International Journal For Technological Research In Engineering
Volume 5, Issue 7, March-2018 *ISSN (Online): 2347 - 4718*

AUTOMATIC NUMBER PLATE RECOGNITION USING CONNECTED COMPONENT ANALYSIS ALGORITHM

S. Stephy Golda Mercy[1], Dr. I. Muthulakshmi[2]
[1]PG Scholar, [2]Professor
Department of Computer Science and Engineering VV College of Engineering, Tamil Nadu, India.

Abstract: Automatic Number Plate Recognition that can be used as a basis for many real-words ITS (Intelligent transportation system) application. The system is designed to real with unclear vehicles plates, variations in weather and lighting conditions, different traffic situations, and high- speed vehicles. ANPR System consists of three parts; they are Plate Segmentation, Plate Detection and Plate Recognition. The proposed work is used to extract numbers from the number plate. This technology is used in various security purposes for finding stolen cars, traffic management system, usage of cars in terrorist attacks and illegal activities. Character segmentation is the process of extracting the characters and numbers from the license plate. Noises in the image are removed using filtering techniques. Connected-component Analysis (CCA) technique is used for the character recognition .character is coordinated with layout utilizing format coordinating calculation lastly the character is separated. The proposed algorithm is compared with the existing methods and found that it is performed better than the existing methods.
Index Terms: Connected-component analysis(CCA), Support Vector Machine (SVM).

I. INTRODUCTION

Automatic Number Plate Recognition (ANPR) is an image processing technology which uses number (license) plate to identify the vehicle. The objective is to design an efficient automatic authorized vehicle identification system by using the vehicle number plate. The system is implemented on the entrance for security control of a highly restricted area like military zones or area around top government offices e.g. Parliament, Supreme Court etc. The developed system first detects the vehicle and then captures the vehicle image. Vehicle number plate region is extracted using the image segmentation in an image. Connected-component Analysis technique is used for the character recognition. The resulting data is then used to compare with the records on a database so as to come up with the specific information like the vehicle's owner, place of registration, address, etc. The system is implemented and simulated in Mat lab, and it performance is tested on real image. It is observed from the experiment that the developed system successfully detects and recognize the vehicle number plate on real images. ANPR is divided in two main steps: plate detection and plate recognition. Plate detection has the purpose of detecting the location of the plate in the whole camera frame. When a plate is detected in an image, the plate segment is passed to the second step. Plate recognition is defined as the act of identifying someone or something because of previous knowledge, or to formally acknowledge someone. Formal acknowledgment of the political existence of a government, Real time number plate recognition plays an important role in maintaining law enforcement and maintaining traffic rules [5]. Many number plates have different styles and varying state by state. The number plates have one row or two rows of numbers and have six to more than ten letters.

Image Processing is an area that uses several techniques and algorithms in order to interpret and understand the information contained in a digital image [2].Most image processing algorithms consist of a few typical steps viz.

- image pre-processing,
- segmentation,
- feature extraction,
- feature selection
- Classification.

First, number plate image is taken. Image processing step includes removal of noise using filters and to enhance the image.

Image segmentation is the process of partitioning an image into multiple segment or set of pixels used to locate object and boundaries [13]. Each of the pixels in a region is similar with respect to some characteristics such as color, intensity or texture.

Image classification of number plate module is done in order to eliminate operator dependency and to improve the accuracy. In digital image classification the conventional statistical approaches for image classification use only the gray values.

Image processing is a method to perform some operations on an image, in order to get an enhanced image. Nowadays, image processing is among rapidly growing technologies. Image Processing is a technique to enhance
raw images received from cameras placed pictures taken in normal day-to-day life for various applications.

Image processing usually refers to digital image processing, but optical and analog image processing also are possible. This article is about general techniques that apply to all of them. The acquisition of images (producing the input image in the first place) is referred to as imaging.

Intelligent transport systems[20] vary in technologies applied, from basic management systems such as car navigation; traffic signal control systems, container management systems, variable message signs, automatic number plate recognition or speed cameras to monitor applications, such as security and to more advanced applications that integrate live data and feedback from a number of other sources, such as parking guidance and information systems, weather information, bridge de-icing

systems, and the like. Additionally, predictive techniques are being developed to allow advanced modeling and comparison with historical baseline data [6]. Some of these technologies are described in the following sections. An intelligent transportation system (ITS) is an advanced application. Aims to provide innovative services, ITS may improve the efficiency of transport in a number of situations, i.e. road transport, traffic management, mobility, etc.

II. LITERATURE SURVEY

A literature survey is a text of a scholarly paper, which includes the current knowledge including substantive findings, as well as theoretical and methodological contributions to a particular topic. Literature reviews are secondary sources, and do not report new or original experimental work. 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, sno wy 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. This chapter discuss some of algorithm and approaches proposed by various author for efficiently character recognition and also segmentation of ANPR system. Mayunga et al. [8] proposed a radial casting algorithm and it sufficiently extract the outlines of the buildings. But it is quite expensive, manually operated and requires well-trained personnel. Peng et al. [2] concluded that high-resolution imagery is a valuable tool for mapping urban areas and extracting land cover information and his model can work robustly in dealing with images taken from urban areas with buildings and roads of high density and complex structure. Muller and Zaum [10] proposed a seeded-region growing algorithm for segmentation. It works well for small images and the runtime of the approach depends on the image content. Chaudhuri et al. [4] used multispectral IKONOS imagery to determine the approximate location and shape for buildings and roads. Morphology-based approach is used to extract structural information from satellite images. Given the recent availability of the commercial high-resolution satellite imagery, only a few methods for building and road detection/extraction from less than or equal to 1m/pixel resolution imagery have been developed. The effect of resolution on the building extraction was reported in [3]. The common challenges in generating building hypotheses from less than or equal to 1m/pixel imagery include: 1) Low-signal-to-noise ratio and 2) Weak object signal. A rapid automatic building extraction approach from very high resolution optical satellite imagery was presented in [16]. The proposed method conducts building extraction based on distinctive image primitives such as lines and line intersections. Wang et al. [1] proposed an edge-aided re-classification algorithm which effectively eliminate the parking lots and driveways and extract complex road junctions. But it fails to extract the road networks, when the road edges are not clear, heavily shadowed and occluded. However, the main drawback of these approaches is that they are not fully automated. In this paper, an automated building extraction strategy for high-resolution satellite imagery is proposed that utilizes structural, contextual, and spectral information and it does not require any manual operations done by human operator.

III. BUILDING AND ROAD DETECTION ALGORITHM

The basic methodology which gives the systematic and theoretical analysis of the methods for detecting the buildings and roads automatically from the high-resolution satellite images are given below.

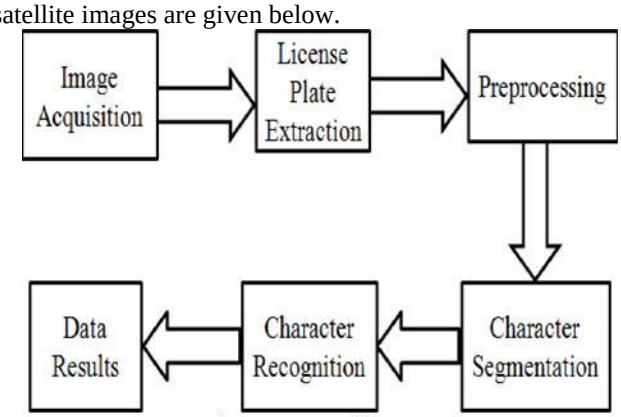

Fig. 1. System Architecture

The system is proposed with new algorithm for number plate detection. First the input color image is converted to gray scale image for easy to handle and simple way to find the location in the number plate. This system uses blurred regions and different font style and sizes using the character reorganization. This work has reduced the error, and time. An everyday increase in the number of cars on roads and highways facing numerous problems for example identification of stolen cars, smuggling of cars, invalid license plate, usage of cars in terrorist attacks and illegal activities.

ANPR is divided in three main part:

PLATE SEGMENTATION

Segmentation is one of the most important processes in the automatic number plate recognition. In the first step (segmentation), we apply different filters, morphological operations, contour algorithms, and validations to retrieve those parts of the image that could have a plate. Segmentation subdivides an image into its constituent regions or objects. That is, it partitions an image into distinct regions that are meant to correlate strongly with objects or features of interest in the image. Segmentation can also be regarded as a process of grouping together pixels that have similar attributes.

Image segmentation is the process of dividing an image into multiple parts [14]. This is typically used to identify objects or other relevant information in digital images.

This is the process were characters are being segmented or divided into more manageable task that could easily be worked on better that the full image. There are many

different ways to perform image segmentation, including steps are

Image Acquisition
The first step of ANPR is image acquisition which means to acquire the input image of vehicle. Image is acquired by digital camera. Images are taken in different background and illumination conditions and at various distances from the camera. Due to poor illumination conditions the acquired image can be of low contrast. Weather conditions (fog, snow, rain) are responsible for introducing "noise" during camera capturing. Different types of images can be acquired during camera capturing that is Light Images, Dark Images, Low Contrast Images, Blurred Images and Noisy Images.Image acquisition is the process of obtaining an image from the camera. This is the first step of any vision based systems. In our current research we acquire the images using a digital camera placed by the road side facing towards the incoming vehicles .Here our aim is to get the frontal image of vehicles which contains license plate.

Image Pre-Processing
Number plate is pre-processed to remove the noise and then the result is passed to the segmentation part to segment the individually characters from the extracted number plate. The preprocessing algorithm helps in improving the quality of the image or the plate image being inputted to the system.

Image Pre-processing involves:
(1). HSV image
HSV is the most common cylindrical-coordinate representations of points in an RGB color model. This representation rearranges the geometry of RGB in an attempt to be more intuitive and perceptually relevant than the Cartesian (cube) representation [5]. HSV is used today in all digital television encoding systems, color pickers, in image editing software, and less commonly in image analysis and computer vision.

(2). Secondary Image
YCbCr, Y′CbCr, or Y Pb/CbPr/Cr, also written as YCBCR or Y'CBCR, is a family of color spaces used as a part of the color image pipeline in video and digital photography systems. Y′ is the luma component and CB and CR are the blue-difference and red-difference Chroma components. Y′ (with prime) is distinguished from Y, which is luminance, meaning that light intensity is nonlinearly encoded based on gamma corrected RGB primaries.

(3). Gray Image
In photography and computing, a grayscale or grayscale digital image is an image in which the value of each pixel is a single sample, that is, it carries only intensity information. Images of this sort, also known as black-and-white, are composed exclusively of shades of gray, varying from black at the weakest intensity to white at the strongest [10].This the process were the 24bit image color is being reduced or converted to 8bit color that could be used and accepted by the number plate system.

RGB to Gray-scale Conversion
In RGB format, each Pixel has three colour components: Red, Green, and Blue. In pre-processing step, the colour image is given as an input and it is converted into grayscale image. also known as black-and-white, are composed exclusively of shades of gray, varying from black at the weakest intensity to white at the strongest .

Canny Edge Detection
The Canny edge detector is an edge detection operator that uses a multi-stage algorithm to detect a wide range of edges in images. Canny edge detection is a technique to extract useful structural information from different vision objects and dramatically reduce the amount of data to be processed. It has been widely applied in various computer vision systems. Canny has found that the requirements for the application of edge detection on diverse vision systems are relatively similar.

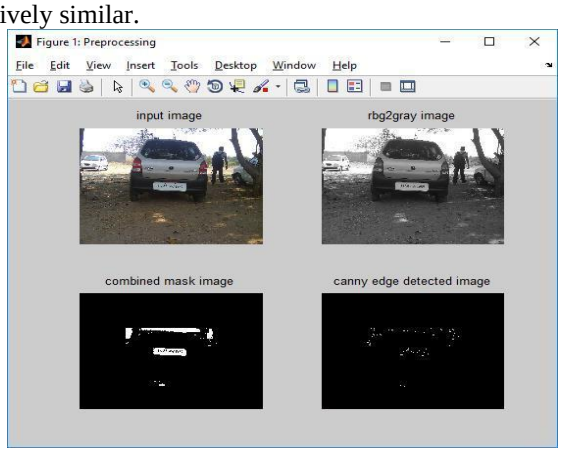

Fig. 2.Pre-processing image

The input image is converted into gray image and edge detection operation is performed.

PLATE DETECTION
Plate detection is an image processing technique for finding the boundaries of objects within images. It works by detecting discontinuities in brightness. Plate detection is used for image segmentation and data extraction in areas such as image processing, computer vision, and machine vision.
The purpose of plate detection is significantly reducing the amount of data in an image and preserves the structural properties for further image processing. Plate detection performs the locating sharp discontinuities in an image. The edge is a boundary between two regions with relatively distinct gray level properties. In edge detection, many operators are defined such as sobel, log, canny, prewitt [10]. The Canny operator was designed to be an optimal edge detector.
It takes as input a gray scale image, and produces as output an image showing the positions of tracked intensity discontinuities. The goal of this section is to elaborate on the methods of finding the vehicles plate's location in captured images.

(1). CCA Algorithm
Connected-component Analysis is an algorithmic application

of graph theory, where subsets of connected components are uniquely labeled based on a given heuristic. CCA is used in computer vision to detect connected regions in binary digital images, although color images and data with higher dimensionality can also be processed.
When integrated into an image recognition system or human-computer interaction interface [19].CCA can operate on a variety of information. Connected Component Analysis, CCA, is one of the most widely selected algorithms for the initial step of character recognition in various segmentation methods. The connected components labeling operator scans the image by moving along a row until it comes to a point p (where p denotes the pixel to be labeled at any stage in the scanning process) for whichV={1}.
When this is true, it examines the four neighbors of p which have already been encountered in the scan Based on this information, the labeling of p occurs as follows:
If all four neighbors are 0, assign a new label to p,
Else if only one neighbor has V={1}, assign its label to p,
Else if more than one of the neighbors have V={1},assign one of the labels to p and make a note of the equivalences.
After completing the scan, the equivalent label pairs are sorted into equivalence classes and a unique label is assigned to each class. As a final step, a second scan is made through the image, during which each label is replaced by the label assigned to its equivalence classes. For display, the labels might be different gray levels or colors. Connected components labeling scans an image and groups its pixels into components based on pixel connectivity.

PLATE RECOGNITION
Plate 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. The features exploited in this process are extracted from the characters in their original size. In other words resizing is avoided, since it results in losing information [22]. For each character candidate, four scale invariant features are calculated.

Support vector machine (SVM)
In machine learning, support vector machines (SVMs, also support vector networks) are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis [23]. Given a set of training examples, each marked as belonging to one or the other of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier (although methods such as Platt scaling exist to use SVM in a probabilistic classification setting). An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. In addition to performing linear classification, SVMs can efficiently perform a non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high-dimensional feature spaces.

B. Character Recognition
The features exploited in this process are extracted from the characters in the in original size. In other words resizing is avoided, since it results in losing information. The characters are normalized. Normalization is to refine the characters into a block containing no extra white spaces (pixels) in all the four sides of the characters [14]. Then each character is fit to equal size for each character candidate, four scale invariant features are calculated. The details of each feature can be described as follows:

(1). Distance to Walls (DTW)
To calculate DTW, the character area is divided into 12 regions (4 rows×3 columns) and in each region the vertical and horizontal distance of active pixels are calculated with respect to the borders of the character area. For each region the shortest and the longest distances are stored shows the calculation scheme for DTW feature. DTW is a 48-dimensionalvector.

(2). Cross-Time Feature (CTF)
This feature includes the six widest white to black and black to white transitions in traversing character horizontally and vertically. For each row and column of a character candidate, the six widest transitions are calculated.

These crossings are sorted and the first three rows and columns with highest and lowest transitions are stored, along with their locations. Storing the locations improves the classification results for random transitions in non-character candidates. The details of calculating this feature is depicted. The gray lines show the black to white crossing widths and white lines depict the white to black crossing widths. For each component, CTF is a 24-dimensionalvector.

(3). Active Region Ratio (ARR)
For calculating this feature the candidate rows and columns are divided into four and three parts respectively. The number of active pixels in each block divided by the total number of active pixels is chosen as a feature for that block.

(4). Height to Width Ratio (HWR)
Alphanumeric characters in any language can be classified under three types, when it comes to height to width ratio: Characters with greater heights, equal heights and widths, and greater widths. These four features are able to recognize vehicles plates with character size of at least 7 × 3 pixels.

Number Plate Extraction
The character segmentation algorithm is used to segment the character. Due to this character segmentation process noise is added and that noise is removed using the filter [12]. Noise reduction is the process of removing noise from a signal. All recording devices, both analog and digital, have traits that make them susceptible to noise. Noise can be random or white noise with no coherence, or coherent noise introduced by the device's mechanism or processing algorithms.

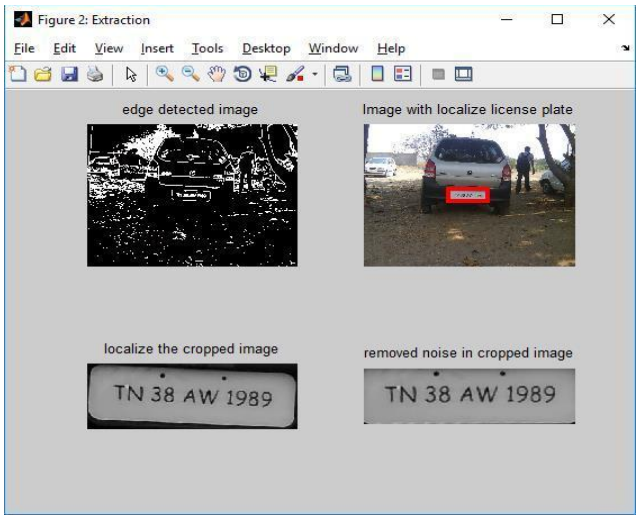

Fig. 3.Number Plate Extraction

The Number Plate Extract from the orginal input image, and remove noise in cropped image. The next step is plate segmentation, it has segment the number plate based on some filters.

Character Segmentation

Connected Component Analysis, CCA, is one of the most widely selected algorithms for the initial step of character recognition in various segmentation methods after detecting the exact location of the plates in a captured image; a binarization process is performed on the detected plates. Several thresholding methods have been proposed for the binarization process at this step the results.

Gabor filters are directly related to Gabor wavelets, since they can be designed for a number of dilations and rotations. However, in general, expansion is not applied for Gabor wavelets, since this requires computation of bi-orthogonal wavelets, which may be very time-consuming. Therefore, usually, a filter bank consisting of Gabor filters with various scales and rotations is created. The filters are convolved with the signal, resulting in a so-called Gabor space. This process is closely related to processes in the primary visual cortex.An input grayscale image (120x120 or 160x160) is densely filtered by a battery of gabor filters at each scale and orientation. Therefore, at each pixel of the input image, filters of each size and orientation are centered.

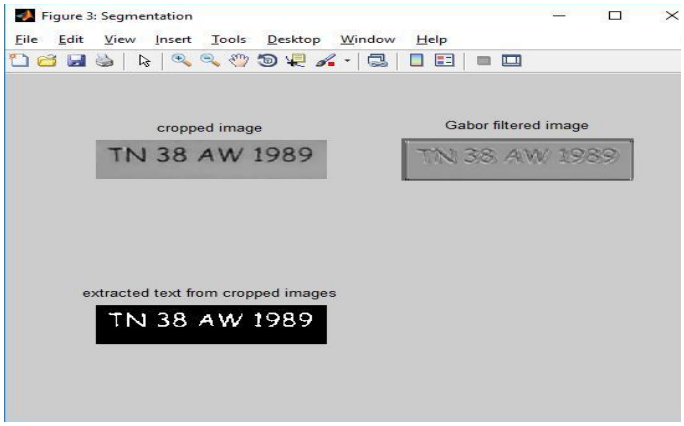

Fig. 4. plate segmentation Image

Segmentation is the process of dividing an image into multiple parts. This is typically used to identify objects or other relevant information in digital images.

Template Matching

Template Matching is one of the most common classification methods. In Template Matching, the features that the classification is based on are the individual pixels [12]. An image is compared with predefined images, which are referred to as templates. The patch image which will be compared to the template image, it compares the number and alphabets with training data set.

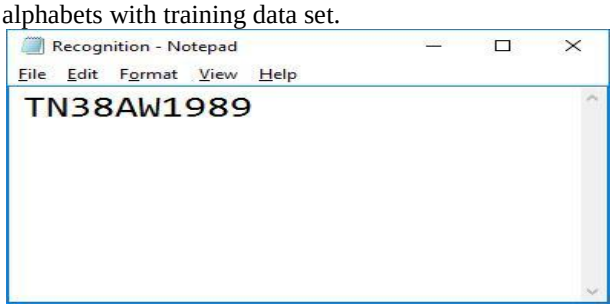

Fig. 5. Final Output

Final output shown in the figure 5. Output will be displayed on notepad.

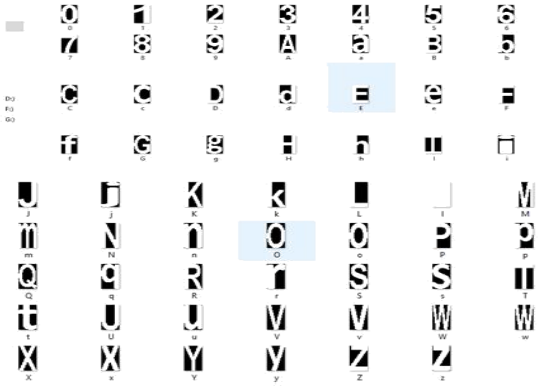

Fig. 6.Training Letters and Numbers

Figure 6 shows the training data set, which is used for template matching.

IV. EXPERIMENT RESULTS

Recognition system works well for low contrast, noisy and blurred as well as dark and light/bright input images. Total 90 vehicle's images are tested. Images are taken in different illumination conditions, at different distances relative to camera and are of different colors and different sizes images. The proposed approach is tested on various real time images of different image categories as using various performance metrics.

V. CONCLUSION

The proposed work is the automatic vehicle identification system using vehicle license plate is presented. The system is implemented in Matlab and it performance is tested on real images. A number plate recognition system is one kind of an Intelligent Transport System. In this document, template matching algorithm has been used to extract the vehicle number plate. The automatic vehicle identification system plays an important role in detecting security threat. Here,

character segmentation for separating individual characters. Finally, match with template using template match algorithm and extract the number plate in notepad. The main advantage of our system is its high detection and recognition accuracies on dirty plates. To achieve reliable evaluations. The system use image processing techniques for identifying the vehicle from the database stored in the computer. The system works satisfactorily for wide variation of conditions and different types of number plates.

ACKNOWLEDGEMENT
This work was supported in part by Anna University recognized research center lab at VV College of Engineering, Tisaiyanvilai, Tamil Nadu, India.

REFERENCES

[1] 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.

[2] S.Du,M.Ibrahim, M.Shehata, and W.Badawy, "Automatic License Plate Recognition (ALPR): A state-of-the-art review," IEEE Trans. Circuits Syst. Video Technol., vol. 23, no. 2, pp. 311–325, Feb. 2013.

[3] C. N. E. Anagnostopoulos, I. E. 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.

[4] K. Deb, V. V. Gubarev, and K.-H. Jo, "Vehicle license plate detection algorithm based on color space and geometrical properties," in Emerging Intelligent Computing Technology and Applications. Berlin, Germany: Springer, pp. 555–564, 2009.

[5] K. Deb, H. Lim, S.-J. Kang, and K.-H. Jo, "An efficient method of vehicle license plate detection based on HSI color model and histogram," in Next-Generation Applied Intelligence. Berlin, Germany: Springer, pp. 66–75, 2009.

[6] 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, July 2013.

[7] Mohammed S. Khalil, FajriKurniawan "License Plate Detection Method for Real-time Video of Low-Cost Webcam based on Hybrid SVM-Heuristic Approach" in Proc. IEEE 11th Int. Conf,pp.321-326.2014.

[8] 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.

[9] Riazul Islam, Kazi Fatima Sharif and SatyenBiswas, "Automatic Vehicle Number Plate Recognition Using Structured Elements"in Proc. IEEE 16th Int. Conf. , vol. 2, pp. 1037–1040,2016.

[10] Choudhury A. Rahman, WaelBadawy"A Real Time Vehicle's License Plate Recognition System"in Proc. IEEE Conference,2003.

[11] S. Kranthi, K. Pranathi, and A. Srisaila, "Automatic number plate recognition," Int. J. Adv. Tech., vol. 2, no. 3, pp. 408–422, 2011.

[12] W. Niblack, An Introduction to Image Processing. Englewood Cliffs, NJ, USA: Prentice-Hall, 1986, pp. 115–116.

[13] R. C. Gonzalez and R. E. Woods, Digital Image Processing. Reading, MA: Addison-Wesley, 1993..

[14] Y.-Q. Yang, J. B. R.-L. Tian, and N. Liu, "A vehicle license plate recognition system based on fixed color collocation," in Proc. Int. Conf. Mach. Learning Cybern., vol. 9. 2005, pp. 5394–5397.

[15] K. Kanayama, Y. Fujikawa, K. Fujimoto, and M. Horino, "Development of vehicle-license number recognition system using real-time image processing and its application to travel-time measurement," in Proc. IEEE Veh. Tech. Conf., May 1991, pp. 798– 804.

[16] J.C. Poon, M. Ghadiali, G. Mao, and L. M.Sheung, "A robust vision system for vehicle license plate recognition using gray-scale morphology," in Proc. IEEE ISIE, 1995, vol. 1, pp. 394–399

[17] N. Otsu, "A threshold selection method from gray-level histograms," Automatica, vol. 11, no. 285–296, pp. 23–27, 1975.

[18] 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.

[19] Y.Wen, Y.Lu, J.Yan, Z.Zhou, K.M.vonDeneen, and P.Shi,"Analgorithm for license plate recognition applied to intelligent transportation system," IEEE Trans. Intell. Transp. Syst., vol. 12, no. 3, pp. 830–845, Sep. 2011.

[20] http://code. google.com/p/bgslibrary/

[21] H. Caner, H. S. Gecim, and A. Z. Alkar, "The characteristic feature based on four types of structural information and their effectiveness for character recognition," IEICE Trans., vol. J68-D, no. 4, pp. 789– 796, 1985.

[22] C. Wolf, J.-M. Jolion, and F. Chassaing, "Text localization, enhancement and binarization in multimedia documents," in Proc. IEEE 16th Int. Conf. Pattern Recognit., 2002, vol. 2, pp. 1037–1040

[23] F. Camastra, "A SVM-based cursive character recognizer," Pattern Recognit., vol. 40, no. 12, pp. 3721–3727, Dec. 2007.

[24] D.Zheng,Y.Zhao,andJ.Wang, "An efficient method of license plate location," Pattern Recognition. Lett., vol. 26, no. 15, pp. 2431–2438, Nov. 2005.

[25] Llorens, Marzal, Palazon, and Vilar, Car License Plates Extraction and Recognition Based on Connected Components Analysis and HMM Decoding, vol. 3522. Berlin, Germany: Springer, 2005, pp. 571–578.

[26] J.-W. Hsieh, S.-H. Yu, and Y.-S. Chen,

"Morphology-based license plate detection from complex scenes," in Proc. IEEE 16th Int. Conf. Pattern Recognit., 2002, vol. 3, pp. 176–179.

