An Enhanced Method For Number Plate Recognition

Ankit Jain[1], Brajesh Patel[2]
[1]Dept of CSE
[2]Asstt. Prof, Dept of CSE
[1, 2] Shri Ram Institute of Technology, Jabalpur, Madhya Pradesh, India.

Abstract- *Over the last decade, the number of vehicles is growing day by day. This increases the problems for traffic police such as red light violations, parking problems, wrong lane violations and toll booth violations. Significant research and development of algorithms in intelligent transportation has grabbed more attention in recent years. An automated, fast, accurate and robust vehicle plate recognition system has become need for traffic control and law enforcement of traffic regulations; and the solution is ANPR. This research will be helpful for traffic police to control these traffic violations. Moreover, it will also be helpful for the other character recognition applications. This research will be able to recognize alphanumeric characters in a given image. This file will have complete information such as image path, date and time of the generated file, total number of characters and extracted characters. This technology will be fast, cost effective and highly accurate.*

This paper is dedicated on an improved technique of OCR based license plate recognition using neural network. An algorithm for recognition of license plate is proposed. The whole system can be categorized under three major modules, namely License Plate Localization, Plate Character Segmentation, and Plate Character Recognition.

Keywords- Automatic Number Plate Recognition (ANPR), Optical Character Recognition (OCR), License Plate (LP), Binary Image, Number Plate Localization (NPL), Segmentation, BPNN.

I. INTRODUCTION

With increasing number of vehicles on roads, it is getting difficult to manually enforce laws and traffic rules for smooth traffic flow. Traffic Management systems are installed on traffic signals to check for vehicles breaking the traffic rules. In order to automate these processes and make them more effective, a system is required to easily identify a vehicle. The important question here is how to identify a particular vehicle? The obvious answer to this question is by using the vehicle's number plate as every vehicle has a unique number through which it is easily differentiated from other vehicles. Vehicles in each country have a unique license number, which is written on its license plate. This number distinguishes one vehicle from the other, which is useful especially when both are of same make and model. An automated system can be implemented to identify the license plate of a vehicle and extract the characters and numbers from the region containing a license plate. The license plate number can be used to retrieve more information about the vehicle and its owner, which can be used for further processing. Such an automated system should be small in size, portable and be able to process data at sufficient rate. Various license plate detection algorithms have been developed in past few years. Each of these algorithms has their own advantages and disadvantages. The main objective of the proposed design is to detect a license plate number from an image which is captured from camera. An efficient algorithm is proposed to detect a license plate under various conditions. This algorithm extracts the license plate data from an image and provides it as an input to the stage of License number Plate Recognition.

Figure1.1: ANPR Block diagram.

This paper focuses on design, deployment and evaluation of a system for two industrial applications: detecting traffic violations at urban intersections and vehicle counting on highways. These two applications are among the most important ones in the ITS industry.

An ANPR system consists of three different modules:

a) Monochrome/Color cameras,
b) IR projector, and
c) The processing board.

In addition to compatibility of interfaces, each section must be chosen properly for a specific application.

In this thesis, a detailed exploration on the important parameters of an ANPR module has been done. A

monochrome camera is usually used with an IR projector for plate detection and recognition. Monochrome camera sensors are capable of providing higher details and sensitivity compared to color camera sensors. An IR projector increases the contrast between plate's characters and plates backgrounds. The IR projector is mostly useful at night to brighten the plates. We also need to synchronize the camera exposure time with the IR projector pulses in order to capture images with clearer plates. IR projector which operates invisibly is an important alternative to the flash lights. In many countries, a color image of the violation scene must be stored as evidence. For such purposes, a color camera is located beside the monochrome camera. Color cameras can be deployed alone in places where controlled lighting conditions exist, e.g., in tunnels. Since a single camera is sufficient in these cases the final cost of the system is reduced. Such systems can also be employed in cases where violations are considered only in the daylight. For example, congestion charging systems need to detect the vehicles that enter a specific zone during specific hours.

Basically, the License Plate Recognition (LPR) process is divided into three main parts: Plate Detection, Character Segmentation, and Character Recognition. Each of these parts plays an important role in the final accuracy. Many problems such as size variations, viewing angle, low contrast plates, vehicles high speed and time consuming algorithms have prevented researchers from introducing a single class of algorithms to solve the problem. There have been, however, many algorithms proposed for each part.

For plate detection, several algorithms have been proposed. Some of these algorithms are based on finding image edges, such as horizontal and vertical edges [1]. For example, in [2], plates are localized using the Canny edge detector. Sobel operator is used by some other methods that work based on detecting image edges. These methods have two main advantages: smoothing the image noise because of the included averaging, and generating thick and bright edges because of the involved differentiation on two rows and columns [3]. The advantage of edge detector methods is their low computational complexity and memory requirements. In some other algorithms, plate detection is performed by finding the borders of a plate using the Hough transform [4], which is a memory and time consuming process. Such method fails in detecting plates without clear borders. Fig. 2 compares the results of applying the Hough transform on both clean and dirty plates. Wavelet analysis has also been utilized for detecting plates [5], [6]. In Wavelet based methods, high-frequency coefficients are used to detect plate candidates. Since these coefficients correspond to the edges, these algorithms suffer from the same disadvantages of edge detection algorithms. In [7], and [8], color is incorporated as an important feature in detecting plates. These algorithms fail on gray-scale images or images with low color disparity. Some detection algorithms are based on a combination of Mathematical Morphology and Connected Component analysis. There are algorithms which first enhance the plate contrast and then apply the detection algorithm. Most of these algorithms are successful in identifying clean plates, but fail when it comes to detecting dirty and low contrast plates. This is due to the fact that these algorithms need a medium to high contrast images for plate detection. Moreover, for dirty plates color is not a reliable feature for plate detection. For character segmentation, there are many algorithms based on morphological operations and connected component analysis (CCA). In such methods, it is necessary to apply a proper thresholding method to obtain a binary image of the plate before any further processing. For example, CCA, which has been used in many research works, depends highly on the previously applied thresholding method. Applying such algorithms on plate candidates relies on appropriate setting of the involved parameters. Unfortunately, at the detection step there is no information about the input plate quality and the parameters cannot be tuned appropriately. Therefore, the recognition accuracy of such algorithms decreases when different plate qualities are involved. For character recognition, many different classification tools and techniques have been utilized so far, such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), Bayes classifier, and K-nearest neighbor and so on. Classifiers are applied to the features extracted from the image segmentation.

II. RELATED WORK

For character segmentation, there are many algorithms based on morphological operations, and connected component analysis (CCA) [9]. In such methods, it is necessary to apply a proper thresholding method to obtain a binary image of the plate before any further processing. For example, CCA, which has been used in many research works, depends highly on the previously applied thresholding method. Thresholding methods like Niblack [10], SAUVOLA [11], Wolf and Jolion [12] and OTSU [13] are good candidates for plate binarization. Applying such algorithms on plate candidates relies on appropriate setting of the involved parameters. Unfortunately, at the detection step there is no information about the input plate quality and the parameters cannot be tuned appropriately. Therefore, the recognition accuracy of such algorithms decreases when different plate qualities are involved. For character recognition, many different classification tools and techniques have been utilized so far, such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), Bayes classifier, and K-nearest

neighbor and so on. Classifiers are applied to the features extracted from the image segmentation.

2.1 Feature Extraction Method

For feature extraction, many methods including character skeleton, active areas, HOG and horizontal and vertical projection, and multiclass AdaBoost approach have been proposed. In some methods, the character recognition system is based on key-points localization like SIFT and SURF. In [12], SIFT is utilized in the plate recognition process. In [9], direction contributively density is used for character recognition. In [13], a comprehensive study of character recognition features is presented. We have proposed a new feature set for the character recognition part of our system. SVM classifier, which has been utilized in various character recognition and classification methods, is employed in our system as well.

Our final system is capable of detecting and recognizing multiple license plates in a single frame in high speed applications, such as highway vehicle detection. Based on the resolution of our cameras and the geometrical setup of our system the maximum number of plates in each captured frame is at most seven. There are several challenges in design and deployment of robust highly accurate ANPR systems. These challenges arise in handling high vehicles speeds, different weather conditions (such as rainy, snowy and dusty), different lighting conditions (such as sunrise and sunset effects on license plates), and camera vibrations. Such phenomenon's and problems make many ANPR systems to present poor recognition results. In addition to that, and for the sake of computational complexity, we have tried many different algorithms for different parts of our ANPR system. For example, algorithms with high computational needs were rejected in spite of their high accuracy. To sum up, we must say that many ITS applications depend on real-world Automatic Number Plate Recognition (ANPR) systems. The performance of ANPR systems are degraded by low clarity of the vehicles plates, variations in weather and lighting conditions, and high vehicle speeds. For instance, most of well-known plate detection and recognition algorithms fail on dirty plates [14]. S. Ramalingam proposed an Automatic Number Plate Recognition model that uses OCR to recognize the individual characters of the number plate. This framework maintains a strategic distance from the requirement for doing any asset escalated field trials by the Police Force.

2.2 Hierarchical License Plate Recognition System

A. Vehicle Detection

HLPRS first detect vehicles and then detect plate on the vehicles. This method can avoid misidentifying traffic signs or advertisements as license plates. Traditional methods such as RCNN and Fast-RCNN are based on sliding windows or selective search to find possible targets, and then use CNN or other methods to determine whether the target is an object. Due to the size and complexity of photos taken on roads, the use of a sliding window can be very time-consuming and can easily catch the wrong information.

B. License Plate Localization

After capturing vehicles, we adopt the SVM to detect vehicle's license plates. SVM is a supervised learning method used for classification and regression analysis. We use the SVM OAR (one against rest) architecture with the HOG values of an image as features to train a classifier. In order to identify license plates correctly, we need to train a classifier that can classify license plates and non-license plates. We have to prepare positive and negative samples during the training phase. The positive samples are the license plates and the negative samples are the vehicle's regions without license plates. Due to the reduced complexity of the negative samples, the recognition rate of license plates is increased.

C. Character Segmentation

After capturing license plates, we have to cut out the area outside the characters, filter out noise, and divided it into single characters for later identification. Figure 2.2 shows an example of a detected license plate.

Figure 2.1: An example of a detected vehicle license plate.

The process of character segmentation consists of several steps. First, the captured image is converted to grayscale and then binarized to eliminate noise. Figure 2.3 shows the noise-removed binarized plate.

Figure 2.2: An example of a binarized plate.

Then, we perform a horizontal projection of the license plate image to determine the position of the characters

arranged on the license plate. As shown in Figure 2.4, the upper and lower borders are removed by horizontal projection.

Figure 2.3: Eliminating the upper and lower borders by horizontal projection.

Finally, we perform a vertical projection on the license plate image to determine the position of each character and then divide it into single characters, as shown in Figure 2.5.

Figure 2.4: Separating characters by vertical projection.

D. Character Recognition

In the final stage, we propose a LPRCNN model to identify blurred and skewed characters as shown in Figure 2.6. The proposed LPRCNN model is composed of two convolutional layers, two maxpooling layers, two fully connected layers, and one output layer. The output layer contains 34 neurons to correspond to the 34 plate characters.

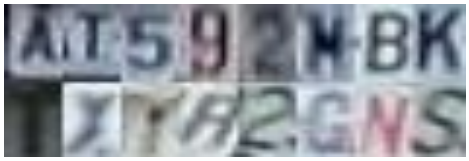

Figure 2.5: Blurred and skewed characters.

III. PROPOSED SYSTEM

The main aim of the proposed work is to detect number plates by recognizing characters using computer vision techniques in OpenCV-Python and KNN. KNN or K-Nearest Neighbor is an algorithm used to recognize characters contained in an image. This process can be, otherwise referred to as Optical Character Recognition (OCR) process.

Proposed work also recognizes multiple license plates from an image. Currently it is beneficial in traffic management at squares having heavy traffic. Countries are constantly working to improve the traffic management system for variety of reasons. The number plate detection is very important and effective in every country keeping in the mind the traffic congestion on the roads. The number plate detection has many applications such as automatic toll collection, tracking of stolen cars, parking lots, etc. The detection of number plate is challenging due to non-uniform nature of lightings in the surroundings. Sometimes the detection of number plate fails due to external environment factors such as snow, rain, extremely cloudy weather.

KNN algorithm is used. KNN is one of the simplest of classification algorithms available for supervised learning. The idea is to search for closest match of the test data in feature space. In pattern recognition, the k-nearest neighbor algorithm is a method to classify characters based on nearest training sets (character set) in the feature space. We started with an initial goal of detecting license plate using contour technique, however this yielded more than expected results. This happened due to state name on the plates, street name boards etc on the image, which was fed to the system. To overcome this challenge we added another module to the system, which was character recognition. To achieve this, we integrated Optical Character Recognition method using kNN algorithm. Using an image file with English alphabets and digits written in few different formats, training algorithms were created and that was loaded and compared to produce much more accurate results.

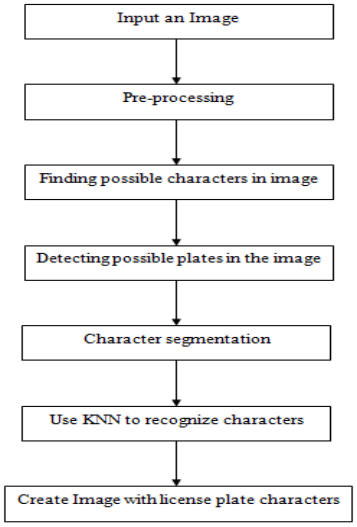

Figure 3.1: Proposed Model.

3.1 Proposed Method

Pre-processing means prepare image to be processed for license plate detection. It contains four steps:

1. Grey Scale Conversion
2. Morphological Operation
3. Image Smoothening
4. Binarization.

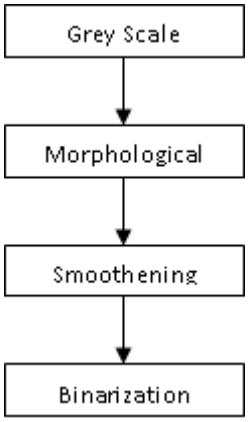

Figure 3.2: Pre-processing steps.

Firstly, proposed system will convert the image into Grayscale image. Morphological operations are used to maximize contrast in the image. There are two operations- top hat and black hat. The top hat filter is used to enhance bright objects of interest in dark background. The black hat operation is used to enhance dark object in white background, so it will introduces more contrast between number plate and the background.

Smoothening is operation which is used to remove noise from image. Basically noise is defined as random variation of brightness. To remove them, smoothening is performed by various filters. Proposed system uses Gaussian blur to remove noise (image smoothing) in gray-scaled image. Image binarization uses thresholding in image to separate an image into a background and foreground.

3.2 License Plate Detection

For License Plate Detection there are three steps:

1. Localize number plate by finding all possible characters in the Image
2. Identify group of matching characters
3. Get list of possible license plates

Now we need to find all the characters present in the image because that will help to localize the number plate, which has a group of characters with same characteristics. For finding the possible characters in the threshold image, we find all the contours. Contours are simply a curve joining all the continuous points (along the boundary), having same color or intensity. The contours are a useful tool for shape analysis and object detection and recognition. Then we are refining the contours list by checking each contour with standard specification for a character like pixel area, aspect ratio, pixel width and height.

This step we rearrange all the possible characters into a list of group of matching characters. If a character that was found does not belong to any group of matches, we omit those, considering that it is not part of a license plate. The matching is performed based on the features like distance between characters, angle between characters, bounding rectangle width, height and area for each character. All these values will be in a limit for license plate characters and this will help to do character grouping. We do this process recursively to find all groups of matching characters.

Finally, we sort the group of matching characters based on size of group, with the assumption that the license plate characters will be the biggest group. Thereafter localize that character group as number plate by drawing a rectangle around the character set.

3.3 KNN for character recognition

In the classification, the K-nearest neighbor algorithm essentially boils down to forming a majority vote between the K most similar instances to a given “unseen” observation. Similarity is defined according to a distance metric between two data points. A popular choice is the Euclidean distance given by:

$$d(x,x') = \sqrt{\left(x_1 - x_1'\right)^2 + \left(x_2 - x_2'\right)^2 + \ldots + \left(x_n - x_n'\right)^2}$$

More formally, given a positive integer K, an unseen observation x and a similarity metricd, KNN classifier performs the following two steps:

• It runs through the whole dataset computing d between x and each training observation. Method will call the K points in the training data that are closest to x the set A. Note that K is usually odd to prevent tie situations.
• It then estimates the conditional probability for each class, that is, the fraction of points in A with that given class label. (Note I(x) is the indicator function which evaluates to 1 when the argument x is true and 0 otherwise) Finally, input x gets assigned to the class with the largest probability.

$$P(y = j \mid X = x) = \frac{1}{K} \sum_{i \in A} I(y^{(i)} = j)$$

Here proposed system will train the system with a set of Upper case alphabets and numerical numbers. After segmenting each character from the license plate, predicting its nearest possible character from the training set.

IV. RESULT

A database comprises of various estimated JPEG and PNG colored images. Add up to 50 pictures are exploited to test the calculation. The pictures are brought with various foundation and also light conditions. Tests demonstrate that the algorithm has great success rate on number plate extraction.

Table 4.1 illuminates number plate recognition success rate between existing and proposed methods:

Method	Total no of images	Success rate (in percentage)
Existing	50	81
Proposed	50	89

Chart that describes above table is shown below:

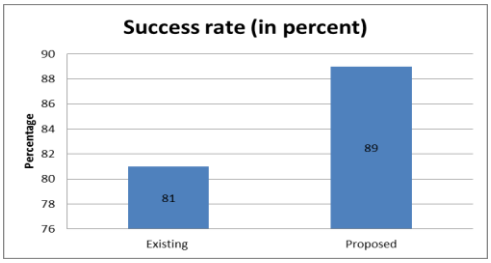

Figure 4.1: Comparison of success rate.

A database also comprises of various estimated JPEG and PNG colored images that contain multiple license plates, so success rate for multiple license plate detection is also compared. Add up to 30 pictures are exploited to test the calculation. The pictures are brought with various foundations. Tests found that proposed system is more successful in finding multiple license plates from an image.

Table 4.2 illuminates number plate recognition success rate for multiple license between existing and proposed methods:

Method	Total no of images	Success rate (in percentage)
Existing	30	62
Proposed	30	87

Chart that describes above table is shown below:

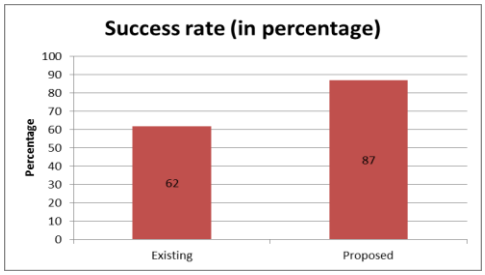

Figure 4.3: Comparison of success rate for multiple license plate detection in an image.

\psychological problem. Untreated depression increases the chance of dangerous behaviors. The significant challenge of detecting depression is the recognition that depressive symptoms may differ from patients' behavior and personality [1]. For clinic depression, doctors may evaluate the patient via the depression test taken by patients. Apparently, these clinical records are restricted due to many factors, such as age, sex; moreover, they are private and expensive. To overcome such limitations of clinical data, it would be beneficial to use text mining tools to extract and analyze depression symptoms from social media, such as Twitter. Social media generates countless data every day because of millions of active users share and communicate in entire community, it changes human interaction [2,3]. Other than the traditional data, such as literatures, social media data is richer and more accessible [4]. However, investigating this new fast-growth of data requires advanced development tool to discover useful insights. These advanced technologies include Natural Language Processing (NLP) [5], data mining, machine learning, social media analysis and so on. In our research, the goal is to extract and summarize the uncommon but potentially helpful factors that depressive symptom performed from the social media data. Finally, the extracted depression symptoms w

IV. CONCLUSION

License plate recognition plays an important role in intelligent transportation system and it has huge number of practical applications such as automatic toll collections, parking fee payment, detection of vehicle crossing speed limits and thereby reducing road accidents etc. Literatures have been extensively reviewed. Neural network has capability of learning, i.e. how to do tasks based on data given for training. Back propagation is supervised form of learning. In License plate recognition system, features extracted from license plate are used as input to neural networks and these are allowed to propagate forward to generate output. Learning of neural networks is done by loading targets and features extracted from license plate characters. Testing of neural network is done and parameters used for performance evaluation i.e. recognition rate and training time is computed. System simulated found that our proposed method performs better over existing one.

Currently the Car Registration License Plate Detection and Recognition System support the basic functionality to detect license plates, recognize the characters and save the output in a file. This application can be extended in many ways. Some of them are listed below:

- Support to recognize the state: It will be nice if the application could recognize the state to which the vehicle belongs by seeing its license plate.
- There is lots of variety in the style of license plates. Their style varies from country to country. For example, at some places, dark characters are written on light plate; at other places, white characters on black background. Some motor departments use one-row plate and others use two-row plate. There is lots of variety in character sizes and fonts as well. So the project should be trained enough to work with various styles.
- The method should be able to automatically decide the threshold values for character extraction according to the quality of the picture.

Vehicle registration number recognition is beneficial in many areas such as:

- Traffic control in restricted area.
- Car parking management.
- Automatic toll barriers.
- Red light violation
- Access control
- Traffic monitoring
- Border crossing
- Security in military areas

There are some other applications which use similar concepts such as text recognition system from pdf to image files, face detection system which are largely used in camera applications, face recognition for security purposes, iris recognition, finger print recognition, pattern recognition etc. All these applications are based on neural network, data mining and image processing.

REFERENCES

[1] B. Hongliang and L. Changping, "A hybrid license plate extraction method based on edge statistics and morphology," in Proc. IEEE 17th ICPR, 2004, vol. 2, pp. 831–834.

[2] A. Mousa, "Canny edge-detection based vehicle plate recognition," Int. J. Signal Process., Image Process. Pattern Recognit. vol. 5, no. 3, pp. 1–8, 2012.

[3] W. Gao, X. Zhang, L. Yang, and H. Liu, "An improved Sobel edge detection," in Proc. IEEE 3rd ICCSIT, 2010, vol. 5, pp. 67–71.

[4] T. D. Duan, D. A. Duc, and T. L. H. Du, "Combining Hough transform and contour algorithm for detecting vehicles' license-plates," in Proc. IEEE Int. Symp. Intell. Multimedia, Video Speech Process., 2004, pp. 747–750.

[5] P. Kanani, A. Gupta, D. Yadav, R. Bodade, and R. B. Pachori, "Vehicle license plate localization using wavelets," in Proc. IEEE ICT, 2013,pp. 1160–1164.

[6] S. Du, M. Ibrahim, M. Shehata and W. Badawy, "Automatic License Plate Recognition (ALPR): A State-of-the-Art Review," IEEE Trans. Circuits and Systems for Video Technology, vol. 23, no. 2, pp. 311- 325, Feb. 2013.

[7] L. Connie, C. Kim On and A. Patricia, "A Review of Automatic License Plate Recognition System in Mobile-based Platform," Journal of Telecommunication, Electronic and Computer Engineering (JTEC), Vol 10, No 3-2, pp. 77-82, 2018.

[8] N. N. Kyaw, G. R. Sinha and K. L. Mon, "License Plate Recognition of Myanmar Vehicle Number Plates A Critical Review," IEEE 7th Global Conference on Consumer Electronics (GCCE), Nara, pp. 771- 774, 2018.

[9] Rupali Kate, Dr.Chitode J.S "Number Plate recognition using Segmentation" International journal of engineering & technology (IJERT) ISSN: 22780-0181. Vol.1 ISSUE 9-Nov-2012.

[10] P. Anishiya and Prof. S. M. Joans, "Number Plate Recognition for Indian Cars using Morphological Dilation and Erosion with the Aid of Ocrs" International Conference on Information and Network Technology, Singapore, 2011.

[11] M. R. Bai, V.V. Krishna, and J. Sreedevi, "A new morphological approach for noise removal cum edge detection," IJCSI International Journal of Computer Science Issues, Vol. 7, Issue6, November 2010.

[12] Hontani, H., and Koga, T., "Character extraction method without prior knowledge on size and information", Proceedings of the IEEE International Vehicle Electronics Conference (IVEC'01), pp. 67-72, 2001.

[13] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. You only look once: Unified, real-time object detection. ArXiv preprint arXiv: 1506.02640, 2015.

[14] Jayash Jagtap, Sushil Holambe," Multi-Style License Plate Recognition using Artificial Neural Network for Indian Vehicles" ICICET, IEEE -2018.