Vision based vehicle detection with occlusion handling

Himanshu chandel[1], Sonia Vatta[2],

[1]Bahra University, National Highway 22, Waknaghat, Himachal Pradesh, 173234
himanshu.chandel01@gmail.com@email.com

[2]Bahra University, National Highway 22, Waknaghat, Himachal Pradesh, 173234
busoniacse@gmail.com

Abstract: In recent years, automotive manufacturers have equipped their vehicles with innovative Advanced Driver Assistance Systems (ADAS) to ease driving and avoid dangerous situations, such as unintended lane departures or collisions with other road users, like vehicles and pedestrians. To this end, ADAS at the cutting edge are equipped with cameras to sense the vehicle surrounding. This dissertation investigates the techniques for monocular vision based vehicle detection. A system that can robustly detect and track vehicles in images. The system consists of three major modules: Histogram of oriented gradient (HOG) is used as the main feature descriptor which is shape oriented, a machine learning part based on support vector machine (SVM) for vehicle verification, to make the system biological human eye driven lastly a technique is applied for texture analysis by applying the concept of gray level co-occurrence matrix (GLCM).More specifically, we are interested in detection of cars from different camera viewpoints, diverse lightning conditions majorly images in low light and normal day light conditions further handling the low level occlusion. The images has been pre-processed at the first step to get the optimum results in all the conditions. In this thesis work, HOG is the main feature extraction method secondly, a texture based feature vector is calculated with the method of gray level co-occurrence matrix and to the end Support Vector Machine is used as a classifier. Experiments have been conducted on 1198 low light car images dataset with side and rear view, secondly normal day light conditions with 1657 image dataset, the dataset contains a mixed samples of partial and self-occlusion. For car images the classifier contains 2 classes of images with the combination of positive and negative images with the test and train segments. Due to length of long feature vector we have made the best filtration for deducing its size using different cell sizes for more accuracy and efficiency. Results will be presented and future work will be discussed.

Keywords: Vehicle, histogram, occlusion, vision, texture, detection.

1. Introduction

Advanced driver assistance systems (ADAS) are technologies that provide a driver with essential information, automate difficult or repetitive tasks, and lead to an overall increase in car safety for everyone. Some of these technologies have been around for a long time, and they have already proven to result in an improved driving experience and better overall road safety. In recent years, there has been a significant increase in the number of the embedded electronic processors in automobiles from 1% in 1980 to over 22% in 2007 [1]. Among the different embedded sub-systems, vision-based advanced driver assistance systems (ADAS) are becoming more popular in recent times because of the availability of low-cost, high-resolution and pervasive cameras [2]. In order to meet the increasing demand in vision-based ADAS, manufacturers such as Texas Instruments etc. are releasing newer embedded platforms that are specifically catered for implementing vision algorithms in automobiles [3]. Navigation systems are used in a wide range of applications such as: Automotive Navigations System, Marine Navigation System, Global Positioning System, Surgical Navigation System, Inertial Guidance system and Robotic mapping.in addition to navigation systems advance driver assistance system is there to maintain safe speed, driving within lane, collision avoidance and at last reducing the severity of accident. There are a wide range of issues and difficult tasks in the domain of navigation systems such as complex backgrounds, low-visibility, weather conditions, cast shadows, strong headlights, direct sunlight during dusk and dawn, uneven street illumination and the problem of occlusions on which this thesis is concentrating. In modest term 4 levels should be accomplished to make a robust navigation system: detection, localization, recognition and understanding. In this research work a number of experiments were carried out on different real world data sets for evaluation and analysis of the methods applied to obtain the best results for detecting vehicles. In the field of surveillance, autonomous navigation, scene understanding occlusion is one of the most common problem and due to this problem lots of vision based algorithms lacks in robustness.

The problem of occlusion is commonly defined under object tracking and object detection. The results are obtained by using the shape features with Histogram of orientation gradient and calculating the texture features such as: entropy, contrast, correlation, homogeneity and energy.

By implementing gray level co-occurrence matrix. Further the training of the features data set is done by support vector machines using statistica and results are obtained on the rate of car detection on different datasets. The dataset which we used in the experiment is collected online by utilizing the images of size 64*64 pixels with positive and negative samples that is car and non-car images which includes occluded samples. The feature vector window with cell size of 3*3 is used for histogram of orientation gradients on the images with the output of feature vector with 1*14400, secondly to minimize the feature vector rather using principal component analysis or any other technique we implemented the window size of 32 * 32 to get the feature vector of length 1*144. Texture is also one of the important part in a biological vision system, the concept of gray level co-occurrence matrix is used, further concatenating the feature vector on the basis of entropy, contrast, correlation, homogeneity and energy with 3 different colour channels red, green and blue. The concatenation of feature vector finally, gives us the feature vector which is further classified and tested on the basis of support vector machines using Statistica. The dataset has been trained and tested using support vector machines with 2 different kernel types: RBF and linear. The overall results has been presented with more than 90% of accuracy rate in the overall system.

2. Literature review

In computer vision, the higher level of understanding with the camera as a sensor involves the tasks such as: place recognition, scene understanding, object detection, object categorization and action recognition. The human perception is yet very intelligent as compared to machines, lots of research is going to make the real world applications by which it will be possible to give a vision to machines. Wide range of applications are there in the field of computer vision such as: Unmanned Aerial Vehicles (UAV's), Unmanned Ground Vehicles (UGV's), Unmanned underwater vehicles (UUv's),Advanced driver assistance systems (ADAS), Automated visual Traffic Surveillance (AVTS), Autonomous Navigation, Pedestrian Classification, Scene Understanding. In computer vision the sensor we use is camera, to give vision to machines researchers work commonly on two platforms: monocular [4] or binocular (stereo) [5] vision. Now, the concept comes what is the difference between one image of a scene and on the other hand two image of the same scene taken from different viewpoints. The two images let us infer the depth information by means of geometry, with the help of triangulation technique [6].

Object tracking and detection is a classical research area in the field of computer vision from decades. Numerous kinds of applications are dependent on the area of object detection, such as advance driving assistance system, traffic surveillance, scene understanding, autonomous navigation etc. Many challenges still exist while detecting an object such as illusion, low visibility, cast shadows and most importantly occlusions of object. Occlusions occur under two categories, firstly its, self-occlusion which means that, from a certain viewpoint, one part of an object is occluded by another part. Secondly, its inter-object occlusion which means when two objects being tracked occlude each other. We will review various occlusion handling methods that involved single and multiple cameras according to their application. In short, the objective of this thesis is to deliberate in detail the problem of car detection and occlusion in object tracking and further provide a best solution for the problem. Many computer vision algorithms suffer due to the presence of occluded objects in a scene. The region, which is occluded though, depends on the camera viewpoint. In some scenarios angle of the camera can define which part is occluded and which one is not, hence minimization approach, temporal selection, graph cut method and sum of squared distance are followed for handling the same problem of occlusion [7]. The very basic technique for predicting occlusion in tracking was started from the geometric information [8, 9]. One solution to occlusion problem is to use the fusion of multiple cameras to determine depth and further estimating the occluded part [10]. Depth estimation itself is a challenging problem since local features alone are insufficient to estimate depth at a point and needs to consider the global context of an image, hence a polynocular stereo algorithm is used for occlusion handling [11]. Rosales and Sclaroff presented the method of temporary and trajectory prediction for the same problem [12].Geometry of a scene plays a vital character in handling occlusion, the technique of epipolar lines and disparity map are used by Davi et.al to get the robust solution [13]. Little and Gillet put the technique of suppression using ordering constraint [14]. Identifying the occlusion regions, by defining the two occlusion maps, that is occlusion map1 showing parts of image 1, that are not visible in image 2 and similarly occlusion map 2 in image 2, while the displacement field between the images act as a key part in this proficiency [15]. Occlusion is the basic element that limits the information in an image. In vision based systems researchers commonly deal with binocular or stereo vision. Binocular is the process of obtaining depth information from a pair of cameras. In stereo vision occlusion occurs when a portion of the picture visible on one image is occluded in the other by the scene itself or, a section of the scene near the image boundary moves out of the field of persuasion on the other picture. The core challenge in tracking is to accurately detect the object in a wide range of environments such as sunny and rainy conditions in driving assistance systems. For autonomous navigation systems, whether its unmanned ground vehicles (UGV), unmanned aerial vehicles (UAV) or unmanned underwater vehicles (UUV), vision is playing a crucial role in path planning and autonomous navigation. In comparison to range sensors such as LIDAR and SONAR, camera is gaining attention for vision based systems as cameras not only deliver geometry but also provide rich appearance information. One of the simplest method for detecting occlusion is cross checking and extrapolation [16]. Mehmood, Nawaz and Rao in addition to the Bhattacharya Coefficient threshold used the normalized cross correlation method for handling occlusion. The same technique was followed further, Normalized cross correlation is used at every step followed by subtracting the mean and dividing the standard deviation [17, 18]. Asymmetric information as well as geometric, photometric constraints and robust cost aggregation for occluded candidates have also been used for handling occlusions [19]. The main objective of [20] is to detect and handle occlusion inside a tunnel and on real road using video as an input. In this paper Position of trajectories is estimated using Relative descriptive histogram, which is an extended approach of HOG [21]. Image segmentation is the core step for various computer vision algorithms and the same technique has been adopted by Enzweiler, Eigenstetter, Schiele and Gavrila for handling occlusion using Mean shift algorithm. Artificial neural networks are trained on intensity depth and motion features. The expert weights are computed that are related to the degree of visibility of the associated component [22]. In stereo vision occlusion is the classical research problem as images are captured from different angles and some pixels are available only in one image. The same approach is used for handling occlusion that is left and right disparity maps are obtained with the help of optimization based on modified constant space belief propagation [23, 24]. Scene understanding started with the dream of constructing machines that can see like humans to understand general principles and current situations from images, but it has become lot more extensive than that. Applications such as image search engines, autonomous driving, computational photography, vision for graphics, human machine interaction, were unexpected and other applications keep rising as scene understanding technology develops. As a center setback of high level computer vision, while it has enjoyed some great success in the past 50 years, a great deal more is needed to achieve a perfect understanding of visual scenes. Due to increase of traffic on roads, intelligent traffic surveillance systems are being implemented in various countries for highway monitoring and city road management system. Traditionally, shadow detection techniques have been employed for removing shadows from the background and foreground, but Sadeghi & Fathy used it as a novel feature for vehicles detection and occlusion handling. They have used photometric characteristic of darkest pixels of strong shadows in traffic images for occlusion handling. Multilevel framework is adopted by Zhang, Jonathan Wu, Yang and Wang for handling car occlusion in traffic; firstly occlusion is detected by evaluating the compactness ratio and interior distance ratio of vehicles, and then the detected occlusion is handled by removing a cutting region of occluded vehicles. On interframe level, occlusion is detected by performing subtractive clustering on the motion vectors of vehicle. Next comes the tracking level, occluded vehicles are tracked by using a bidirectional occlusion reasoning mechanism [25].

3. Shape analysis

The shape can be thought of as a silhouette of the object (e.g. obtained by illuminating the object by an infinitely distant light source). There are many applications where image analysis can be reduced to the analysis of shapes (e.g. organs, cells, machine parts, characters). Shape analysis methods play an important role in systems for object recognition, matching, registration, and analysis. The goal of shape analysis is to provide a simplified representation of the original shape so that the important characteristics of the shape are preserved. The work important in the above sentence may have different meanings in different applications. This general definition implies that the result of analysis can be either numeric (e.g. a vector) or non-numeric (e.g. a graph). The input to shape analysis algorithms are shapes (i.e. binary images).. Research in shape analysis has been motivated, in part, by studies of the human visual form perception system. Methods for shape analysis are divided into several groups. Classification is done according to the use of shape boundary or interior and according to the type of result. Four resulting combinations are boundary scalar, boundary space domain, global scalar, and global space-domain methods. The last few decades have resulted in an enormous amount of work related to shape analysis. As a kind of global shape description, shape analysis in transform domains takes the whole shape as the shape representation.

Generally, low level features such as colour, texture, shape, corner, etc., are used to represent the approximate perceptual representation of an image, using which similarity and dissimilarity of the images are computed. But it is found that the perceptual representation of an image in terms of low level features fails to capture entire semantic information of an image and it is often difficult to model accurately.

In this work The Histogram of oriented gradient approach is utilized first to figure out the best features. It operates on a dense grid of uniformly spaced cells and used in local contrast normalization on overlapping blocks for improved accuracy. Basic idea behind HOG is that the appearance and shapes of local objects within an image can be well described by the distribution of intensity gradients as the votes for dominant edge directions. Such descriptor can be obtained by first dividing the image into small contiguous regions of equal size, called cells, then collecting histogram of gradients directions for the pixels with each cell and at last combining all these histograms. In order to improve the detection accuracy against varied illumination and shadowing, local contrast normalization can be applied by computing a measure of the intensity across larger region of image, called a block and using resultant value to normalize all cells within the block. There are 2 basic variants of HOG descriptors commonly used: Rectangular and Circular HOG.

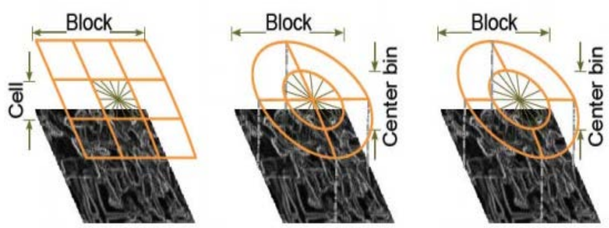

Figure 1: Variants of HOG descriptors. (a) A rectangular HOG (R-HOG) descriptor with 3 × 3 blocks of cells. (b) Circular HOG (C-HOG) descriptor with the central cell divided into angular sectors as in shape contexts. (c) A C-HOG descriptor with a single central cell.

The process starts from computation of image gradients, this is done by applying 1D centered, point discrete derivative

ISSN 2229-5518

mask in both the horizontal and vertical directions, which in specific are filter kernels of following form.

[-1, 0, 1] and $[-1,0,1]^T$

There are many more complex masks, such as sobel, prewitt, canny or diagonal mask, but these mask result in poorer performance. Then magnitude and orientation at each pixel I(x, y) is calculated by:

$$G_{mag}(x,y) = \sqrt{G_x^{\,2}(x,y) + G_y^{\,2}(x,y)}$$

$$\Theta\,(\boldsymbol{x},\boldsymbol{y}) = \text{arc tan}\,\frac{G_y(x,y)}{G_x(x,y)} + \frac{\pi}{2}$$

$\boldsymbol{G_x(x,y)}$ and $\boldsymbol{G_y(x,y)}$ are gradient values at each pixel in horizontal and vertical direction. For colour images, the channel with the largest magnitude gives the pixels dominant magnitude and orientation.

Orientation Binning

In this histograms for each cell are created. Cells are pixel regions that are either rectangular or radial in shape and histogram bins are evenly expanded from 0° to 180° or from 0° to 360°, so every histogram bin has a spread of 20°. Every pixel in the cell casts a weighted voting into one of the 9 histogram bins that its orientations belongs to. As for the weight of votes, it can either be the gradient magnitude itself, or some function of the magnitude for example square root or square of the gradient magnitude, or some clipped version of the magnitude. Generally the gradient magnitude is directly used.

Descriptor blocks

To obtain the robustness against various illumination and contrast, the gradient strengths must be locally normalized. This leads to grouping the cells into larger pixel regions called blocks. These blocks overlap with neighboring blocks, so that each cell can contribute its orientation distribution more than once. As mentioned before, totally 4 block geometries exist, with one most commonly used: Rectangular HOG blocks. R-HOG blocks are usually square grids and the optimal parameters are found to be 2*2 cell blocks of 8*8 pixel cells. Besides, a Gaussian spatial window can be applied to each block before histogram voting so that the weight of each pixel around the edge of the block can be significantly suppressed.

Block Normalization

There are 4 different ways proposed to normalize the blocks: Let V denote the non-normalized feature vector that collects all cell histograms from a given block, $||V_k||$ denotes its k-norm for k=1, 2. The normalization schemes helps in normalizing the bins to get the optimum results of feature vectors.

4. Texture Analysis

Texture is an important cue for biological vision systems to estimate the boundaries of objects. Also, texture gradient is used to estimate the orientation of surfaces. For example, on a perfect lawn the grass texture is the same everywhere. However, the further away we look, the finer this texture becomes – this change is called texture gradient. For the same reasons, texture is also a useful feature for computer vision systems.

Gray level co-occurrence matrix is computed, firstly we separated the intensity in the image into a small number of different levels. Normalization of the matrix is done by determining the sum across all entries and dividing each entry by this sum. This co-occurrence matrix contains important information about the texture in the examined area of the image. The GLCM is created from a gray The GLCM is created from a gray-scale image.

We calculated and examined the following parameters for Red, Green and Blue colour channels.

$$\text{Energy} = \sum_{a,b} P^2(a,b)$$

$$\text{Entropy} = \sum_{a,b} P(a,b)\log_2 P(a,b)$$

$$\text{Contrast} = \sum_{a,b} |a-b|^\kappa\, P^\lambda(a,b),\ \text{usually}\ \kappa = 2,\ \lambda = 1$$

$$\text{Correlation} = \frac{\sum_{a,b}[(ab)P(a,b)] - \mu_x\mu_y}{\sigma_x\sigma_y}$$

$$\text{Homogeneity} = \frac{P(a,b)}{1+|a-b|}$$

5. Classifier

A support vector machine is a classifier defined separating a hyper plane. SVM's belong to a family of generalized linear classifiers. The foundations of Support Vector Machines (SVM) have been developed by Vapnik [26] and gained popularity due to many promising features such as better empirical performance. SVMs are commonly used to solve the classification problem. SVM is a useful technique for data classification. A classification task usually involves with training and testing data which consist of some data instances [27].

SVM is related to statistical learning theory which was first introduced in 1992. SVM becomes popular because of its success in handwritten digit recognition. It is now regarded as an important example of "kernel methods", one of the key area in machine learning. In this thesis we are using 1 D feature vector for the classification problem of cars and non – car.

Each instance in the training set contains one target values and several attributes. The goal of SVM is to produce a model which predicts target value of data instances in the testing set which are given only the attributes [64]. Classification in SVM is an example of Supervised Learning. Known labels help indicate whether the system is performing in a right way or not. This information points to a desired response, validating the accuracy of the system, or be used to help the system learn to act correctly. A step in SVM classification involves identification as which are intimately connected to the known classes. This is called feature selection or feature extraction. Feature selection and SVM classification together have a use even when prediction of unknown samples is not necessary. They can be used to identify key sets which are involved in whatever processes distinguish the classes [28]. Support vector machines are used to identify the 2 classes, as in this thesis we are using car and non-car. A good decision boundary is a key for this algorithm. Consider a 2 classes such as car and non- car, which is a linearly separable classification problem.

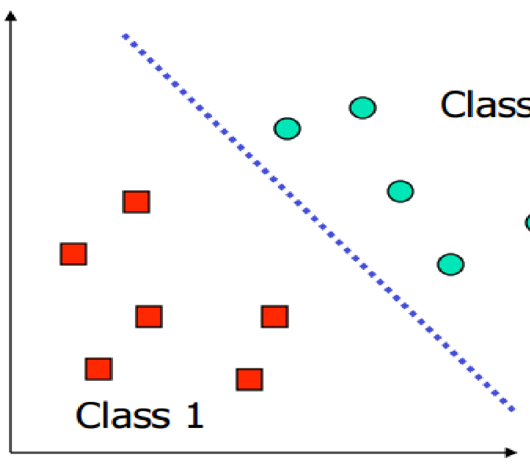

Figure 2: Linearly separable classification

6. Results

Our method was tested on a number of challenging sequences .The resolution of the images used for testing was 64 × 64 pixels. The algorithm has been developed using Matlab and for training purpose Statistica is used. The algorithm is run on an Intel Core I5, 3.00 GHz machine.

This section presents the experimental results obtained using a trained and testing model. The first one, based on the method suggested by [21] builds the appearance model based on shape feature vector and normalizing the histograms based on gradient. The feature vector of Gray level co-occurrence matrix based on 3 different colour channels, Red, Green and Blue. The Entropy, Contrast, Correlation, Energy and Homogeneity are calculated for each image to get the best results using texture and shape features.

The histogram of orientation gradient is implemented with the 20 * 20 cell size, to collect the best feature vector. The classification is performed using Statistica and implementing Support vector Machines (SVM).

Handling occlusion was one of the most critical parts in this work. Vehicle tend to be on roads and in different traffic conditions so the chances of occluding one car with another are high further depending on self-occlusion, as camera viewpoint angle is different . To solve this problem, some methods try to position the camera in the ceiling or at a vertical angle in the hopes of decreasing the chances of occlusion in surveillance systems. Although the vertical angle of a camera can solve occlusion problems, it is not applicable in all potential uses of vehicle tracking. The texture analysis gave us the optimum results for handling occlusion and with both shape and texture features, we were able to get the accuracy of more than 90% in both the conditions of real world scenarios.

6.1 Low light conditions

Figure 3: Sample of low light condition dataset

6.2 Kernel Type – RBF

Table 1: Classification Summary – Support Vector Machine

Class Name	Total	Correct	Incorrect	Correct (%)	Incorrect (%)
Car	645	602	43	93.33333	6.666667
Non - Car	645	620	25	96.12403	3.875969

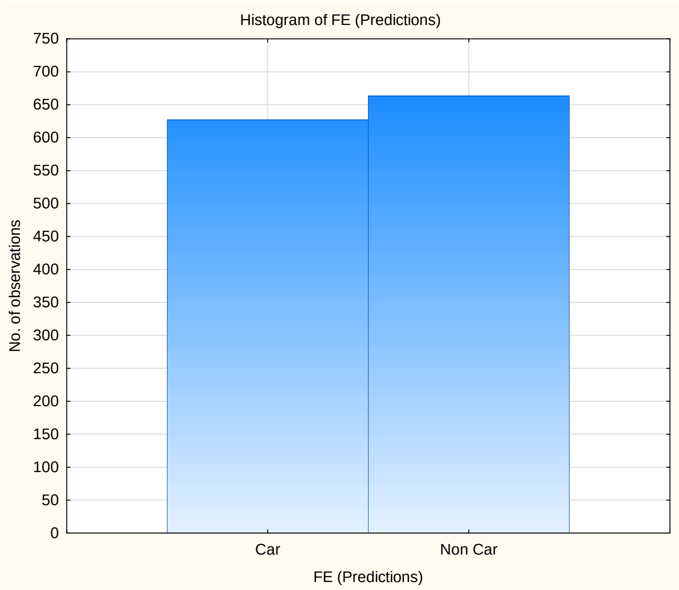

Figure 4: Histogram of observed dataset

```
Dataset Book1 final features to svm:
  Dependent: FE
  Independents: B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S.
  Sample size = 1104 (Train), 1290 (Test), 2394 (Overall)

Support Vector machine results:
  SVM type: Classification type 1 (capacity=10.000)
  Kernel type: Radial Basis Function (gamma=0.006)
  Number of support vectors = 306 (264 bounded)
  Support vectors per class: 150 (Car), 156 (Non Car),

  Class. accuracy (%) = 95.924(Train), 94.729(Test), 95.280(Overall)
```

Figure 5: Overall performance of detector

6.3 Kernel Type – Linear

Table 2: Classification Summary – Support Vector Machine

Class Name	Total	Correct	Incorrect	Correct (%)	Incorrect (%)
Car	645	591	54	91.62791	8.372093
Non - Car	645	605	40	93.79845	6.201550

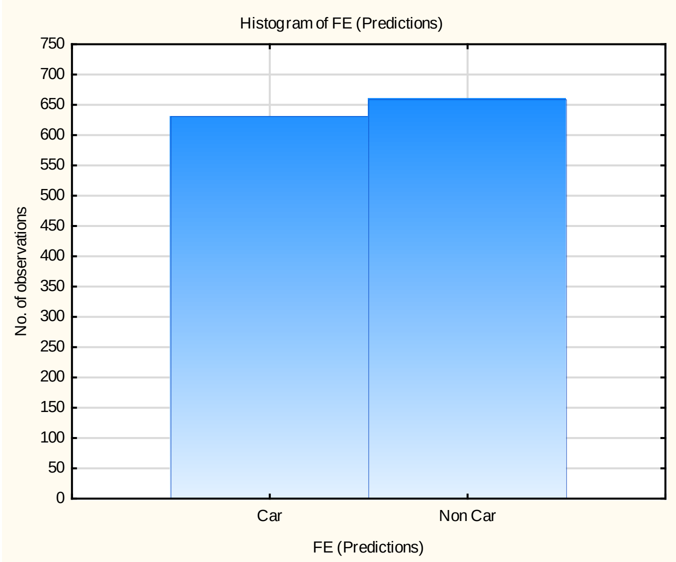

Figure 6: Histogram of observed dataset

```
Dataset Book1 final features to svm:
  Dependent: FE
  Independents: B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S.
  Sample size = 1104 (Train), 1290 (Test), 2394 (Overall)

Support Vector machine results:
  SVM type: Classification type 1 (capacity=10.000)
  Kernel type: Linear
  Number of support vectors = 197 (9 bounded)
  Support vectors per class: 92 (Car), 105 (Non Car),

  Class. accuracy (%) = 98.822(Train), 92.713(Test), 95.530(Overall)
```

Figure 7: Overall performance of detector

6.4 Normal day light conditions

ISSN 2229-5518

Figure 8: Sample of normal daylight condition dataset

6.5 Kernel Type – RBF

Table 3: Classification Summary – Support Vector Machine

Class Name	Total	Correct	Incorrect	Correct (%)	Incorrect (%)
Car	892	878	14	98.43049	1.569507
Non - Car	900	864	36	96.00000	4.000000

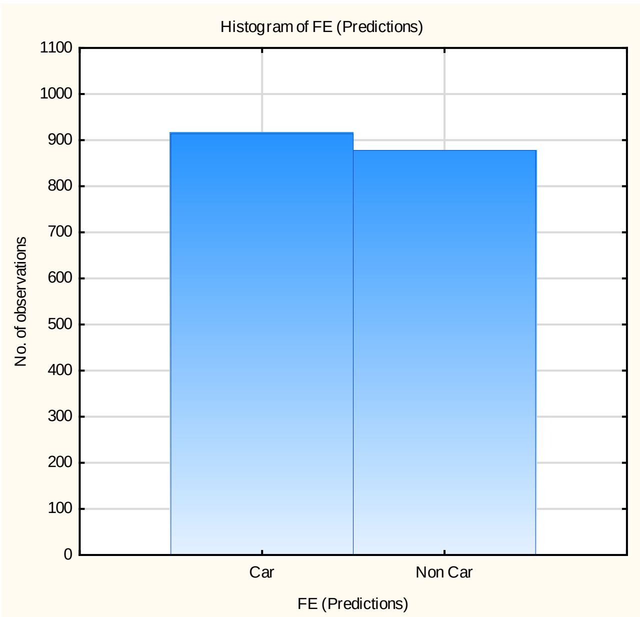

Figure 9: Histogram of observed dataset

```
Dataset Sheet1 in Book4 to SVM:
  Dependent: FE
  Independents: B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S.
  Sample size = 1533 (Train), 1792 (Test), 3325 (Overall)

Support Vector machine results:
  SVM type: Classification type 1 (capacity=10.000)
  Kernel type: Radial Basis Function (gamma=0.006)
  Number of support vectors = 252 (208 bounded)
  Support vectors per class: 128 (Car), 124 (Non Car),

  Class. accuracy (%) = 97.847(Train), 97.210(Test), 97.504(Overall)
```

6.6 Kernel Type – Linear

Table 4: Classification Summary – Support Vector Machine

Class Name	Total	Correct	Incorrect	Correct (%)	Incorrect (%)
Car	892	858	34	96.18834	3.811659
Non - Car	900	870	30	96.66667	3.333333

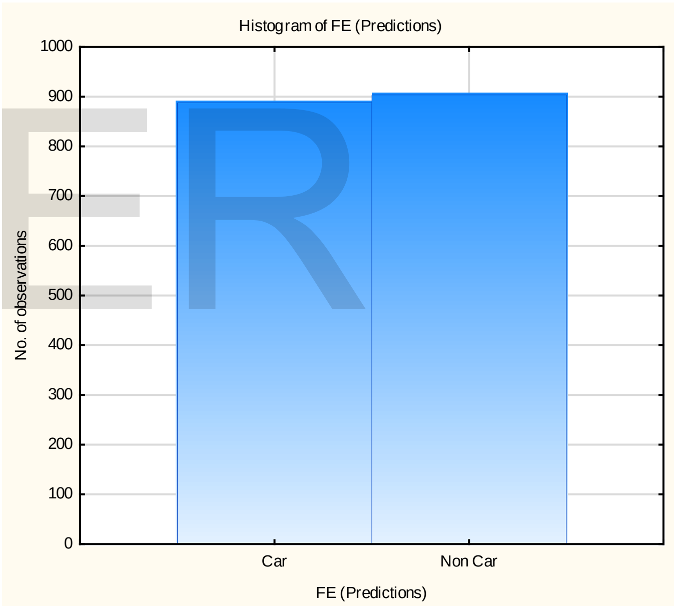

Figure 11: Histogram of observed dataset

```
Dataset Sheet1 in Book4 to SVM:
  Dependent: FE
  Independents: B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S.
  Sample size = 1533 (Train), 1792 (Test), 3325 (Overall)

Support Vector machine results:
  SVM type: Classification type 1 (capacity=10.000)
  Kernel type: Linear
  Number of support vectors = 167 (7 bounded)
  Support vectors per class: 93 (Car), 74 (Non Car),

  Class. accuracy (%) = 99.543(Train), 96.429(Test), 97.865(Overall)
```

Figure 12: Overall performance of detector

References

ISSN 2229-5518

[1] Andrej Karpathy and Li Fei-Fei, “Deep visual-semantic alignments for generating image descriptions,” arXiv preprint arXiv: 1412.2306, 2014.

[2] Sing Bing Kang, Szeliski R. , Jinxiang Chai, “Handling occlusions in dense multi-view stereo,” Computer Vision and Pattern Recognition, 2001, pp. I-103 - I-110 vol.1.

[3] M. Zeeshan Zia, Michael Stark and Konrad Schindler, “Are Cars Just 3D Boxes? – Jointly Estimating the 3D Shape of Multiple Objects,” IEEE Conference on Computer Vision and Pattern Recognition, 2014 , pp. 3678 – 3685.

[4] H. Bassmann, Ph.W Besslich, “Monocular computer vision,” Third International Conference on Image Processing and its Applications, 1989, pp. 107 - 111.

[5] Massimo Bertozzi and Alberto Broggi, “GOLD: A Parallel Real-Time Stereo Vision System for Generic Obstacle and Lane Detection,” IEEE Transactions on image processing, vol. 7, no. 1, 1998, pp. 62 - 81.

[6] Triangulation https://en.wikipedia.org/wiki/Triangulation_(computer_vision)

[7] Sing Bing Kang, Szeliski R. , Jinxiang Chai, “Handling occlusions in dense multi-view stereo,” Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001, pp. I-103 - I-110 vol.1.

[8] P. Wunsch and G. Hirzinger, “Real-Time Visual Tracking of 3-D Objects with Dynamic Handling of Occlusion,” Proceedings of the IEEE International Conference on Robotics and Automation, 1997 , pp. 2868 - 2873 vol.4

[9] Fabienne LathuiIiere and Jean-Yves Herve, “Visual Tracking of Hand Posture with Occlusion Handling,”15th International Conference on Pattern Recognition,2000, pp. 1129 - 1133 vol.3

[10] Ting - Hsun Chang,Shaogang Gong and Eng Jon Ong, “Tracking multiple people under occlusion using multiple cameras,” *BMVC,2000.*

[11] Yasuyuki SUGAYA and Yuichi OHTA, “Stereo by the integration of two algorithms with/without occlusion handling,”15th International Conference on Pattern Recognition,2000

[12] Ro´merRosales and StanSclaroff, “Improved Tracking of Multiple Humans with Trajectory Prediction and Occlusion Modeling,” IEEE Conf. on Computer Vision and Pattern Recognition,1998

[13] Davi Geiger,Bruce,Ladendorf,Alan Yuille, “Occlusions and Binocular Stereo,”International Journal of Computer Vision,1995

[14] James J. Little and Walter E. Gillet, “Direct Evidence for Occlusion in Stereo and Motion,” Computer Vision — ECCV, 1990

[15] Juyang Weng,Narendra Ahuja,Thomas S. Huang, “Two view matching,” IEEE,1988

[16] Geoffrey Egnal and Richard P. Wildes, “Detecting Binocular Half-Occlusions:Empirical Comparisons of Five Approaches,”IEEE Transactions on pattern analysis and machine intelligence, VOL. 24, NO. 8, 2002

[17] Yaowen Guan, Xiaoou Chen, Deshun Yang, Yuqian Wu, “Multiperson tracking by detection with local particle filtering and global occlusion handling,” IEEE International Conference on Multimedia and Expo (ICME), 2014

[18] Tao Yang, Stan Z.Li, Quan Pan, Jing Li, “Real-time Multiple Objects Tracking with Occlusion Handling in Dynamic Scenes,” IEEE Computer Society Conference on Computer Vision and Pattern Recognition,2005

[19] Dongbo Min and Kwanghoon Sohn, “Cost Aggregation and Occlusion Handling With WLS in Stereo Matching,” IEEE Transactions on image processing, VOL. 17, NO. 8,2008

[20] Bing-Fei Wu et al. , “A Relative-Discriminative-Histogram-of-Oriented-Gradients-Based Particle Filter Approach to Vehicle Occlusion Handling and Tracking,” IEEE Transactions on Industrial Electronics, VOL. 61, NO. 8,2014

[21] Navneet Dalal and Bill Triggs, “Histograms of Oriented Gradients for Human Detection,” IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005, pp. 886 - 893 vol. 1.

[22] Markus Enzweiler,Angela Eigenstetter,Bernt Schiele,Dariu M. Gavrila, “Multi-Cue Pedestrian Classification With Partial Occlusion Handling,” IEEE Conference on Computer Vision and Pattern Recognition, 2010

[23] Woo-Seok Jang , Yo-Sung Ho, “Efficient depth map generation with occlusion handling for various camera arrays,” Signal, Image and Video Processing, Volume 8, Issue 2, pp 287-297,2014

[24] Woo-Seok Jang and Yo-Sung Ho, “Efficient Disparity Map Estimation Using Occlusion Handling for Various 3D Multimedia Applications,” IEEE Transactions on Consumer Electronics, 2011

[25] Whei Zhang,Q.M Jonathan Wu,Xiaokang Yang, “Multilevel framework to detect and handle Occlusion,” IEEE Transaction on Intelligent Transport Systems,VOL.9,NO.1,2008

[26] Corrina Cortes and Vladimir Vapnik, “AT & T Labs Research, USA

[27] Duda R. and Hart P., "Pattern Classification and Scene Analysis", Wiley, New York 1973.

[28] Nello Cristianini and John Shawe-Taylor, “An Introduction to Support Vector Machines and Other Kernel-based Learning Methods”, Cambridge University Press, 2000.