

Smart Number Plate Recognition System Enhanced by Machine Learning

Dharanya Subramanian¹, Kavinn Raj G¹, Madhu Priya S¹, Premamala T²

U.G. Student, Department of IT, Velalar College of Engineering and Technology, Erode, Tamil Nadu, India¹

Associate Professor, Department of IT, Velalar College of Engineering and Technology, Erode, Tamil Nadu, India²

ABSTRACT: The usage of vehicles is raising exponentially in the day to day life, which leads to violation of traffic rules, ingoing to restricted areas, theft of vehicles, abnormal number of accidents, which in-turn results in linear increase in crime rates. This enforces the need for number plate recognition system. In this paper, a Smart Number Plate Recognition System that makes use of the emerging technology is introduced. The developed system aims on increasing the efficiency that uses image processing technology for identification of vehicles. It is an automated system enhanced by machine learning to recognize vehicles without authorization. The approach used to segment the images is CCA where all the connected pixels belong to the same object. The machine learning model is trained using SVC and it undergoes 4-fold cross validation to measure accuracy of the model.

KEYWORDS: Image processing, CCA-Connected Component Analysis, SVC - Support Vector Classifier, Cross validation

I. INTRODUCTION

The research on automatic number plate recognition system has prioritized first in recent years due to its uniqueness in identifying the vehicles effectively. The smart number plate recognition system can be used in solving various issues in real time scenarios. The common problem faced by various commercial and residential organizations in the current scenario is illegal car/vehicle parking inside their premises. Some of the fields like tolls, residences, business complexes and parking spaces can make use of the number plate recognition system for updation. In addition to this, the smart number plate recognition system can be extended for roads specifically in traffic zone and prohibited area that requires police patrolling, where the movements can be monitored. It can also monitor the vehicles that violate the traffic rules and can be extended such that the unauthorized vehicles can be recognized and maintaining a record of registration number of such vehicle.

It is used to replace the manual process of writing the registration number of the vehicle in the computer. In order to obtain an appropriate recognition, the smart number plate recognition system that comprises of three main topics. Firstly, the number plate has to be located correctly from the given input. Secondly, the located number plate undergoes segmentation of characters. Finally, the characters are recognized using Optical Character Recognition with the help of trained machine learning model. Determining the exact location of the number plate in the captured image (from the input) is the basic step. The unwanted spots are removed from the area where the number plate is located by using the connected component analysis algorithm. The developed system automates the process of identifying vehicle's registration number, making it fast, and effective in terms of cost and performance.

II. RELATED WORK

Various approaches for the detection, localization and classification of different objects in images have been proposed in the literature. The problem with the Automatic Number Plate Recognition system is about recognizing the license plate number from images of a vehicle. The solution requires comprehensive image processing that involves identification of artifacts and pattern recognition. Some of the challenges in recognizing correct numbers are plate variations across countries, plate location within the image, different sizes, different color combinations, varying font sizes, inclination and plate contamination due to other frames, images or screws. Besides image variations, the amount of light and background can worsen the problem further.



Let us consider the various approaches used in the existing system.

1) Tesseract Engine and OpenCV

This methodology uses adaptive threshold for highlighting the characters and remove the context. A component algorithm is first applied to the transformed binary image from the original image, in order to remove unnecessary image spaces. A special algorithm called Image Scissoring is used that makes use of Tesseract, an Optical Character Recognition Engine, which returns ASCII to the license number. The entire system was developed with the help of OpenCV.

2) Neural Networks

This methodology implements the method of forward context feed for classification of characters. Backward-propagation algorithm is used to build the neural network. The pre-processing measures include standardization, scale and edge detection. The horizontal and vertical graph and component survey can tackle the character fragmentation problem.

3) MATLAB

MATLAB provides an approach which is based on Sobel Edge's detection system and efficient morphological process. This approach is simplified by using the Surround Box method to separate both letters and numbers used in the number pad. The template matching methodology is used to recognize numbers and characters after the template is fragmented. The entire system was implemented with MATLAB.

Table1 : Literature Review

S. No	Title	Pros	Cons
1	<i>Hilario Seibel</i> - Eyes on the Target: Super-Resolution and License-Plate Recognition in Low-Quality Surveillance Videos (2018)	- Can add any climate. - Not littered with factors like light, speed, font size and designs.	- Inadequate accuracy - Complicated processing
2	<i>Anumol Sasi</i> - Automatic Car Number Plate Recognition (2017)	- The system is developed for accident prevention and road safety.	- Expensive - Difficult to maintain and install
3	<i>Rahim Panahi</i> - Accurate Detection and Recognition of Dirty Vehicle Plate Numbers for High-Speed Applications (2016)	- Image Enhancement - Efficient Character Segmentation and Recognition	- Difficulties in Blur images Broken Number Plate. - Confusions between some characters like O and D
4	<i>Hitesh Rajput</i> - An Automated Vehicle License Plate Recognition System (2015)	- Will be used for big style of license plates. - Better accuracy	- Cannot withstand severe climate. - Challenges in using OCR.
5	<i>Amir Hossein Ashtari</i> - An Iranian License Plate Recognition System Based on Color Features (2014)	- The system is predicated on colour features. - Uses colour Template Matching. - Uses speed control camera.	- Photos of vehicles taken by speed control camera, aren't standard. - Cannot cope up with all climate.

III. METHODOLOGY

The proposed system aims at detecting the number plate of a car and recognizes its characters in an efficient and cost effective way. The methodology followed in the proposed system has been explained in terms of block diagram is shown in Fig.1.

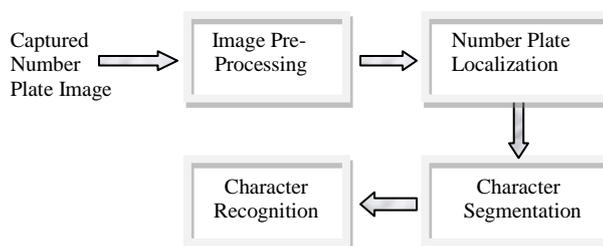


Fig.1. Block Diagram for Implementation of the Proposed System

The methodology followed in extracting the number plate in smart number plate recognition system is stated in the following steps.

- a. Detect License Plate from the given dataset.
- b. Perform segmentation of characters.
- c. Train a Machine Learning model to predict characters.
- d. Prediction of characters in License Plate.

The proposed system comprises of four basic modules which are explained below.

1) Number Plate Localization

This is the first and the most important stage of the system. This is the stage where the position of the license plate is determined. The input at this stage is an image or video footage of the vehicle and the output is the number plate of the vehicle. If the input is a video footage it undergoes preprocessing. Here the video footage is converted into frames and depending upon the time and length of the video, the frame with clear picturization of car is selected and fed as input.

Factors to be considered for detecting the number plate

- Number Plate size
- Number Plate location
- Number Plate background
- Screw

The approach used to segment the image is Connected Component Analysis (which was explained in the upcoming section). Initially, the car image is converted into a gray scale image where each pixel is between 0 & 255. The gray scale image is further converted into a binary image in which a pixel is either complete black or white. This can be done by maintaining threshold in the pixel values of 0 (black) for all pixels in the input image with luminance less than threshold value and 1 (white) for all other pixels. All the connected regions in the binary image are identified using the concept of CCA and mapped. The overall mapped region is marked by a rectangle and thus the number plate can be detected.

2) Character Segmentation

The number plate detected in the Number Plate Localization module is given as input to the Character Segmentation module. This is the stage, where the characters on the number plate are mapped out and segmented into individual images. For segmenting the characters, the connected component analysis algorithm was used. The extracted number plate is divided into different images, each containing one isolated character. The binary image is scanned and the pixels are labelled as per the connectivity conditions of current pixel. Each character has been resized from the previous step to the standard height and width in order to be used in the recognition process. The CCA algorithm is used for better performance and it is very useful for automated image analysis.

The steps that are used to segment the characters of the number plate are as follows:

- Extend the contrast of the gray scale image over the entire range of gray levels available (0-255).
- Threshold the number plate image.
- Search for connected components in the image in which, each of them will be assigned a special label in order to distinguish between different connected components in the image.

- Each character from the previous step has to be resized to the standard height and width in order to be used in the recognition process.

3) Training a Machine Learning Model

For recognizing the characters, we have to use the concept of machine learning. Machine Learning can be briefly stated as the branch of AI that deals with data and processing of data to discover pattern that can be used for future prediction. It can be further classified into supervised and unsupervised learning. In supervised learning, the known (training) dataset is used to make predictions. The training datasets have been collected from various sources and the model is trained with the help of the datasets collected.

The machine learning model is trained using the Support Vector Classifier which is also called as the support vector clustering algorithm has the following characteristics.

- Fits to the data we provide.
- Partition the datasets and group them accordingly.
- Clusters can be defined as connected components of the graph.
- Designed for pair-wise comparisons and produce accurate estimates.

The trained model undergoes 4-fold cross validation to determine the accuracy of the model. Then the model can be persisted to a file and can be used for prediction.

4) Character Recognition

This was the last module of the Smart number plate recognition system that makes use of the trained Machine learning model. It uses the concept of supervised learning that makes use of the training datasets.

The basic requirements for the recognition of characters are listed below.

- Collecting the required training datasets.
- Choosing a supervised learning classifier.
- Training a machine learning model.
- Validating the trained model and determining its accuracy.

All of these requirements are satisfied in the previous module, i.e. Training a Machine Learning Model module. Among various available classifiers, Support vector classifier has been used because of its better performance.

Some of the algorithms and techniques used for the development of the proposed number plate recognition system are explained as follows.

1) Connected Component Analysis

The algorithm used all-over the implementation of the proposed system is the Connected Component Analysis algorithm. It is an algorithmic application that makes use of the graphical theory. The algorithm is used for detection and segmentation of the number plate from the vehicle's image in the Number Plate Localization phase, for segmenting the characters from the detected number plate in the Character Segmentation phase and for recognizing the segmented characters in the Character Recognition phase.

In connected component analysis, all the connected pixels that belong to the same object are collectively known as connected regions. The pixels that have same value and are adjacent to each other are connected, which results in connected pixels. Instead of Connected Component Analysis, approaches like edge detection and morphological operations can be explored. Basically, CCA has added advantage on grouping and labelling connected regions on the foreground.

A graph is constructed from the input data provided, containing the vertices and connecting edges as parameters. The vertices comprise of information required for the comparison heuristic and the edges point out the connected neighbours required for the graph that can be traversed later. The labelling of vertices is based on the connectivity and relative values of their neighbours. The working of the algorithm can be explained in the following steps.

Step 1: Start

Step 2: Get the pre-processed input (image/video frame)

Step 3: Scan the binary digital image pixel by pixel.

Step 4: If pixel is not background, check neighbours

a) If neighbours are already labelled, assign neighbour's parent label to main label.

b) Is none of the neighbours is labelled, then assign new label to pixel.

2) Support Vector Classifier

Support vector classifier is a supervised machine learning algorithm that predicts the future outcome based on the training datasets. The developed system makes use of supervised learning because the datasets required for the number plate recognition system is already known, since the vehicle’s number plate only consists of alphanumeric characters, i.e. uppercase alphabets (A-Z) and digits (0-9). Support vector classifier is used because of its high accuracy when compared to other classifiers.

The main objective of the Support Vector Classifier is to perform segregation over the given dataset in an optimized way that acquires the best possible outcome. The margin is defined as the distance between the either nearest points. While selecting the hyper-plane, a hyper-plane with the maximum possible margin between support vectors in the given dataset should be selected. The selection process can be described in the following steps:

1. Generate hyper-planes for optimized segregation of the classes.
2. Among the generated hyper-planes, select the right hyper-plane with the maximum segregation from the either nearest data points.

Thus the algorithms used for the implementation of the proposed system are explained briefly.

IV. EXPERIMENTAL RESULTS

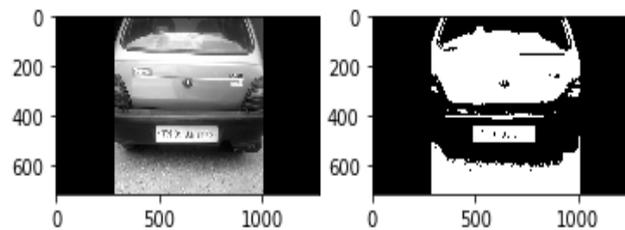
The Figures given below show the results of number plate detection from an input image/video and recognition of characters from the detected number plate.

Figs. 2 (a), (b), (c) shows the result obtained from the Number Plate Localization phase. Figure (a) is the image of the car, obtained after its gray-scale and binary image conversion. Figure (b) is the result obtained after applying the Connected Component Analysis Algorithm for identifying and mapping the connected components. Figure (c) is the input image with the number plate detected.

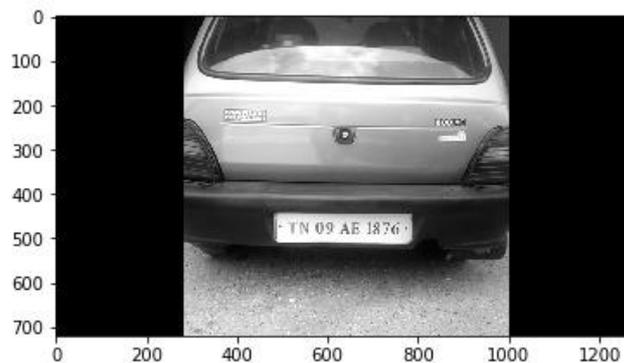
Fig. 3 shows the result obtained from the character Segmentation phase, where each character of the detected number plate are segmented.

Fig. 4 shows the 4- fold cross validation result obtained from Training the Machine Learning Model Phase which is used to determine the accuracy of the model.

Fig. 5 shows the final outcome i.e. the successfully predicted characters of the detected number plate image.



(a)



(b)

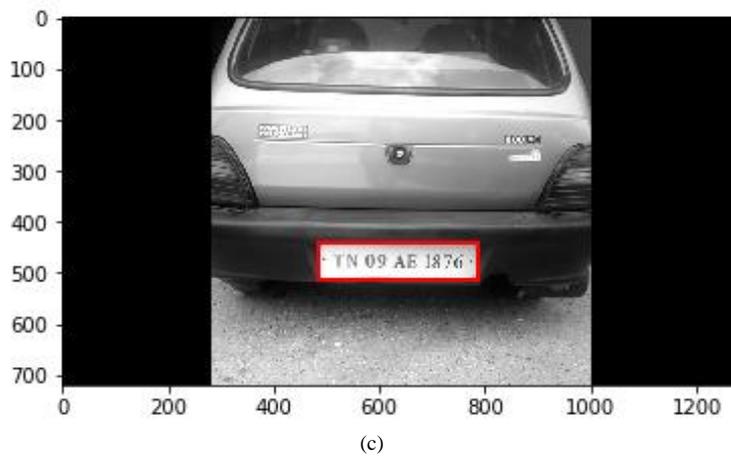


Fig.2. Number Plate Localization (a) Gray scale and Binary image (b) Image after applying CCA algorithm (c) Image after detecting the number plate of the image

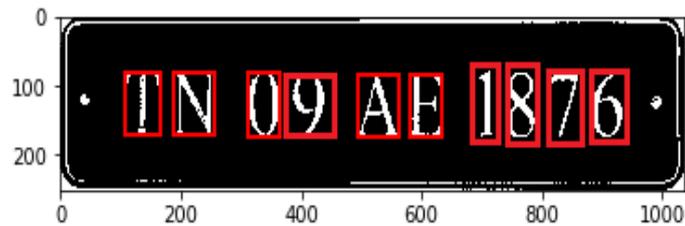


Fig. 3 Character Segmentation - The CCA algorithm is used where the binary image is scanned and the pixels are labelled as per the connectivity conditions of current pixel.

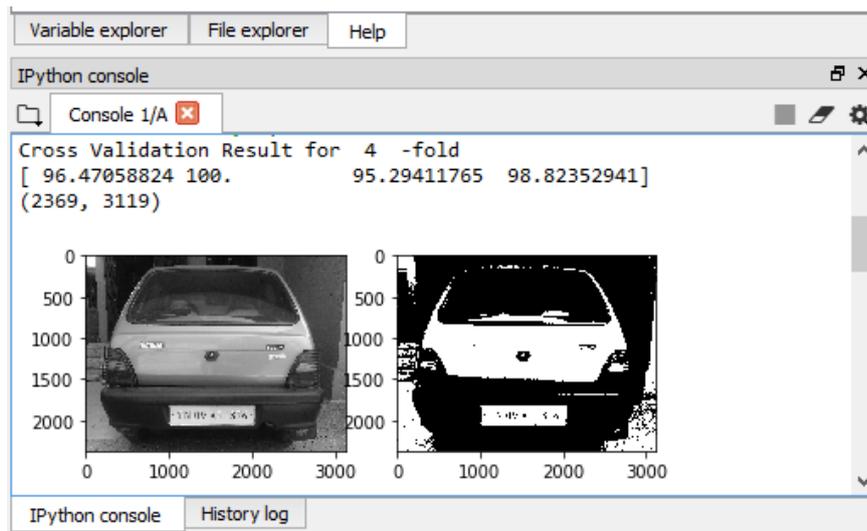


Fig. 4 Training a Machine Learning Model – The model undergoes 4-fold cross validation to determine its accuracy.

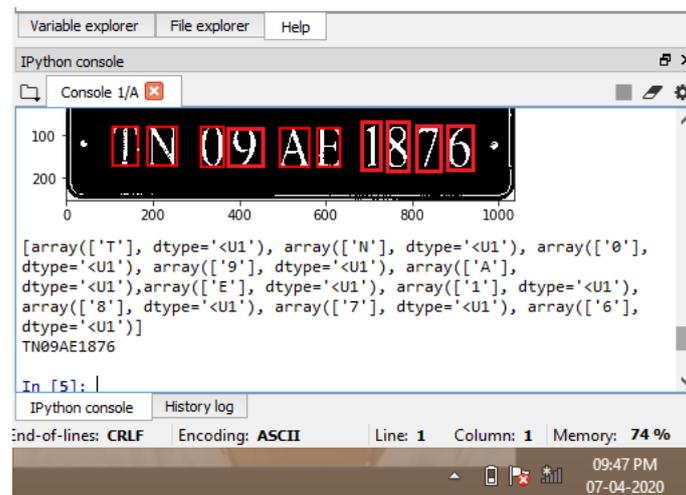


Fig. 5 Character Recognition – The final outcome of the proposed system where the characters predicted from the number plate image is displayed in textual form.

V. CONCLUSION

In this paper, an efficient, cost-effective, Smart Number Plate Recognition System enhanced by Machine Learning has been presented to provide an effective system for detecting and recognizing the number plate for uniquely identifying the vehicles. Each stage of the proposed system that includes number plate detection, Character Segmentation and Recognition has been implemented effectively using the Connected Component Analysis Algorithm and trained machine learning model. The Machine learning model has been trained using the Support Vector Classifier with the help of the datasets collected and tested using 4-fold cross validation which acquires nearly 98.53% overall accuracy in training the supervised learning model. In the detection phase, contours were obtained from the image and analyzed for representing a possible number plate. Several values representing boundary condition for the possible number plate contours are hardcoded. After applying the necessary boundary conditions an accuracy of 92% was achieved for number plate detection phase. An overall accuracy of 89% is achieved for the proposed system. The algorithm used here accelerates the process. In addition, it also improves the number plate detection and extraction of characters, under certain conditions. Accuracy of classifier is mainly affected due to confusion between '0', 'O' and '1', 'I' in case of blurred and low-quality images. The accuracy can be improved by training the classifier with larger dataset of various types of character images.

REFERENCES

- [1] Hilario Seibel, Siome Goldenstein, Anderson Rocha.: Eyes on the Target: Super-Resolution and License-Plate Recognition in Low-Quality Surveillance Videos - IEEE Paper, 2018.
- [2] Anumol Sasi, Swapnil Sharma, Alice N. cheeran: Automatic Car Number Plate Recognition - IEEE paper, 2017.
- [3] Rahim Panahi, Iman Gholampour.: Accurate Detection and Recognition of Dirty Vehicle Plate Numbers for High-Speed Applications - IEEE Paper, 2016.
- [4] Koshti Sanket, Prof. R.A.Patil.: Review of Car-License- Plate Detection Methods, Global Journal Of Engineering Science And Researches, ISSN 2348 – 8034, April 2016.
- [5] Hitesh Rajput, Tanmoy Som, Soumitra Kar.: An Automated Vehicle License Plate Recognition System - IEEE paper, 2015.
- [6] Divya.K.N, Dr.Ajit Danti.: Recognition of Vehicle Number Plates and Retrieval of Vehicle Owner's Registration Details, International Journal of Innovative Research in Technology & Science, ISSN:2321-1156, 2015.
- [7] Ganesh R.Jadhav, Kailash J. Karande.: Automatic Vehicle Number Plate Recognition for Vehicle Parking Management System - IEEE paper, 2014.
- [8] Muhammad Tahir Qadri, Muhammad Asif.: Automatic Number Plate Recognition System For Vehicle Identification Using Optical Character Recognition, Published in International Conference on Education Technology and Computer, ISBN - 978-0-7695-3609- 5/14, 2014.
- [9] Dhiraj Y. Gaikwad, Pramod B. Borole.: A Review Paper on Automatic Number Plate Recognition System, International Journal of Innovative Research in Advanced Engineering (IJIRAE) Volume 1 Issue 1, April 2014.



- [10] Mr. G. T. Sutar, Prof. Mr. A.V. Shah.: Number Plate Recognition Using an Improved Segmentation, International Journal of Innovative Research in Science, Engineering and Technology Volume 3, Issue 5, May 2014.
- [11] R.Parisi, E.D.Diclaudio.: Car Plate Recognition by Neural Networks and Image Processing, Published in International Journal of Engineering Science and Advanced Technology Volume-2, Issue-3, 1736 – 1452 Feb 2014.
- [12] Priyamvadha Krishnakumar.: Automatic Number Plate Recognition (ANPR) through smart Phones using Image Processing Techniques, IOSR Journal of VLSI and Signal Processing, Volume 4, Issue 4, Ver. III, August 2014.
- [13] M. S. Sarfraz, A. Shahzad, M. A. Elahi, M. Fraz, I. Zafar, and E. A. Edirisinghe.: Real-time automatic license plate recognition for CCTV forensic applications, J. Real-Time Image Process., vol. 8, no. 3, pp. 285–295, Sep. 2013.