



# IJRASET

International Journal For Research in  
Applied Science and Engineering Technology



---

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

---

**Volume: 10    Issue: IV    Month of publication: April 2022**

**DOI: <https://doi.org/10.22214/ijraset.2022.41198>**

**[www.ijraset.com](http://www.ijraset.com)**

**Call:  08813907089**

**E-mail ID: [ijraset@gmail.com](mailto:ijraset@gmail.com)**

# Vehicle Number Plate Recognition System Using TESSERACT-OCR

Akhil Songa<sup>1</sup>, Rahul Bolineni<sup>2</sup>, Harish Reddy<sup>3</sup>, Sohini Korrapolu<sup>4</sup>, Vani Jayasri Geddada<sup>5</sup>

<sup>1, 2, 3, 4, 5</sup>Department of Computer Science and Engineering, GITAM

**Abstract:** *With the increase in the number of vehicles, automated systems to store vehicle information are becoming increasingly necessary. Communication is critical for traffic management and crime reduction, and it cannot be overlooked. Automatic vehicle identification using number plate recognition is a reliable method of identifying vehicles. It requires a lengthy time and a lot of practice to develop satisfactory results using present algorithms that are based on the idea of learning. Even so, accuracy is not a significant concern.*

*It has been devised as an efficient approach for recognizing vehicle number plates, which is included in the suggested algorithm. The technique is intended to address the difficulties of scaling and recognition of the position of characters as long as the accuracy is maintained. Automatic Number Plate Detection is a unique application in Machine Learning as it detects images and converts them to text form.*

*The algorithm detects and captures the vehicle image and extracts the vehicle number plate using image segmentation. The extracted image is later sent to optical character recognition technology for character recognition. This system is implemented in areas like traffic surveillance, military zones, apartments, etc.*

**Keywords:** *Optical Character Recognition, tesseract ocr, matplotlib, Number Plate Recognition.*

## I. INTRODUCTION

The rapid increase in the number of automobiles on the road began in the late 1900s. Due to the high number of vehicles on the road, license plate identification has become a time-consuming operation, as the information received from license plates is largely used in traffic monitoring, parking, border monitoring, law enforcement, and other applications, among other things. This recognition challenge can be divided into five steps, each of which is discussed below.

Picture recognition, image preprocessing, license plate localisation, character segmentation, and optical character recognition. The process of capturing an image of a license from a vehicle image is called image recognition. The process of brightening, normalizing, and modifying a picture is called image preprocessing. The process of finding and recognizing symbols in a picture is called image segmentation.

And the technology used for image to text conversion is OCR. Where traffic norms and standards differ from one country to the next, each country has its own set of guidelines to follow. In turn, this results in the algorithm being tuned to a specific region. For example, there are 12 different types of fonts that can be used on Indian license plates, and there are even more options for license plates from other nations.

The general structure of an Indian license plate consists of two letters for the state code, three letters for the district code, and four digits for the vehicle code. In contrast, in countries such as the United States of America, there is no standard when it comes to fonts, and they are designed with a great deal of variation in the contrast between text and background. In cases where the number plate and the background are nearly identical, it is difficult to determine the location, and variations in the brightness and contrast of the image also result in inaccurate results. The use of morphological techniques helps to get the right information and get accurate results.

The work is broken into numerous sections, which are as follows:

- 1) Raw image should be used as input.
- 2) Image binarization is the process of dividing an image into two halves.
- 3) Using the mid-filtering method, you may reduce noise.
- 4) The histogram equalizer can be used to improve contrast.
- 5) Localization of the plate Segmentation of characters

## II. LITERATURE SURVEY

Vehicle detection has become an integral part of the digital world for finding the penalties of the vehicle. A collection of studies have been performed to classify the vehicles such as bicycles, trucks, cars, etc. The Optical Character Recognition (OCR) technique was employed, which is a popular utilized mechanism that converts scanned pictures of handwritten or printed text into machine encoded text and is a commonly used technology. It is suggested here to employ a feed-forward neural network to perform optical character recognition. Two unique actual image datasets of the characters are often used for testing and training the claimed neural network. A total of two unique picture data-sets were employed to simulate outer-world circumstances that the neural network will be exposed to throughout the testing phase. Artificial Neural Networks (ANNs) are a type of intelligent computer architecture that is commonly employed in pattern recognition applications. The multilayer feed-forward neural network is the most often used artificial neural network (ANN), and contains the fundamental internal design that can sort out the given inputs to a list of goal categories. To organize the inputs of neural networks, the tasks performed in and typically use Feature extraction and Binary pixel; the prior is the regularly executed method for neural networks and results in accurate performance even in the tough conditions; the latter is the latest method for neural networks and accomplish a great results in the arduous environments. Although, feature extraction is typically a time-consuming process that requires extensive computing or numerous phases to complete. Similar approaches are available that employ additional processes in the training phase or later receiving the outcomes of a neural network in order to deal with tough characters that are members of ambiguous character sets, and these methods are also available. Specifically, for Confusing characters (for example, I/1, B/8 and O/D), additional training is employed. In addition, the comparison of distinct sections of ambiguous characters is conducted in . Single stage classifiers and multistage classifiers are two types of statistical classifiers that may be distinguished. Character characteristics are recovered from the elastic mesh in the work , and the full address character string is used as the target of investigation. The SVM (Support Vector Machine) program was tested on Japanese vehicle number plates, and the attributes to detect Kana (Japanese script), numbers, and the group of characters that describe the area were used to train the SVM integration. The identification percentages for strings of characters, numerals, Kana, and are 97.8 percent, 99.5 percent, 98.6 percent, respectively, according to the results. A two-stage hybrid optical character recognition system is introduced with the goal of improving the recognition rate.

## III. PROCESS OF ANPR

### A. Image Acquisition

Input image captured through camera.



Fig. 1 Captured Image

### B. Image Pre-processing

- 1) *RGB To Gray Conversion*: The use of colour images does not aid in the identification of significant edges and other characteristics. Processing an RGB image is complicated and takes a long time, therefore we must first convert the coloured image to a grayscale image.



Fig. 2 RGB to Gray Conversion

- 2) *Image Enhancement*: The purpose of adaptive histogram equalisation is to improve the contrast of a picture (gray colour image). In this step, we will create many histograms, each representing a different part of the image. In contrast to a typical histogram, which shows just one histogram for the whole image, this one shows the complete image.
- 3) *Noise Removal*: To remove noise in the image.

### C. Edge Detection

An edge is a line drawn between two places that have grey level qualities that are significantly different from one another. It is capable of detecting discontinuities in the intensity levels. The first stage in the recognition of a plate is to determine the size of the plate (which is a rectangle); thus, we must determine the edge of a rectangular plate. The edges of the image are emphasised with the help of the sobel operator. As a result, the quantity of data in the image is reduced, and the necessary data is processed for subsequent use.

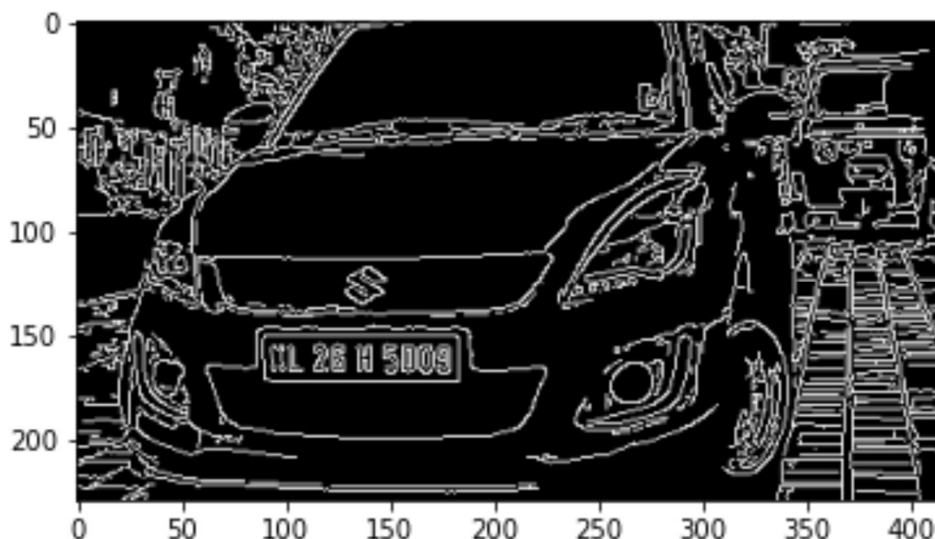


Fig. 3 After edge detection

#### D. Morphological Image Processing

The output of the structuring element must be the same size as the input. By using dilation and by adding pixels to the object's border, we may make the boundaries appear thicker than they actually are. Using the Shrinking process, thinning the picture to remove any unnecessary bits is accomplished

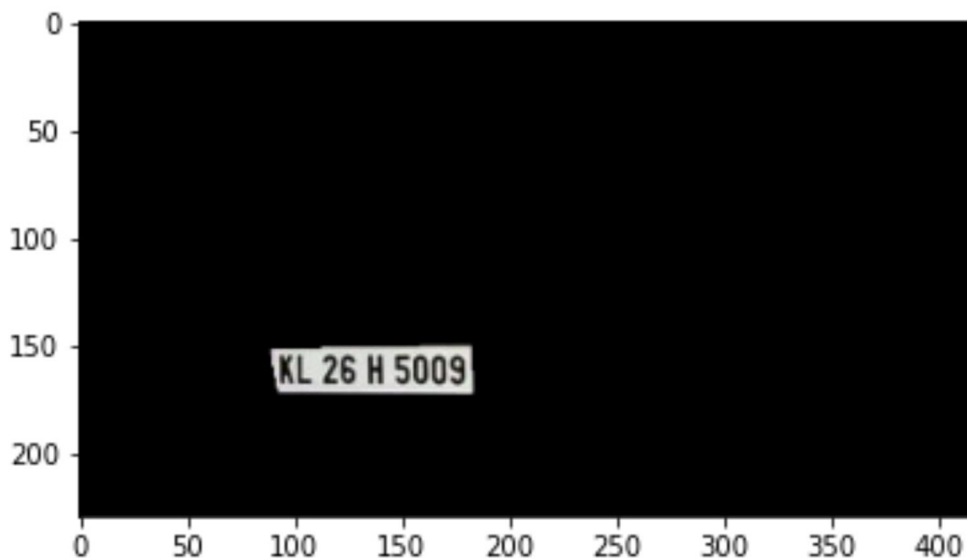


Fig. 4 Number plate recognized

#### E. Threshold

If the specified threshold value is exceeded or not, two separate levels are assigned to pixels that fall above and below the threshold value. In order to distinguish an item from a background image, the image is transformed to binary form. The technique of determining the grey level threshold is straightforward. To compare pixels in a picture, the value of the threshold (T) is determined and used. Aside from that, it converts the input binary image (K) into an output binary image (F) that is then segmented. Global threshold is used to segment the histogram of an image using a single threshold value, as opposed to local threshold. When we talk about threshold, we are referring to the volume of grey level that falls between the baseline border, which is located between the pixels found in the foreground and the background..

#### F. Segmentation

In between the number plate extraction and character recognition, character segmentation serves as a transitional step. There are several segments for the different characters on a number plate region. Various factors, such as illumination variations, plate frames, and rotation, all contribute to the inability to complete the segmentation task. A segmentation approach is sometimes referred to as a boundary box analysis in some circles. Characters are allocated to related components using this approach, and these characters are then extracted using the boundary box analysis. When the amount of noise in the image has been reduced, the segmentation procedure is complete.

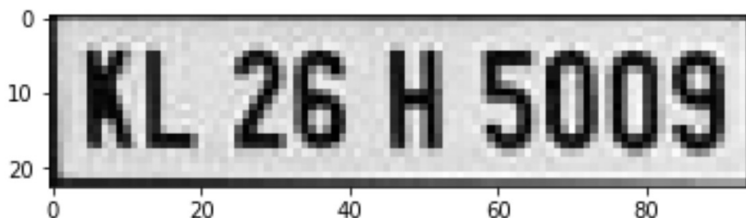


Fig. 5 Number plate segmented image

**G. Character Recognition**

Character recognition is accomplished by employing feature extraction to extract the characteristics of characters and their different categorization procedures, which is a step in the character recognition process. The identification of characters from the number plate is accomplished through the use of a machine learning system.

**Text: KL 26 H 5009**

Fig. 6 PyTesseract OCR output

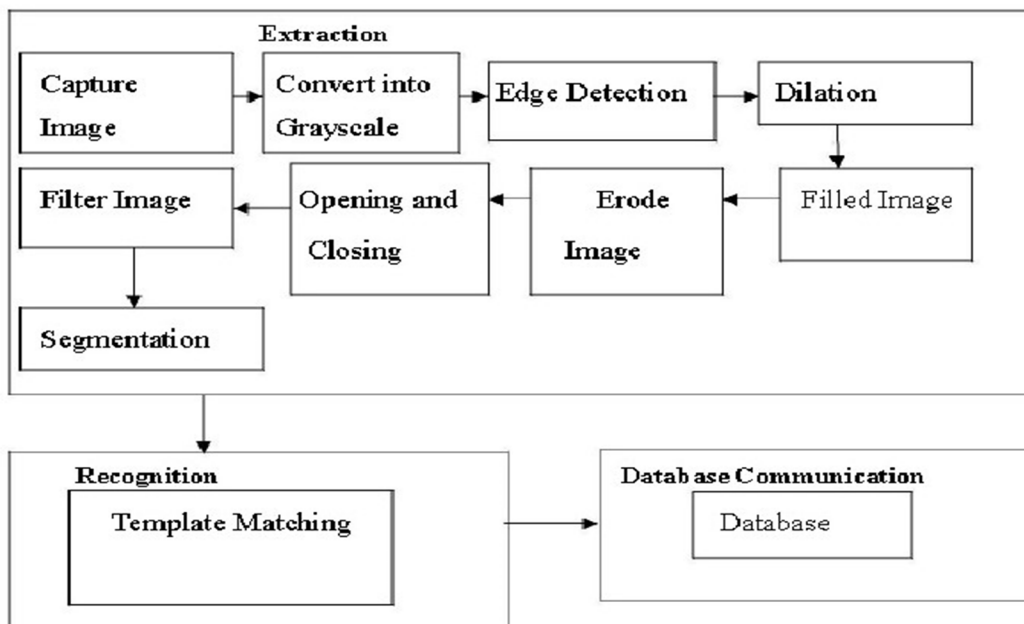


Fig. 6 Flow chart of Automatic Number Plate Recognition

**IV. CONCLUSIONS**

After reviewing a number of studies, we've come to the conclusion that there are a number of methods for identifying a vehicle's license plate. In conjunction with the Sobel edge detection method and automatic license plate recognition for detecting edges and fill spaces fewer than 8 pixels, employed to identify and recognize automotive license plates.

In order to identify the car, the system uses a variety of image processing algorithms that are stored on the computer's hard drive. In order to assess the program's capabilities, Matlab is used, and photo data from around the world is used. According to the counterfile findings, there are a wide variety of illumination situations in which the system may identify and recognize a car using its license plate. In addition, statistical analysis can be used to assess the likelihood that a vehicle's license plate will be observed and recognized.

**REFERENCES**

- [1] Optasia Systems Pte Ltd, "The World Leader in License Plate Recognition Technology" Sourced from: www.singaporegateway.com/optasia, Accessed 22 November 2008.
- [2] J. W. Hsieh, S. H. Yu, and Y. S. Chen. Morphology based license plate detection from complex scenes. 16th International Conference on Pattern Recognition (ICPR'02), pp. 79-179, 2002.
- [3] V. Kasmat, and S. Ganesan, "An efficient implementation of the Hough transform for detecting vehicle license plates using DSP's," IEEE International Conference on Real-Time Technology and Application Symposium, Chicago, USA, pp. 58-59, 2005.
- [4] S.H. Park, K.I. kim, K. Jung and H.J. Kim, "Locating car license plate using Neural Network," Electronic Letters, Vol. 35, No. 17, pp. 1474 - 1477, 1999. [5] K.K. KIM, K.I., KIM, J.B. KIM, and H.J. KIM, "Learning-Based Approach for License Plate Recognition" Proceeding of IEEE Signal Processing Society Workshop, Vol. 2, pp.614-623, 2000.
- [5] FAOUR, REHAM & Shanwar, Bassel & Zarka, Nizar. (2016). RECOGNITION OF VEHICLE NUMBER PLATE USING MATLAB. 10.13140/RG.2.1.1459.8640.



10.22214/IJRASET



45.98



IMPACT FACTOR:  
7.129



IMPACT FACTOR:  
7.429



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24\*7 Support on Whatsapp)