Name Plate Recognition of Indian Cars Using OCR

1st Shad Been Sharif Sami
Department of Computer Science and Engineering (CSE)
School of Data and Sciences (SDS)
Brac University
Dhaka, Bangladesh
shad.been.sharif@g.bracu.ac.bd

2nd Tasfia Jahan
Department of Computer Science and Engineering (CSE)
School of Data and Sciences (SDS)
Brac University
Dhaka, Bangladesh
tasfia.jahan.nahin@g.bracu.ac.bd

3rd Shihab Sharar
Department of Computer Science and Engineering (CSE)
School of Data and Sciences (SDS)
Brac University
Dhaka, Bangladesh
shihab.sharar@g.bracu.ac.bd

4th Annajiat Alim Rasel
Senior Lecturer
Department of Computer Science and Engineering
Brac University
Dhaka, Bangladesh
annajiat@bracu.ac.bd

5th MD. Mustakin Alam
Department of Computer Science and Engineering (CSE)
School of Data and Sciences (SDS)
Brac University
Dhaka, Bangladesh
md.mustakin.alam@g.bracu.ac.b

***Abstract*—Cars can have different kinds of number plate which mainly carries the license number of that vehicle . This paper mainly represents how the number plates can be easily recognized and can be verified in case of any rule violation cases or on any fatal street collision case or in any car robbery cases . Bangladeshi car owners has the tendency of changing the typeface and design of the number plate . This paper is mainly going to focus on designing a recognition system which will be able to detect any kind of number plates . To increase the effectiveness of scanning the number plates we have used Optical Character Recognition (OCR) in this case . In this paper we mainly focused on detecting Indian number plates using OCR methods and analyzed our data with the Indian car license number datasets .**

***Index Terms*—Optical Character recognition (OCR) , Automatic Number Plate Recogniton (ANPR) , , Digital Signal Processing (DSP) , feature Based Number Plate recognition**

I. Introduction

The easiest way to uniquely identify a car is by using its registered license plate. The License plate is usually a piece of metal, rectangular in shape with a mixture of alphanumeric characters that form a uniquely identifiable ID. License plates are crucial in the identification of vehicles for several use cases, such as, parking, border control, detecting and monitoring traffic, travel and airport parking. Although it is a known fact that license plates are faked or swapped, it is still a global effort to use license plates as a means of uniquely identifying a vehicle. As personal home security has become crucial over the years, this concept holds significance not only in state matters but also in the realms of personal security. Automatic Number Plate Recognition (ANPR) technique has garnered fame in recent years with the progress in camera, storage and most importantly processor technology. English being one of the most common languages being used across the world is also often used for license plates in most countries. To simplify the process, a License Plate Recognition(LPR) system consists of a camera and a computer, where the computer runs a software designed to identify a vehicle in 4 steps. The primary task of this software is to identify the vehicle using its Vehicle Identification Number (VIN), this is split into 4 distinct segments, capturing the vehicle's image, detecting the number plate itself, character segmentation and lastly, character recognition.The fourth component makes use of optical character recognition (OCR) and You Only Look Once (YOLO3) technique for extracting characters from an image and then matches with a database that has information on every alphanumeric character, to identify the individual characters. The first stage, which is to take a picture of the car, may seem quite simple, but it is actually a very demanding operation as taking a picture of a moving vehicle is very hard . Also there are more issues like dust and dirt on number plates , damaged number plates can generate blur images . Even in some cases the numbers also becomes blur . This kind of issues are needed to be taken into consideration .

In this paper we have used - Optical Character Recognition (OCR) , MATLAB, color segmentation, online license plate matching - these techniques for our ANPR .

II. Literature Review

Number Plate recognition is necessary to identify those cars which has break traffic rules . There are many approaches

for number plate recognition which mainly differs depending on the characters used in that language . For example in [1] Omran has implemented an automatic license plate identification system for detecting Iraqi car license plates . He mainly worked with three types of number plates according to the size of them and he did plate identification by segmenting each numbers and characters into separate images using optical character recognition (OCR) system .

In [2] Laroca have tried to construct a trustworthy detection and recognition system for transitional, standard car license plates, which are frequently observed in emerging nations. By using cutting-edge technology, such as machine learning (ML) models, it is possible to improve the readability of license plates that are designed or printed in a variety of fonts and styles. They have used the latest version of You Only Look Once (YOLO) which is an object direction approach in their work. The license plate placement is decided upon after the raw image has undergone pre-processing to improve its quality and is separated into grid cells of the proper size. After post-processing the data, industry-recognized criteria are used to determine the accuracy of the proposed model.

In [3] Quadri have used OCR for recognizing characters. To determine precise details like the owner of the car, the location where it was registered, its address, etc., the obtained data is then utilized to compare with the entries in a database. Using actual photos, the system's performance is evaluated after being developed and simulating it in Matlab. From the experiment, it can be seen that the created system correctly locates and identifies the car number plate on real photos.

In [4] For extracting characters, Shidol has worked on character segmentation, character identification, and number plate extraction.The Sobel filter, morphological methods, and linked component analysis are used to extract license plate numbers. Character recognition is accomplished using a support vector machine (SVM).The segmentation accuracy in this case is 80%, while the recognition rate is 79.84%.

In [5] a computer technology is used to show the number plate recognition method. Digital gray-level photographs of cars are thresholded using the Niblack algorithm, which was found to perform better than other binarization techniques previously utilized in similar systems. A simple yet very effective rule-based approach is used to detect the position and size of number plates. A special dimension reduction method is used to condense the 225 features in the neural network inputs to just 50 features. Each of the six parallel small networks can recognize six different characters. The device can identify single and double line plates in a variety of lighting conditions and with a slight rotation. 86.1% of all license plates can be successfully identified through their approach .

In [6] the challenge of distinguishing number plates is taken into account in India, where the norms for number plate standards are hardly observed. The system combines techniques made specifically for Indian number plates, such as "Feature-based number plate Localization" for locating the number plate, "Image Scissoring" for character segmentation, and "Statistical feature extraction" for character identification. The device can identify single and double line license plates in a range of illumination conditions with a success rate of about 82%.

In [7] ANPR and its practical applications are included . This also uses the OCR technique. The accuracy of the ANPR system was found to range from 75% to 85% for Indian license plates .

In [8] the author uses the BAM (Bi-directional Associative Memories) technique for number plate identification. All methods for number plate recognition share the broad processing steps listed below. Trigger, Image Capture, Vehicle Presence, Finding Plate, Character Segmentation, and Recognition Process were the stages that were taken. Using edge detection, color detection, and signature analysis, the license plate was discovered. The characters were recognized using the OCR technique.

Edge-based texture analysis is used in this research [9] to discover candidate plate regions. Using an edge detector on the plate area, search for straight lines to locate the plate's edges. Plate rectification, segmentation, Go to Grayscale glyphs after detecting text blobs. Using a feed-forward artificial neural network, OCR is carried out.

The image filtering for the license plate is shown in this study [10] after the picture acquisition using the camera (RGB). The ANPR algorithm's next step is to segment the license plate in a picture. After that, the characters are identified using OCR.

III. Methodology

The vehicle detection module's inductive sensor, which detects the presence of a vehicle, is a metal wire loop embedded beneath the road. When a vehicle crosses the loop, the induced current modifies to reveal the existence of the vehicle. The DSP is therefore stopped, which prompts the IR camera to begin recording. The acquired image is processed byDigital Signal Processing (DSP) to determine the vehicle's license plate using the various image processing methods mentioned before. The DSP delivers the license number in ASCII format, which is used to access the entire database including all relevant data on the vehicle. Fig.1 depicts every phase in detail.

A. *Image pre-processing*

Several techniques, such as the multiple interlacing algorithm, Fourier domain filtering, and color image processing, are advised for number plate localisation. These algorithms do not operate for Indian license plates properly since they assume the presence of elements like a plate border, plate color, and character color. Therefore, we developed and used a "Feature-based number plate localization" [11] method that is appropriate for Indian conditions. This tactic uses a number of algorithms that were developed based on the shared traits of both characters and license plates. Using Ostu's method, the input grayscale image (fig.2) is adaptively transformed into a binary image (fig.3) for pre-processing. Compared to other methods that employ adaptive binarization like Niblack's technique , this one is more appropriate .

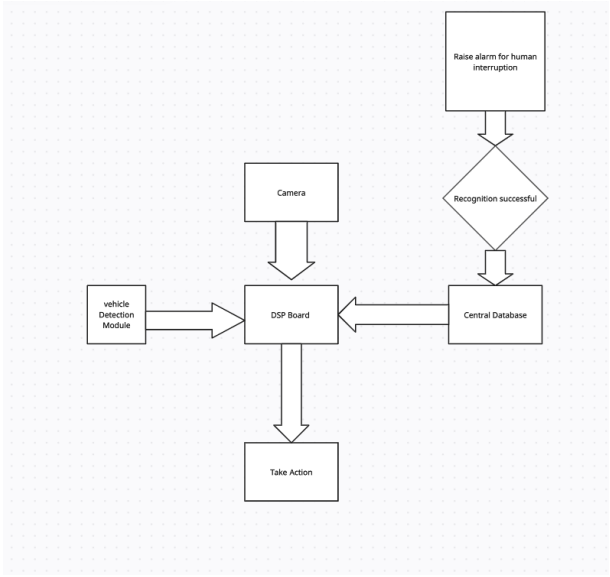

Fig. 1. flow chart

Fig. 2. gray-scale image

Fig. 3. binarized image

Fig. 4. step 1 outcome

Fig. 5. step 2 outcome

Fig. 6. step 3 outcome

Fig. 7. step 4 outcome

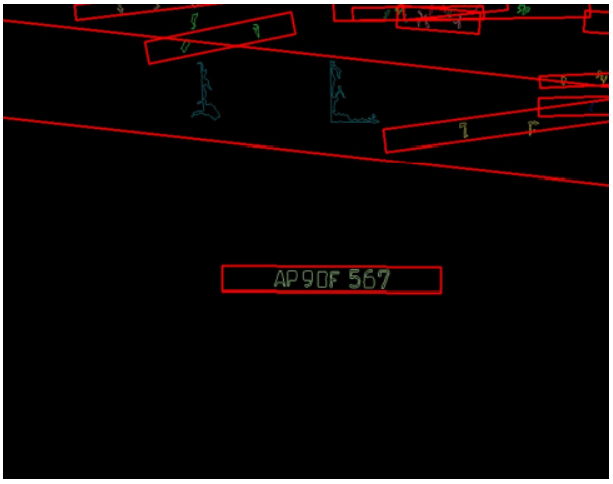

Fig. 8. step 5 outcome

*a) **Step 1**:* Rolling across the binary image is a mask with the dimensions of the largest character possible and the shape of an inverted "L." A position is considered to be a potential character location at each increment if the following conditions are met: The mask must have at least one white pixel, as must the immediately next row and column (fig 4).

*b) **Step 2**:* Calculated character sizes for those who made the cut. Locations are eliminated if they are less than 50% of the largest character size (fig.5).

*c) **Step 3**:* Each likely character's white pixel density is estimated. Only when it exceeds 40% of the total number of pixels is the location saved (fig. 6). To accomplish time optimization, all of the previous steps are completed in a single loop.

*d) **Step 4**:* The white pixel density is determined for a set of rows with a height equal to the highest number plate height that is feasible. A portion of the region is deleted if it is below a specified threshold (fig. 7).

*e) **Step 5**:* The white pixel density is determined for a set of columns whose width is equal to the widest number plate width that can be used. That region is again disregarded if it is below a particular threshold (fig. 8).

Fig. 9. step 6 outcome


```
14 possible plates found

license plate read from image = AP9BF567
```

Fig. 10. outcome of image scissoring

*f) **Step 6**:* It is computed how many characters will fit on the finalized number plate locations. The area is discarded if there are fewer than four characters. If two number plate areas are discovered in close proximity to one another and have about the same amount of characters, those areas are combined. These procedures allow you to precisely pinpoint the number plate in the image and get rid of any additional background noise. The number plate is now taken out of the input binary image (fig. 9) and eroded with squares of size 2X2 to prevent character overlap before segmentation.

B. Character Segmentation

Character segmentation is recommended using a variety of techniques, including blob coloring and the peak-to-valley method . These techniques, however, are not appropriate for Indian license plates because they take a long time and don't produce satisfactory results when the characters overlap. A novel "Image Scissoring" algorithm is designed to offer reliability and time-optimization. This algorithm copies the vertically scanned and scissored portion of the license plate into a new matrix at the row where there are no white pixels (i.e., a blank row).In order to find a blank row, this scanning method is continued, and various scissored areas are as a result obtained in various matrices. There are two possible row configurations for Indian license plates. As a result, there can only be two matrices at a time. Height comparisons are used to eliminate erroneous matrices. The previous matrix is eliminated if any matrix's height is less than one-fourth of the height of the tallest matrix. Individual characters are

segmented using width as a threshold as the same process is repeated horizontally on each matrix (fig. 10).

C. *Pre-recognition character enhancement*

Segmented characters are extracted from the input grayscale images in this step. Then, using Ostu's technique, each character is binarized adaptively. The binary character is then scissored in place. These actions aid in streamlining the subsequent recognition procedure .

D. *Character Recognition and Syntax Checking*

The ANPR system's most crucial phase is now. Character identification can be done directly against templates . For typeface variants, which are frequently seen on Indian license plates, this method, however, has a relatively low success rate. The characters can be categorized using artificial neural networks such as BPNNs. However, they don't offer hardware or time optimization. Therefore, statistical feature extraction has been applied.This method extracts fourteen traits from each of the twelve equal sections that make up the character at the beginning. Binary edges (2X2) of fourteen different sorts make up the characteristics. To choose the correct character, the greatest value of correlation between the feature vector created in this manner and the feature vectors of all the stored templates is determined . Finally, syntax checking is carried out to make sure that any fake characters are not interpreted as a legitimate license number.

IV. Indian License Plate Dataset

In this paper we introduce an Indian Number (licence) plate dataset [12] with 16,192 images and 21683 number plates . Among them 472 images were picked in order to train the model. 183 went into validation and the rest of 190 went into testing. The pre-processing step includes resizing and feature extraction. No further augmentation was put in before the training process.

TABLE I
ACCURACY RATE

Result Analysis	
Process Steps	Success Rate
Character Recognition	85%
Character Segmentation	95%
Number Plate Localization	87%

V. Result Analysis

A series of Indian car photographs with vastly different lighting conditions that were not used during testing were used to test the system. Average processing time for recognition is 2 seconds. The code can be optimized to make this even better. 82% of the plates were properly identified when non-English number plate lettering and severely deformed number plates are eliminated. The performance of each portion is as follows: 85% for character recognition, 95% for character segmentation, and 87% for number plate localisation (Table -1). The accuracy of each character is given in Table-2 .

TABLE II
CHARACTER ACCURACY

Accuracy percentage of Characters	
Character	Cumulative%
A	7.7%
P	14.8%
9	21.9%
B	27.8%
F	33.1%
5	37.9%
6	42.6%
7	47.3%

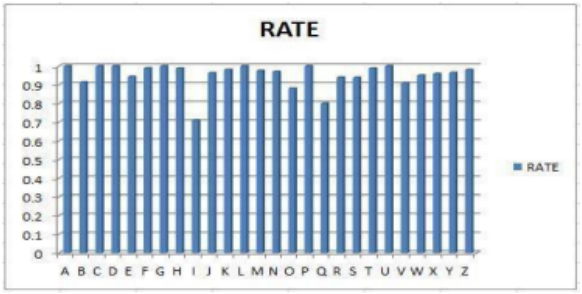

Fig. 11. accuracy of each character

VI. Conclusion

The device functions admirably for a wide range of lighting situations and various sorts of number plates frequently encountered in India. It is unquestionably a superior option than India's current manual systems. There are currently several limitations on criteria like vehicle speed, script on the license plate, and image skew that can be effectively removed by improving the algorithms. Future research will focus on the remaining issues, improving the design and the algorithms to enable number plate recognition that is suitable for commerce in the Australian setting. This paper introduces a real-time automobile license plate identification system and highlights several potential uses for the technology.The steps involved in the recognition process have been identified while discussing the system architecture. The experiment makes it clear that the suggested technique is a viable option for real-time recognition. The experiment described in this paper was conducted under perfect conditions. Research under various weather scenarios is ongoing. For traffic surveying or other application-specific goals covered in the article, the prototyped system will be coupled with the junction surveillance video system.

References

[1] Omran, S. S., & Jarallah, J. A. (2017, March). Iraqi car license plate recognition using OCR. In 2017 annual conference on new trends in information & communications technology applications (NTICT) (pp. 298-303). IEEE.

[2] Laroca, R., Severo, E., Zanlorensi, L. A., Oliveira, L. S., Gonçalves, G. R., Schwartz, W. R., & Menotti, D. (2018, July). A robust real-time automatic license plate recognition based on the YOLO detector.

In 2018 international joint conference on neural networks (ijcnn) (pp. 1-10). IEEE.

[3] Qadri, M. T., & Asif, M. (2009, April). Automatic number plate recognition system for vehicle identification using optical character recognition. In 2009 International Conference on Education Technology and Computer (pp. 335-338). IEEE.

[4] Shidore, M. M., & Narote, S. P. (2011). Number plate recognition for indian vehicles. IJCSNS International Journal of Computer Science and Network Security, 11(2), 143-146.

[5] Coetzee, C., Botha, C., & Weber, D. (1998, July). PC based number plate recognition system. In IEEE International Symposium on Industrial Electronics. Proceedings. ISIE'98 (Cat. No. 98TH8357) (Vol. 2, pp. 605-610). IEEE.

[6] Kulkarni, P., Khatri, A., Banga, P., & Shah, K. (2009, April). Automatic number plate recognition (anpr) system for indian conditions. In 2009 19th International Conference Radioelektronika (pp. 111-114). IEEE.

[7] Kashyap, A., Suresh, B., Patil, A., Sharma, S., & Jaiswal, A. (2018, October). Automatic number plate recognition. In 2018 international conference on advances in computing, communication control and networking (ICACCCN) (pp. 838-843). IEEE.

[8] Rahman, C. A., Badawy, W., & Radmanesh, A. (2003, July). A real time vehicle's license plate recognition system. In Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance, 2003. (pp. 163-166). IEEE.

[9] Hamey, L. G., & Priest, C. (2005, December). Automatic number plate recognition for Australian conditions. In Digital Image Computing: Techniques and Applications (DICTA'05) (pp. 14-14). IEEE.

[10] Mutholib, A., Gunawan, T. S., & Kartiwi, M. (2012, July). Design and implementation of automatic number plate recognition on android platform. In 2012 International Conference on Computer and Communication Engineering (ICCCE) (pp. 540-543). IEEE.

[11] Kulkarni, P., Khatri, A., Banga, P., & Shah, K. (2009, April). Automatic number plate recognition (anpr) system for indian conditions. In 2009 19th International Conference Radioelektronika (pp. 111-114). IEEE.

[12] https://www.kaggle.com/datasets/kedarsai/indian-license-plates-with-labels