IJournals: International Journal of Software & Hardware Research in Engineering
ISSN-2347-4890
Volume 3 Issue 8 August, 2015

A NOVEL APPROACH FOR AUTOMATIC NUMBER PLATE RECOGNITION SYSTEM FOR INDIAN NUMBER PLATES

Author: Keerti Puloria[1]; Mr. Sachin Mahajan[2]

Affiliation: PG Scholar, CSE Department, J.I.T. Borawan (M.P.)[1];

HOD of IT Department, J.I.T. Borawan (M.P.)[2]

E-mail: kpuloria@gmail.com[1]; er_sachin1980@yahoo.com[2]

ABSTRACT

This paper presents the Automatic Number Plate Recognition (ANPR) System for Indian Cars. ANPR is a real life application. Basically ANPR systems are consist of Vehicle image capture, Preprocessing, Extraction of number plate, Character segmentation and Character recognition steps. Image capturing is usually done by existing closed-circuit television or road-rule enforcement cameras, or special cameras specifically designed for the task. ANPR systems should be quickly and should process license plates under different environmental conditions, such as indoors, outdoors, day or night time successfully. We are using median filter, histogram equalization to pre-process the vehicle image. Segmentation technique is based on discontinue properties of pixels in which we detect points, lines and edges of number plate. For edge detection we use "Canny Edge Detector" operator. And for character recognition we have applied SVM technique. Experiments are performed using database of car images and the method gives 98.33% correct classification rate.

Keywords: SVM, Canny Edge Detector, OCR, Grey Level Image, Binary Level Image.

1. INTRODUCTION

Automatic number plate recognition (ANPR) is a mass surveillance method which was invented in 1976 at the Police Scientific Development Branch in the UK that uses OCR i.e. optical character recognition technique on images to read vehicle number plates. ANPR is also known as automatic vehicle identification, car plate recognition, automatic license plate recognition, and optical character recognition (OCR) for cars. They can use existing closed-circuit television (Figure 1) or road-rule enforcement cameras, or ones specifically designed for the task.

Figure 1: Closed Circuit Television

ANPR can be used to store the images captured by the cameras as well as the text from the license plate, with some configurable to store a photograph of the driver. Systems commonly use infrared lighting to allow the camera to take the picture at any time of the day. It can be used in various fields like in police forces, electronic toll collection on pay-per-use roads and cataloging the movements of traffic or individuals.

ANPR systems may also be used by:

- Section control, to measure average vehicle speed over longer distances.
- Border crossings
- Automobile repossessions
- Petrol stations to log when a motorist drives away without paying for their fuel.
- A marketing tool to log patterns of use
- Targeted advertising, a-la "Minority Report"-style billboards.
- Traffic management systems, which determine traffic flow using the time it takes vehicles to pass two ANPR sites

ISSN-2347-4890
Volume 3 Issue 8 August, 2015

- Analyses of travel behavior (route choice, origin-destination etc.) for transport planning purposes
- Drive through Customer Recognition, to automatically recognize customers based on their license plate and offer them the items they ordered the last time they used the service.
- To assist visitor management systems in recognizing guest vehicles.
- Police and Auxiliary Police
- Car parking companies
- Hotels

Different ANPR technologies are used due to plate variation from place to place. And the variations of the plate types like location, quantity, size, color, occlusion, font etc. or environments cause like illumination and background challenges in the detection and recognition of license plates.

Basically ANPR consist of following steps as shown in figure 2:-

a) Vehicle image capture
b) Preprocessing
c) Extraction of number plate
d) Character segmentation
e) Character recognition

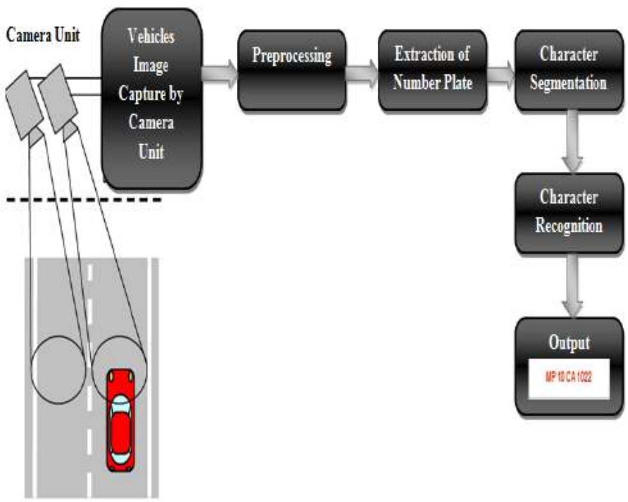

Figure 2: Steps of ANPR system

1.1 Vehicle Image Capture

The first step i.e. to capture image of vehicle looks very easy but it is quite exigent task as it is very difficult to capture image of moving vehicle in real time in such a manner that none of the component of vehicle especially the vehicle number plate should be missed. For this purpose existing closed-circuit television or road-rule enforcement cameras, or special cameras specifically designed for the task. Vehicle Image Capture is shown in figure 3.

Figure 3: Vehicle Image Capture

1.2 Pre-processing

This step is essential to enhance the input image and making it more suitable for the next processing steps. In this step eliminate as much background noise as possible, contrast enhancement and de-blurring to optimize the localization algorithm and also save the processing time.

1.3 Extraction of Number Plate

In this stage, the location of the license plate is identified and the output of this stage will be a sub-image that contains only the license plate. This is done in two main steps.

a. Locating a large bounding rectangle over the license plate.
b. Determining the exact location of the license plate.

1.4 Character Segmentation

After locating the LP and skew correction, next step is the segmentation of characters. Character segmentation is the procedure of extracting the characters from the LP image. Almost all the papers that had been surveyed [1], [3] used horizontal and vertical projection to segment the characters.

1.5 Character Recognition

After segmenting the characters, the next step is character recognition. Almost all ANPR systems are using different types of machine learning techniques are used for character recognition like ANN, SVM, HMM etc. other method is templates machining method. In ANPR mostly used templates machining. If we used machine learning techniques so we required feature extraction.

2. LITERATURE REVIEW

An algorithm proposed by *Aksoy et.al. 2000 [1]* to improve the performance of recognition. This algorithm trains character samples and obtains the rules that are used to recognize the numbers on number plates. But drawback of this method was that it is not robust to image rotation, translation and scaling and it cannot distinguish digits 6 and 9 without additional observation.

Tran Duc Duan et.al. 2005 [2] proposed an efficient boundary line-based method combining the Hough transform and Contour algorithm. But this method still has a few errors when dealing with bad quality plates.

Optical character recognition technique is used by *Muhammad Tahir Qadri et.al. (2009) [3]* for the character recognition. But the OCR methods used in this project for the recognition is sensitive to misalignment and to different sizes.

Rani Thakur et.al. (2012) [4] proposed a new license plate recognition approach based on the Radial Basis Function Neural Networks (RBFNN). But accuracy is not acceptable in this method and some problematic features like distance, light and corner are restricted.

3. SVM TECHNIQUE

Support Vector Machine is supervised Machine Learning technique. Support Vector Machine (SVM) was first heard in 1992, introduced by Boser, Guyon, and Vapnik. Support vector machines (SVMs) are a set of related supervised learning methods used for classification and regression. They belong to a family of generalized linear classifiers. In another terms, Support Vector Machine (SVM) is a classification and regression prediction tool that uses Machine Learning theory to maximize predictive accuracy while automatically avoiding over-fit to the data. Support Vector machines can be defined as systems which use hypothesis space of a linear functions in a high dimensional feature space, trained with a learning algorithm from optimization theory that implements a learning bias derived from statistical learning theory. Support vector machine was initially popular with the NIPS community and now is an active part of the Machine Learning research around the world. It is also used for many applications, such as ANPR systems, Iris Recognition, hand writing analysis, face analysis and so forth, especially for pattern classification and regression based applications. The foundations of Support Vector Machines (SVM) have been developed by Vapnik and gained popularity due to many promising features such as better empirical performance. SVMs were developed to solve the classification problem, but recently they have been extended to solve regression problems.

3.1 SVM for Classification

SVM is a useful technique for data classification. Even though it's considered that Neural Networks are easier to use than this, however, sometimes unsatisfactory results are obtained. A classification task usually involves with training and testing data which consist of some data instances. 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.

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.

3.2 Strength and Weakness of SVM

The major strengths of SVM are the training is relatively easy. No local optimal, unlike in neural networks. It scales relatively well to high dimensional data and the trade-off between classifier complexity and error can be controlled explicitly. The weakness includes the need for a good kernel function. Support Vector Machines have significant theoretical advantages over other Machine Learning methods. They don't require any independence assumptions, unlike Naive Bayes. SVMs are capable of discovering non-linear separating boundaries between classes, while Maximum Entropy can discover only linear ones. Finally, even though SVMs accept only numerical features, categorical features can be used too, by mapping them to numerical features using one out-of-m encoding. In practice, this means that one can use both numerical and categorical features to describe the instances of the problem, which cannot be done easily with Maximum Entropy or Naive Bayes[41].

3.2.1 Strengths

The main strengths of the SVM approach are:

3.2.1.1 Maximization Of Generalization Ability

In training a multilayer neural network classifier, the sum-of-squares error between outputs and desired training outputs is minimized. Thus, the class boundaries change as the initial weights change. So does the generalization ability. Thus, especially when training data are scarce and linearly separable, the generalization ability deteriorates considerably. But because a support vector machine is trained to maximize the margin, the generalization ability does not deteriorate very much, even under such a condition.

3.2.1.2 No Local Minima

A multilayer neural network classifier is known to have numerous local minima, and there have been extensive discussions on how to avoid a local minimum in training. But because SVM is formulated as a quadratic programming problem, there is a global optimum solution.

3.2.1.3 Robustness to Outliers

Multilayer neural network classifiers are vulnerable to outliers because they use the sum-of-squares errors. Thus to prevent the effect of outliers, outliers need to be eliminated before training, or some mechanism for suppressing outliers needs to be incorporated in training. In support vector machines the margin parameter C controls the misclassification error. If a large value is set to C, misclassification is suppressed, and if a small value is set, training data that are away from the gathered data are allowed to be misclassified. Thus by properly setting a value to C, we can suppress outliers.

3.2.2 Weaknesses

The weaknesses of support vector machines explained so far are as follows.

3.2.2.2 Extension to Multiclass Problems

Unlike multilayer neural network classifiers, SVM use direct decision functions. Thus an extension to multiclass problems is not straightforward, and there are several formulations. One of the purposes of this book is to clarify relations between these formulations.

3.2.2.3 Long Training Time

Because training of a SVM is done by solving the associated dual problem, the number of variables is equal to the number of training data. Thus for a large number of training data, solving the dual problem becomes difficult from both the memory size and the training time.

3.2.2.4 Selection Of Parameters

In training a support vector machine, we need to select an appropriate kernel and its parameters, and then we need to set the value to the margin parameter C. To select the optimal parameters to a given problem is called model selection. This is the same situation as that of neural network classifiers. Namely, we need to set the number of hidden units, initial values of weights, and so on. In support vector machines, model selection is done by estimating the generalization ability through repeatedly training support vector machines. But because this is time-consuming, several indices for estimating the generalization ability proposed.

4. PROPOSED SYSTEM

The car number plate has up to ten characters. Usually the number plate consists of two main sections. The upper section contains main information of the number plate, and the lower part is for the name of the state. In order to speed up the process, we use histogram projection to separate number plate into two groups. The first group usually consists of three or four letters and three or two digits. The second group mainly includes the name of the state. Therefore, two sets of SVMs are designed according to these two groups of characters. One set of SVMs is designed for recognizing characters of number plates and the other one is designed for characters representing the state. In the experiments shown in [13], it is concluded that “one against all” (OAA) could obtain higher accuracy than method of “one against one” (OAO). In the following experiments, only OAA method is adopt.

For real time character recognition of number plates, there are many factors causing misrecognition. For example, the numbers may also appear slanted due to the orientation of the video system, the illumination condition may vary according to the time of day and the changing weather, and the characters in number plate may be obscured by rust, mud, peeling paint, and fading colour. In addition, the contrast between characters and number plate surfaces can be affected by their colours. Therefore, the recognition system must be robust to many changes in dealing real time images .For ANPR system we are proposing following system as shown in figure 4.

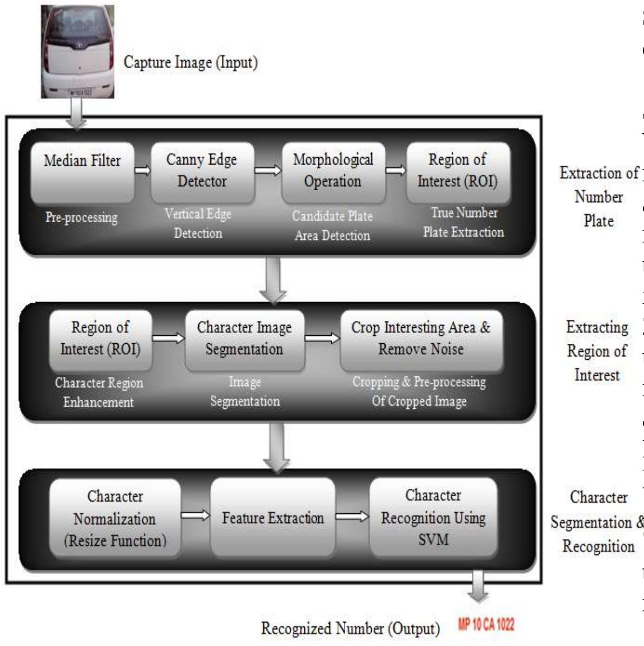

Figure 3: Proposed System

Furthermore the recognition system must be fast and not too expensive in real-life application. In order to solve these problems mentioned above, in our SVM-based recognition system, two kinds of SVMs are set up first. Each SVM has one type of number samples as one positive label and all or some of the other samples as another negative label. After training, each SVM gets its own values of parameters. The decision value of the testing sample will be calculated based on the values of parameters obtained. The final recognition result will be achieved according to the class that gives the maximum decision value. In order to recognize a number plate we are proposing following algorithm.

4.1 Algorithm

Step 1. First of all capture the vehicle's image.

Step 2. Pre-process the image of number plate.

Step 3. Segment and normalize the character Image.

Step 4. Extract the feature vector of each normalized candidate.

Step 5. Train SVMs based on saved sample database.

Step 6. Recognize the number plate by the set of SVMs trained in advance.

Step 7. If there are no more unclassified samples, then STOP.

Otherwise, go to Step 6.

Step 8. Add these test samples into their corresponding database for further training.

5. RESULT ANALYSIS

We have proposed a new ANPR system to recognize the Indian car number plates. With new algorithms for detection, segmentation, and recognition stages considering various features of the license plates. These algorithms are implemented using MATLAB 7.1 and applied on 25 grey scale images of Indian vehicles snapped under different illumination conditions, such as night, sunny, and cloudy. We also test the algorithms by using the vehicle images containing English license plates taken from an international license plate database. We also implement the ANPR systems presented in [3–5] using MATLAB 7.1. Experiment set up and parameter values of these algorithms in detection, segmentation, and recognition stages are shown in figures and graphs.

5.1 Snapshots for Number Plate Recognition

5.1.1 Snapshot of Main form

This screen contains Main Form which displays different options.

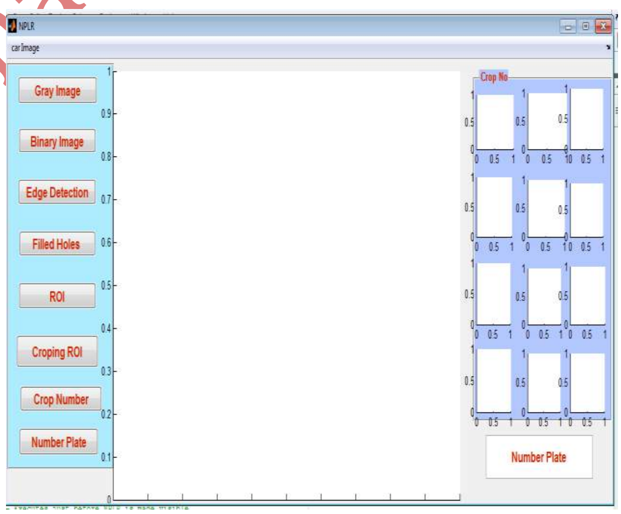

Figure 4: Main form

5.1.2 Snapshot of Loading Image

This screen shows that user has to first select number plate from file selector. File selector contain number of images of car number plates. User can choose any image of number plate from this file selector folder to load image.

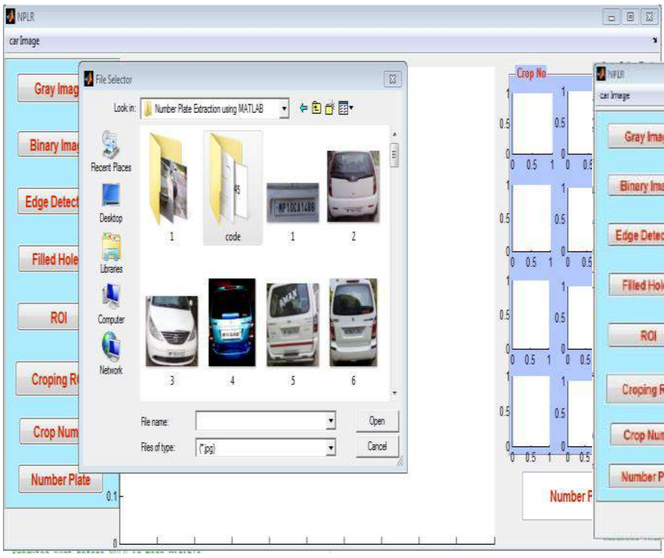

Figure 5: Loading Image

5.1.3 *Snapshot of RGB Image*

This screen shows the loaded image selected by user which is in RGB format. After loading image in RGB format user can apply options one by one, which are on left side of the screen to process image.

Figure 6: RGB Image

5.1.4 *Snapshot of Gray Image*

This screen shows that the loaded image is a colored image so to convert that color image into gray level image user has to press button grey image. Then the colored image which contains red, green, blue colors is converted into grey level image. After converting it into gray level image it has only black, white and gray in color of 8-bit format. The main purpose of converting RGB image into Grey image is to reduce number of colors of number plate and to improve processing speed and efficiency of operation.

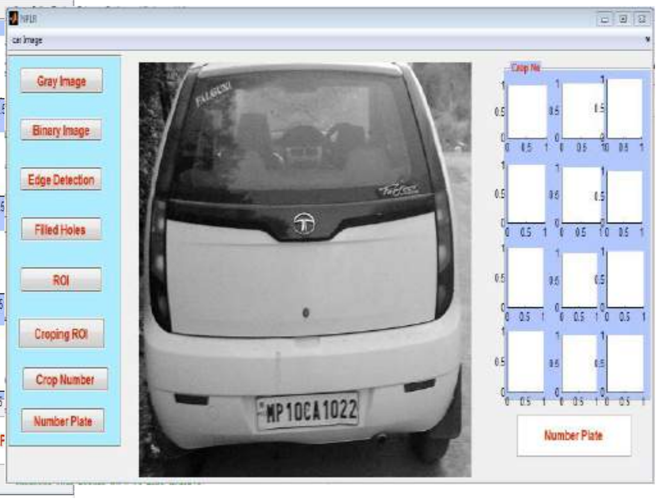

Figure 7: Gray Image

5.1.5 *Snapshot of Binary Image*

This screen shows that grey level image is now converted into binary image by pressing the button Binary Image. An image consists of numeric values between 0 - 255. Here the numerical value of the picture is reduced to two values with binary level. Thus, an 8 - bit image is converted into 2 - bit format.

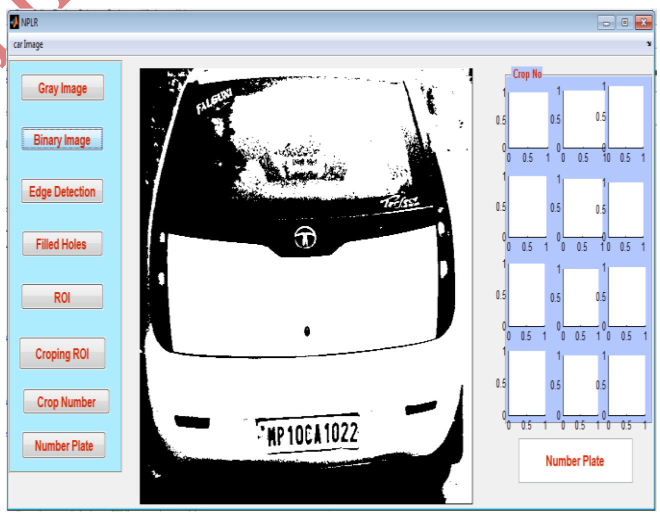

Figure 8: Binary Image

5.1.6 *Snapshot of Edged Image*

This screen shows the edge detection of binary image. User press the button Edge Detection and edge detection process in a binary image is then determine the frontiers of all represented objects based on automatic processing of the binary level information in each present pixel because the image edge is most basic and most obvious feature of any image.

Figure 9: Edged Image

5.1.7 Snapshot of Filled Holes Image

This screen shows the Filled holes option pressed by user is used to fill small holes including numbers of number plate so that number plate area will be large to isolate from figure.

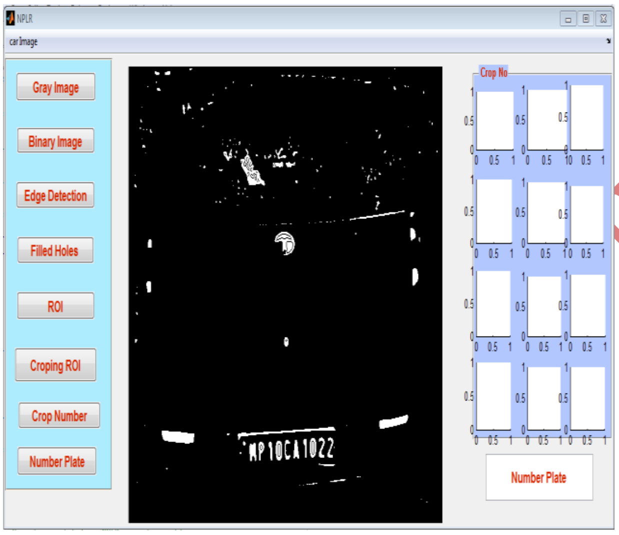

Figure 10: Filled Holes Image

5.1.8 Snapshot of ROI Image

This screen shows the ROI image of vehicle. ROI stands for Region of Interest. When ROI button is pressed by user the algorithm successfully detects the ROI that only contain vehicle number plate.

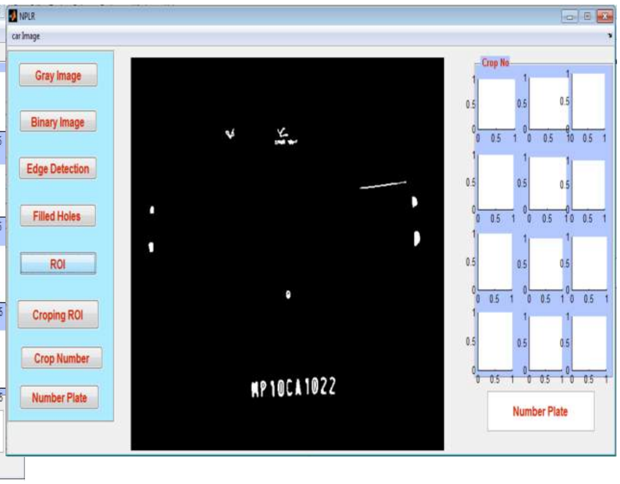

Figure 11: ROI Image

5.1.9 Snapshot of Cropping Area image

This screen shows that, after finding Region of Interest necessary area or part of image are cropped from whole image.

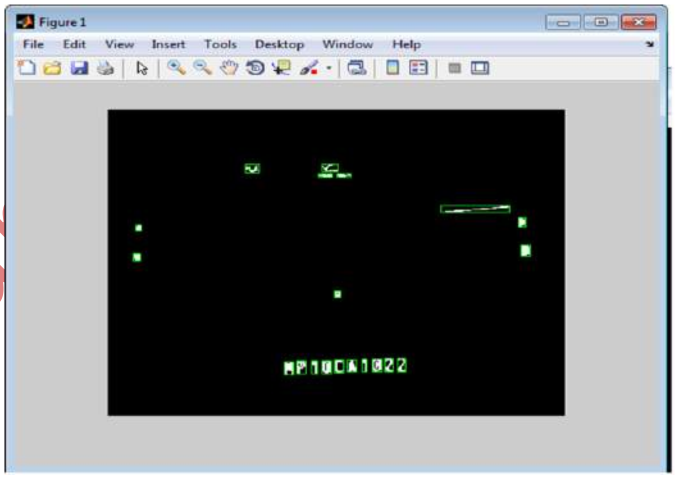

Figure 12: Cropping Area Image

5.1.10 Snapshot of Cropped Number

This screen shows that after cropping interested area of image, the number and alphabets are cropped which are display in rectangular boxes at right side of image.

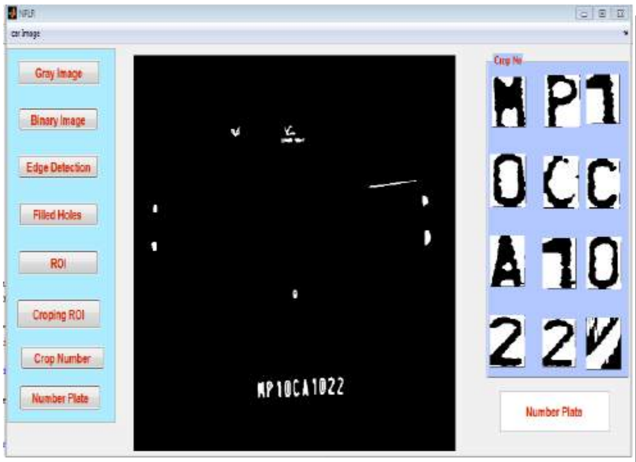

Figure 13: Cropped Number

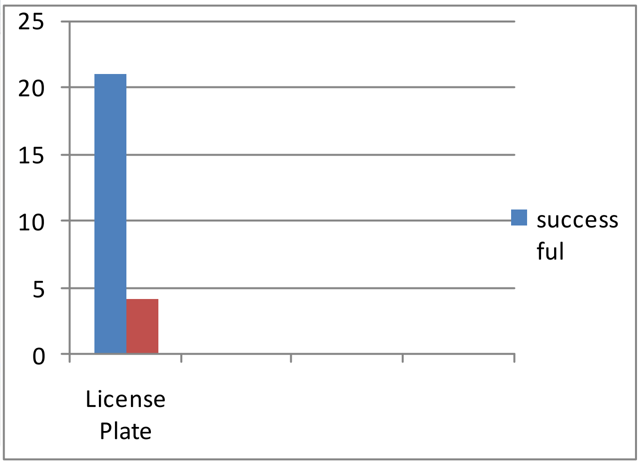

Figure 15: For Front Side of Car Number Plate

5.1.11 Snapshot of Number Plate

This screen shows that after cropping numbers user press the button Number Plate and the numbers and alphabets of number plate are display in rectangular box below at right section of screen. And this is the required numbers and alphabets of a car's number plate which we want to recognize.

Figure 14: Number Plate

5.2 Result Analysis

Result obtain are as follows:-

5.2.1 Graph for Front Side of Car Number Plate

The below graph show that when the image is captured from front side of the car around 21 number plates out of 25 images are recognized successfully while 4 number plates can't be determined.

5.2.2 Graph for Back Side of Car Number Plate

The below graph show that when the image is captured from back side of the car around 19 number plates are recognized successfully out of 25 images while 6 number plates can't be determined.

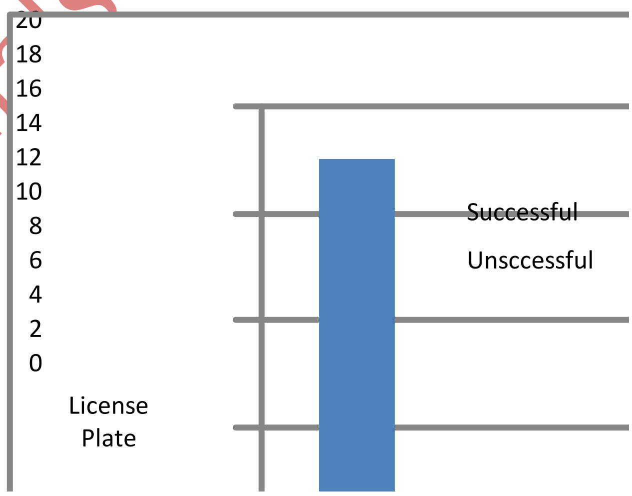

Figure 16: For Back Side of Car Number Plate

5.2.3 Graph for Overall Successful Number Plate

The below graph show the overall successful rate of number plate recognition. Here the 40 number plates are recognized successfully out of 50 images while 10 number plates can't be determined.

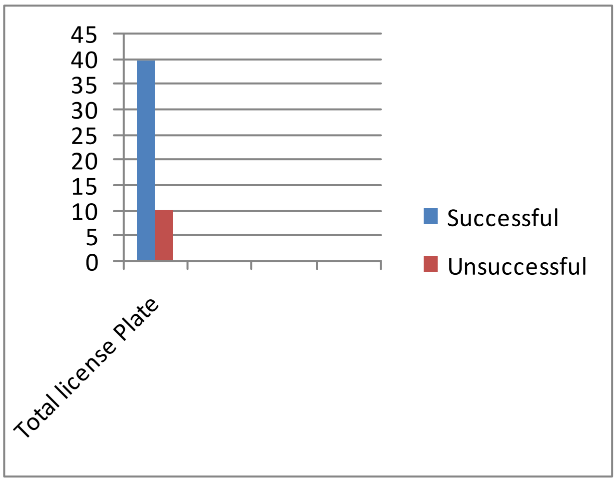

Figure 17: Overall Successful Number Plate

5.3 Result for Number plate character

Test results indicate that results for the characters "B" and "D" are very similar. It may be stated that character recognition algorithm ignores the curve in the middle while processing the character from the left side. The reason of this ignoring is this curve's being very tiny in real plate screens. What is more, algorithm assumes this curve as a noise and handles it as being straight.

Due to small differences between the upper and lower parts of the "D" character, it couldn't be recognized and this caused misidentification of the character.

Very similar results were encountered for the characters "6", "G", "0" and occasionally these characters were mixed with each other. Identification rate of the characters "5"and "S", "2" and "Z", "I" and "1", "U" and "H" were found very close. Due to this, in some occasions, they couldn't be distinguished from one another.

6. CONCLUSION AND FUTURE WORK

Here in this thesis, the methods for traffic surveillance have been presented and the work on motion detection, license plate extraction and character recognition is carried out. In motion detection, a study on different background subtraction available in the literature has been studied and their performance tests on the different video test sequence are given. The fitness coefficient and error coefficient is also calculated for all the methods. It should be noted that robust motion detection is a critical task and its performance is affected by the presence of varying illumination, background motion, camouflage, shadow, and etc.

In license plate extraction the strength and weakness of the different extraction algorithm have discussed which are available in the literature and comparisons of all the methods have been done. Proposed two methods for extraction of license plate i.e. edge detection method and the block variance technique are presented. The block variance algorithm has been tested on 90 images and giving 87.4% accuracy measure. In character extraction template matching (OCR) algorithm is used for extraction and different algorithms that are presented in literature survey are also studied. For improving the performance of template matching algorithm the format of license plate is studied.

This integrated system locates tracks and extracts traffic parameters in real time. Furthermore, the system can utilize any existing traffic surveillance infrastructure without further modification or tuning (except for the camera calibration that calculates image metrics). Overall, the system was found to work satisfactorily and the background reconstruction algorithm added robustness to the process. In normal traffic conditions the system responded well and the outcome results regarding vehicle license plate and trajectory were accurate enough. The experiments carried out showed that the proposed algorithm is capable of real time operational working due to its low complexity. The background reconstruction algorithm allows the unobstructed operation of the system without human intervention. The system works well either in real time mode or in already stored video.

In future work, we aim to focus on night surveillance and to improve the existing algorithms reported in literature. However, the other segments of our suggested system should be improved, focusing on the occlusion handling, vehicle matching procedure and also focus on improving the accuracy measure for character recognition by using the concept of neural network for recognizing all font type of a character by using back propagation algorithm. In this, first the network is trained and to train the network, the input and target are required. After the network had been successfully trained, the segmented character in license plate can now inputted the neural network to simulation. Ideally, the input characters will compare with the data that trained in neural network, and then outputted the ASCII code for corresponding input character.

7. REFERENCES

[1] Aksoy, M. S., Cagil, G. and Turker, A. K. (2000) "Number-plate recognition using inductive learning". *Robotics and Autonomous Systems, Elsevier, Vol.33, pp.149-153.*

[2] Tran Duc Duan, Tran Le Hong Du, Tran Vinh Phuoc, Nguyen Viet Hoang, "Building an Automatic Vehicle License-Plate Recognition System", *Intl. Conf. in Computer Science – RIVF'05 February 21-24, 2005, Can Tho, Vietnam.*

[3] Muhammad Tahir Qadri and Muhammad Asif , "Automatic Number Plate Recognition System For Vehicle Identification Using Optical Character Recognition",*International Conference on Education Technology and Computer, 2009.*

[4] Rani Thakur, Dr. Manish Manoria, "RBFNN Approach for Recognizing Indian License Plate", *International Journal of Computer Science and Network (IJCSN) Volume 1, Issue 5, October 2012 www.ijcsn.org ISSN 2277-5420.*

[5] V. Shapiro, G. Gluhchev, and D. Dimov, "Towards a multinational car license plate recognition system," Mach. Vis. Appl., vol. 17, no. 3, pp. 173– 183, Aug. 2006.

[6] G.Sun, G.Li, L.Xu and J.Wang, "The Location and Recognition of Chinese Vehicle License Plates under Complex Backgrounds," Journal of Multimedia., vol. 4, no.6, pp.442-449 Dec 2009.

[7] D.H. Ballard, "Generalizing the Hough Transform to Detect Arbitary Shapes," Pattern Recognition, vol. 13, no. 2, pp. 111-122, 1981.

[8] H. Caner, H. S. Gecim, and A. Z. Alkar, "Efficient embedded neural network- based license plate recognition system," IEEE Trans. Veh. Technol., vol. 57, no. 5, pp. 2675–2683, Sep. 2008.

[9] C. Anagnostopoulos, I. Anagnostopoulos, V. Loumos, and E. Kayafas, "A license plate-recognition algorithm for intelligent transportation system applications," IEEE Trans. Intell. Transp. Syst., vol. 7, no. 3, pp. 377–392, Sep. 2006.

[10] Y.Wen, Y.Lu, J.Yan, Z.Zhou, K.M.von Deneen, and P.Shi, "An Algorithm for license Plate Recognition Applied to intelligent Transportation system," IEEE Trans.Intell.Transp.Syst., vol. 12, no.3, pp. 830-845, Sept- 2011.

[11] Christos Nikolaos E. Anagnostopoulos, Ioannis E. Anagnostopoulos, Vassili Loumos, and Eleftherios Kayafas, "A License Plate-Recognition Algorithm for Intelligent Transportation System Applications," pp. 377-392, 2006.

[12] A.Tihar, A.Adnan, M.Fahad "License Plate Recognition for Pakistani License plates", Canadian journal on image processing., vol. 1, no.2, April 2010.

[13] Ying Wen et al., "An Algorithm for License Plate recognition Applied to Intelligent Transportation System," IEEE Transactions of Intelligent Transportation Systems, pp. 1-16, 2011.

[14] M.I.Khalil,"Car Plate Recognition Using the Template Matching Method", International Journal of Computer Theory and Engineering, Vol. 2, No. 5, October, 20101793-8201.

[15] Kaushik Deba, Md. Ibrahim Khana, Anik Sahaa, and Kang-Hyun Job, "An Efficient Method of Vehicle License Plate Recognition Based on Sliding Concentric Windows and Artificial Neural Network", Science Direct, CSIT-2012.