. E. Rybski, D. Huber, D. D. Morris, and R. Hoffman. Visual classification of coarse vehicle orientation using histogram of oriented gradients features. In *Intelligent Vehicles Symposium (IV), 2010 IEEE*, pages 921–928. IEEE, 2010.
- [20] Z. Sun, G. Bebis, and R. Miller. Quantized wavelet features and support vector machines for on-road vehicle detection. In *Control, Automation, Robotics and Vision, 2002. ICARCV 2002. 7th International Conference on*, volume 3, pages 1641–1646. IEEE, 2002.
- [21] E. Tasnim and A. Ghani. *Pineapple distribution classification using RGB and fuzzy*. PhD thesis, Universiti Malaysia Pahang, 2010.
- [22] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I–511. IEEE, 2001.