respectively. The trained vehicle detector is tested on the test data and efficient results are obtained from vehicle detection problem. In addition, the success detection rate of the trained detector has been tried to be maximized as much as possible and experimental analysis comparisons are made with the results obtained from the methods.. Overall section will detail the proposed method, wheras will illustrate implemnetaiton details and also the exerimental results. Finally , the study will be concluded in section 4.

## 2 Methods used for Vehicle Detection

The working method consists of six main stages. These are respectively; loading the data set, the design of the convolutional neural network, configuration of training options, training of the Faster R-CNN object detector, evaluation of trained detector. These stages and conventional and the faster R-CNN methods will be discussed in this section.

### 2.1 R-CNN (Regions with Convolutional Neural Network Features)

The R-CNN approach combines two basic concepts. From these, the first is to carry out efficient convolutional neural networks from bottom to up region proposals to locate and dismember objects. Next, when the label training data is insufficient, it is followed by a supervised training for a field-specific fine tuning task, which provides significant performance improvement. The method is named R-CNN (CNN-enabled regions) because Regional proposals are combined with CNNs.

The working object detection system composed of three modules. Firstly, it categorically produces independent region proposals. These essentially describe the candidate detection set that can be used by the detector. The second module includes a convolutional neural network, producing an attribute vector of constant length from every region. The third module, on the other hand, includes a cluster of linear SVMs that are specific to the class used for assortment of regions [8].

#### 2.1.1 Region Proposal

Various recent studies have provided methods to produce categorical independent zone recommendations.These methods have examples such as the objectness of image windows [1] , selective Search for Object Recognition [3] , category independent object proposals [4] , object segmentation using constrained parametric min-cuts [5] , Multiscale combinatorial grouping [6] and so on [7]. These methods establish cells by implementing convolution neural network with square cuts. Although R-CNN is not based on the specific zone proposal method,

#### 2.1.2 CNN (Convolutional Neural Network) for Feature extraction

In this study, a feature vector of size 4096 were extracted from each region proposal with Caffe deep learning framework. Features were calculated by forwarding the average output 227x227 red-green-blue image with five convolution layers and two completely connected layers.

In order to calculate an attribute in a region proposal, the image data is first converted to a form compatible with CNN. (In this study, fixed entrances of 227 * 227 pixels in size are used.). Then,the most simple of the possible transformations of the random-shaped regions was selected. Here, all the pixels in a tight bounding box around the candidate area are resolved unto the required size, regardless of the size or aspect ratio. Before dissolving, the tight bounding box was expanded to provide w pixels skewed picture content around the box at the skewed dimension (w = 16 was used). In addition, a simple bounding box regression was used to expand the localization performance within the application [13]. This is shown in the following equation (1). The details of this equation can be seen in [8].

$$w_* = \underset{w_x}{argmin} \sum_{i}^{N} (t_*^i - w_*^T \emptyset_5(P^i))^2 + \lambda ||W_*||^2 \quad (1)$$

#### 2.1.3 Classify Regions

In this study, selective search was performed on test images to obtain approximately 2000 region proposals at test time. Each proposal has been resolved and advanced through CNN for the calculation of attributes. Then, for each class, each produced attribute vector is scored using the trained support vector machine (SVM). Considering the scored regions, greedy non-maximum suppression is applied independently when there is high-intersection (IoU) overlap with the selected zone with a higher rating over a learned threshold for a rejected region.