License Plate Recognition System Using KNN Method

Alya M. Issa
College of Electronic Technology
aliameftahgaroom@gmail.com

Ali M. Obaid
Libyan Petroleum Institute
alimeftahobaid@gmail.com

Yonis A.Esoud
College of Technical Sciences
yonisabdulqadergodged@gmail.com

Abstract: nowadays tracking vehicles is a very cumbersome task due to their voluminous numbers. As a consequence an automatic identification system becomes a matter of practical necessity. To achieve such an objective, we propose an automatic car License Number Plate Recognition System (LNPRS). The system consists of three parts; the first has the role of preprocessing the image and prepare it for later stages. The second logically segments the license plate sub-image from the original front image of the vehicle. The third part deals with the sub-image that encompasses the license number, for recognition through a set of recognition algorithms based on Knn-Nearest Neighbor method. The final outcome would be the Vehicle Identification Number, VIN.Accordingly, the status of every vehicle would be easily tracked. Experiments on our system produced an overall 91% recognition rate.

Keywords: License Number Plate Recognition System LNPRS, KNN K-Nearest Neighbor, VIN Vehicle Identification Number.

I. Introduction

Typical vehicles tracking system would consist of a camera, a computer contain recognition software. The system acquires the image of the tracked vehicle –from the front or the rare, analyses the image, and obtains the license-plate number or Vehicle Identification Number, VIN.

The quality of the acquired images is a major factor in the success of the LNPRS. LNPRS as a real-life application has to quickly and successfully process license plates under different environmental conditions such as indoors, outdoors, day/night time.

Automatic License Plate Recognition is a computer vision technology that efficiently identifies cars number plates from images without the need for human mediation.

In recent years, it has become more and more important, this is due to the following main factors:-

The growing number of vehicles on the roads, traffic law enforcement, and crimes resolution, as it helps to identify the vehicles of the offenders, difficulty managing busy parking cars area in terms of entry, exit and security.

II. Related Works on LNPRS

In general, License-Number Plate Recognition System, LNPRS, in its essence is more or less a numeral recognition system with added capability for recognizing isolated symbols. Thus, research on the subject is relatively abundant. We shall trace the most resent activities. In [1] researchers provide a typical License-Number Plate Recognition System. The system is composed of the four well-known stages; namely, Image Acquisition, Plate Localization, Character Isolation, and Character Recognition. Each step might include further processing possibility due to the diverse nature of the acquired plate image. They mentioned that LNPRS emerged in the 1980's and currently there are multiple commercial license plate recognition systems available in the markets. In [2] the authors present a state of the art of the research specifically on LNPRS. They mentioned, among others, that the plate rectangle is detected by different techniques including geometrical attributes, Sobel filters to detect edges of color transition, block-based methods, and Hough transform. Some techniques opted for localizing characters directly in the image of the car using local features [3]. Fuzzy rules are also used to extract texture features from the entire car image to recognize characters [4]. In the recognition phase, they mentioned several recognition methods. These methods include template matching after resizing the character shape to a standard size, and features extraction through feature vector. Classification can be performed by different classifiers including those of neural networks and multistage classifiers [5]. Experimental results of practical LNPRS showed 90% overall success in a data set of 16800 images [2].

III. The Method of KNN Nearest Neighbor

KNN is a non-parametric classification algorithm proposed by Cover and Hart in 1968 [6]. The most used measure in finding the K nearest neighbors is the Euclidean distance [7]. Practically, K nearest neighbors are found by calculating the Euclidean distance between the stored templates and the would-be classified example. Euclidean distance is both easy to program and very efficient. The classification of any given example to a certain class is determined by the highest number of votes of the labels of the *k*-near neighbors, where *k* is always a small integer [8].

For instance, if example 1 *x* has k nearest examples in the template space and a majority of them have the same label 1 *y*, then example 1 *x* belongs to 1 *y*.

Euclidean distance is calculated as follows:

Having two vectors $x_i$andx_j, where

$$x_i = (x_i^1, x_i^2, \dots, x_i^n),$$

$$x_j = (x_j^1, x_j^2, \dots, x_j^n),$$

The distance between $x_i$and$x_j$is:

$$D(x_i, x_j) = \sqrt{\sum_{k=1}^{n} (x_i^k - x_j^k)^2}$$

The Realized System

Our realized system is composed of the following stages**:**

As shown in ***Fig 1***, below.

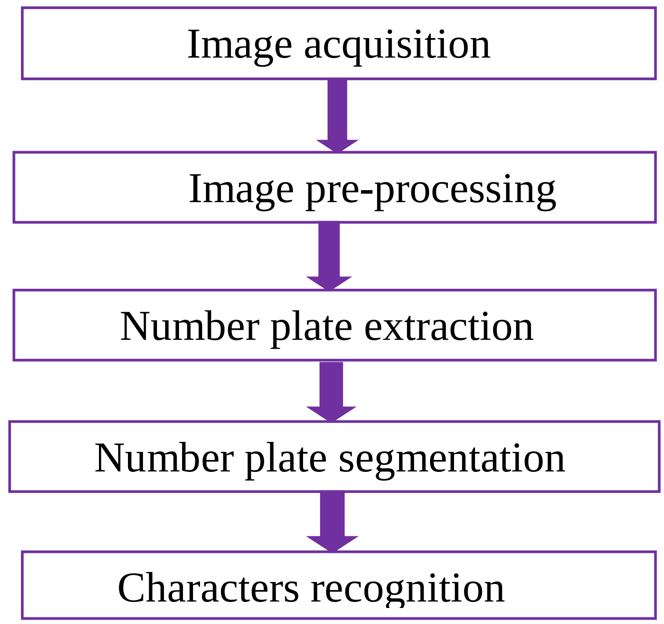

Figure 1, Stages of the LNPRS

Preprocessing

In this stage two basic operations are performed. The first is converting the 24-bit color image to 8-bit grayscale image [9]. The second is removing noise on the grayscale image. These noises appear on the original image in the form of isolated spots or blobs due to physical stains. The purpose of noise removal is to facilitate edge detection, and consequently trace contours. Elimination of noise is usually performed by filtering, our choice was Gaussian filter, a Gaussian filter is a low-pass filter that helps in removing spurious noise, and thus smooth's images.This is done by convolving the image with a low-pass filter kernel. The kernel is a square array of pixels (a small image so to speak), as shown in Figure 2 Gaussian kernel (5, 5).

$$\frac{1}{273}\begin{bmatrix} 1 & 4 & 7 & 4 & 1 \\ 4 & 16 & 26 & 16 & 4 \\ 7 & 26 & 41 & 26 & 7 \\ 4 & 16 & 26 & 16 & 4 \\ 1 & 4 & 7 & 4 & 1 \end{bmatrix}$$

Figure 2 Gaussian kernel (5, 5).

Each pixel in the image gets multiplied by the Gaussian kernel. This is done by placing the center pixel of the kernel on the image pixel and multiplying the values in the original image with the pixels in the kernel that overlap. The values resulting from these multiplications are added up and that result is used for the value at the destination pixel.

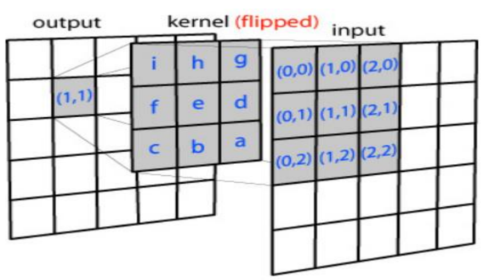

Figure 3 effect the Gaussian kernel.

The gray-scale image is further converted, to a binary image by a suitable threshold value. In order to obtain a binary image. Thresholding is a procedure of transforming an input gray scale image into a binary image, by using a specific threshold value. The goal of thresholding is to mark pixels that belong to the foreground with the same intensity, and those of the background with different intensities.

Localization of the plate

The objective of this stage is to perform the logical cut-off of the license plate from the entire car image. The basic idea for detecting the license plate in the car image is to search the entire image for the very distinctive shapes of numerals [10]. Once any shape is found, the process is repeated for the adjacent area of pixels until we find no more shapes. By the end of this step all character-like regions represent the borderless plate. Since the numerals in the license plate is usually of black on white background (or the opposite), tracing the contour of these objects is easily achieved [11]. However, to eliminate the contours that do not appear to be characters, we subject the contours to more rigors test. In this test the detected contour should have a minimum area of 100 pixels, a minimum width of 2 pixels, a minimum height of 8 pixels, a minimum aspect ratio of 0.25, and maximum aspect ratio of 1.0. We set a condition that the sequence must be greater than four characters, because license plates usually consist of more than four letters or numerals. Therefore, any sequence of numerals that are less than this value will not be considered as license plate, even if all the previous conditions are fulfilled.However, if the largest sequence is obtained, and all conditions are met, then the region of the image can be considered as the license plate.The last step is to determine the frame of the license plate In order to segment the characters. This is achieved by applying these measures:

Height of the plate = height of tallest character × 1.3,

Width of the plate = (sum of width of all characters) × 1.5,

(1.3, 1.5 are padding factors) determined by experimentation.

The plate frame is highlighted with red color and cut from the main image as seen in figure4(c).

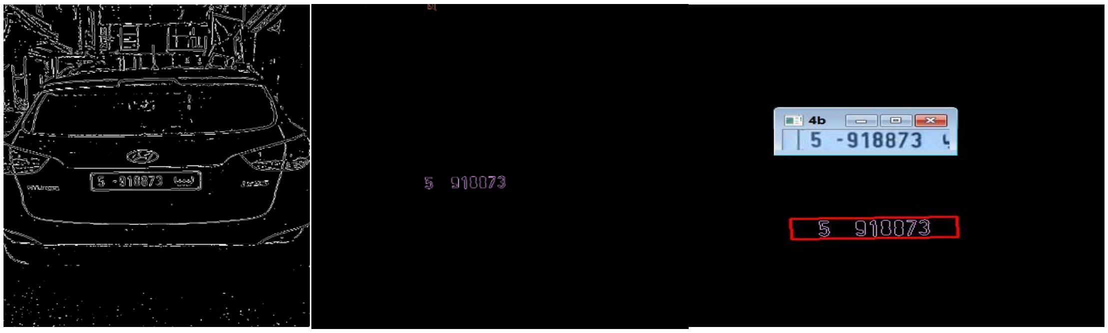

Figure 4 (a) multiple contoursdetected, (b) contours of numbers,(c) plate frame.

License Plate Segmentation using charactercontours

Image segmentation is the process of dissecting an image into different regions based on the characteristics of the pixels. By dividing the image into segments, we can make use of the regions of interest and limit our processing effort to these regions. The final purpose of segmentation is to identify objects or boundaries within the image that help in analyzing it.

Segmentation is one of the most difficult tasks in image processing. The important thing here is to know the location of an object in the image, the exact shape of that object, and what pixels belong to it. To perform segmentation of the characters within the license plate sub-image we use the contour method again to find the borders of each character. Beside that must applying several processing steps to distinguish such contour from other possible contours. These steps are called the morphological operations and their purpose is to clarify further the boundaries of the characters.

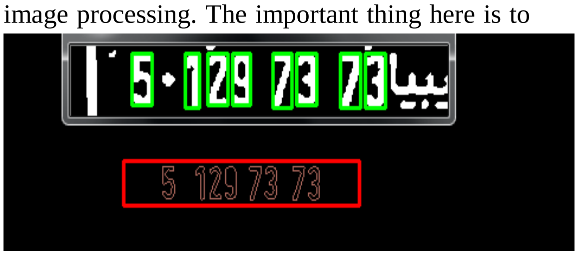

Figure 5 segments of contours of numbers.

Figure 6 isolated small boxes of numbers.

Recognition of the Characters

Recognition is the last stage in the entire process of automatic license plate recognition, ALPR. It also predated by several steps or sub-stages, namely image resizing or scaling and recognition process.

1- Scaling numeral images

As preparation for the recognition stage all small images of numerals will be scaled to a standard size. The standard size should have a height of 30 pixels and a width of 20, which means the new size will have 600 pixels. The scaling is important for the matching process.

2- The stages of the recognition process

The recognition process of numerals using the **k-Nearest Neighbor** algorithm is done in two stages:

First stage in which we train the algorithm, where we enter five different shapes for each numeral. Each time we enter one of these small images, we tell the algorithm the identity of the numeral represented by the small image.

Second stage is where recognition takes place. Here the algorithm takes the segmented and resized image of the numeral (input image) for comparison with the images stored in the images file. Because we chose k= 3 in the KNN algorithm, the input image is matched to the three closest images in the image file [12].

Figure 7 illustrates how the algorithm works with the different k values.

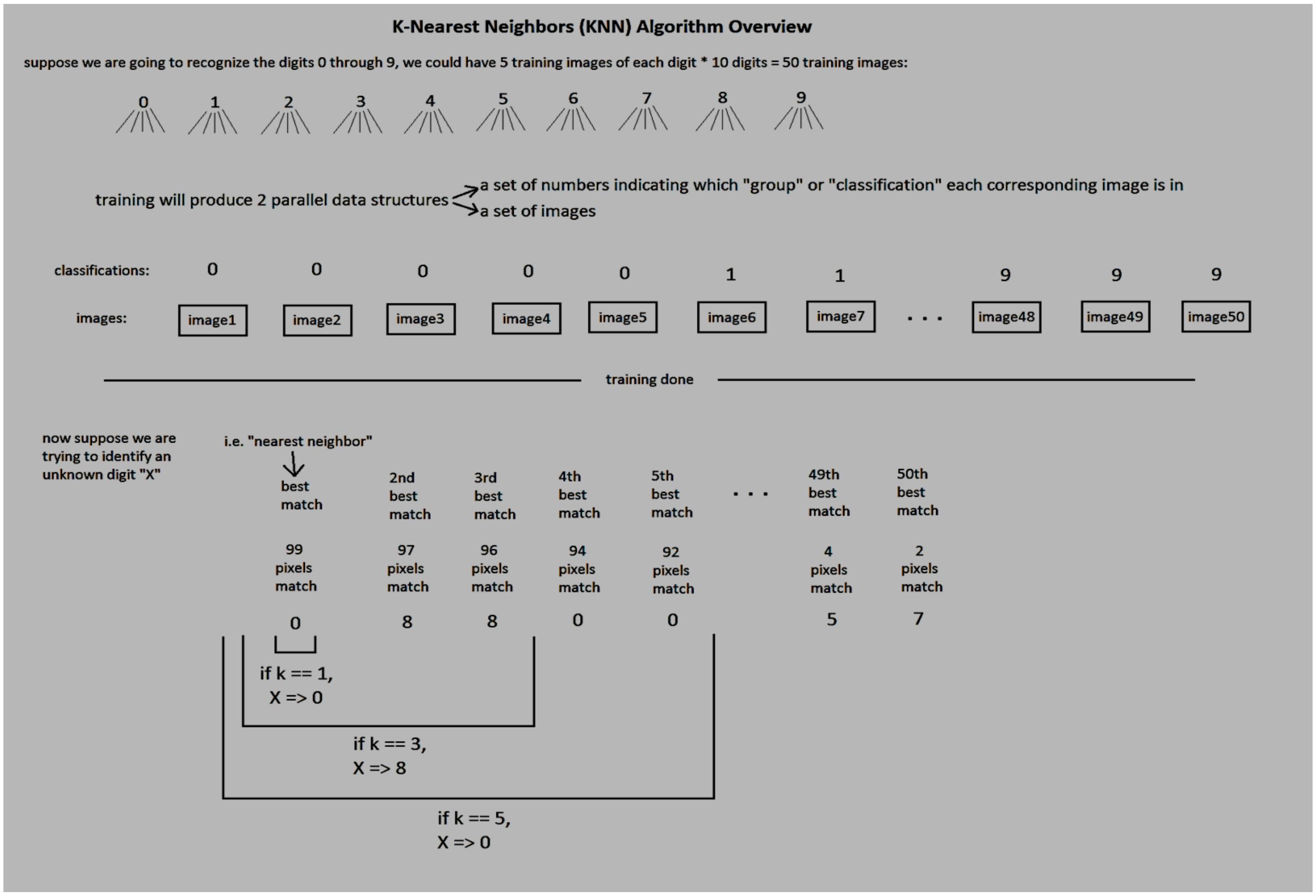

Figure 7 performance of KNN algorithm with different k values.

IV. Experimental Results

The method described in previous section is applied to localize and recognize the license plates of 100 various images which were license plates characters.
The method achieved accuracy over 97% for localizing plates. The recognition system
Implemented by KNN algorithm after segmentation of characters which get accuracy over 94% in image plate, after that identify numbers and achieve an accuracy over 91%.
In the following show a group of pictures with the steps for recognizing them and the final results of the recognition.
These images were selected from 91 images that were identified with succeed.

The pictures have poor quality, deformed, variable lighting, and images have some distortions, images with pins, image with dirt and dust or some oils covering the lower half of the numbers or the plate, and they have been successfully identified.
Figure 8(a) shows a poor quality picture, while picture 8(b) shows the process of extracting the plate thensegment it and identifying the numbers correctly.
Figure 9(a) shows a deformed picture, while picture 9(b) shows the process of extracting the plate thensegment it and identifying the numbers correctly.
Figure 10(a) shows images with pins, while picture 10(b) shows the process of extracting the plate thensegment it and identifying the numbers correctly.

Figure 11(a) shows image with dirt and dust or some oils covering the lower half of the numbers or the plate, while picture 11 (b) shows the process of extracting the plate thensegment it and identifying the numbers correctly.

Figure8(a) poor quality picture (b) results.

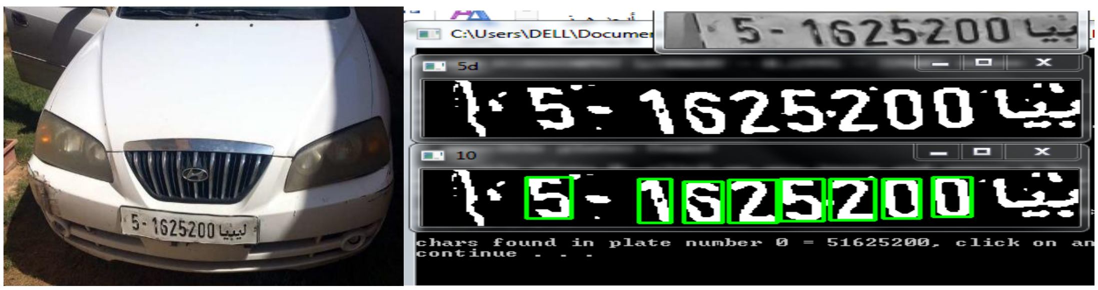

Figure9(a) deformed plate picture (b) results.

Figure10(a) shows images with pins (b) results.

ICTS40608122021-CM5001

Figure11(a) image with dust or some oils covering the lower half of the plate, (b) results.

V. Conclusion

We presented a license number plate recognition system, LNPRS, based on the KNN classification method. We tested the system by a sample of 100 images of car fronts of different car types. The camera took the images in a rather casual way, which means the real operating conditions would be either the same or better. The correct localization of the license plates of the sample images was 97%. The correct segmentation of characters within the plate reached a rate of 94% while the final recognition of numerals was 91% in normal operating conditions. However, if we control the process of image acquisition and guarantee the cleanness of license plates, the correct localization would easily be raised to almost 100%, while the correct recognition of numerals, out of the localized plates could achieve 98%. Misrecognition was mainly due to stains in the plates that forced the system to consider numeral 3 as 8 and sometimes 6 as 5. It can be inferred easily that keeping the quality of the plates is the best way for raising the recognition rates. Having done that, it would let the system jump to almost 100% in localization and not far from that for recognition.

References

[1] WenjingJia, Xiangjian He and Massimo Piccardi, "*Automatic License Plate Recognition: a Review* ", Department of Computer Systems, Faculty of Information Technology, University of Technology, Sydney

[2] Shan Du, Mahmoud Ibrahim, Mohamed Shehata, and WaelBadawy, "*Automatic License Plate Recognition (ALPR): A State-of-the-Art Review*", IEEE Transactions On Circuits And Systems For Video Technology, Vol. 23, No. 2, February 2013.

[3] V. Koval 1, V. Turchenko 1, V. Kochan 1, A. Sachenko 1, G. Markowsky, "*Smart License Plate Recognition System Based on Image Processing Using Neural Network*", IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, 8-10 September 2003, Lviv, Ukraine.

[4] Rentian Huang, HissamTawfik, and Atulya Nagar, "*Licence Plate Character Recognition Using Artificial Immune Technique*", Intelligent and Distributed Systems Lab, Deanery of Business and Computer Sciences, Liverpool Hope University, Liverpool, United Kingdom L16 9JD

[5] Soungsill Park*, Hyoseok Yoon and Seho Park, **"***Multi-Style License Plate Recognition System using K-Nearest Neighbors*", KSII Transactions On Internet And Information Systems VOL. 13, NO. 5, May 2019

[6] G. Shakhnarovich, T. Darrell and P. Indyk, "*Nearest-Neighbor Methods in Learning and Vision: Theory and Practice,*" The MIT Press, 2006.

[7] KhushbuKhamar, "*Short Text Classification Using kNN Based on Distance Function*", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, Issue 4, April 2013

[8] Siddhartha Roy, M.Saravanan2, "*Handwritten Character Recognition using K-NN Classification Algorithm*", Vol-3 Issue-5 2017 IJARIIE-ISSN(O)-2395-4396

[9] E. R. Lee, P. K. Kim, and H. J. Kim, "*Automatic recognition of a car license plate using color image processing*", in Proc. IEEE Int. Conf.Image Process., vol. 2. Nov. 1994, pp. 301–305.

[10] 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.

[11] S. Kim, D. Kim, Y. Ryu, and G. Kim, "*A robust license-plate extraction method under complex image conditions,*" presented at Pattern Recognition, 2002. Proceedings. 16th International Conference on, 2002.

[12] K.H.PavanAkshaylepcha, S.R.Adithya, G.Anuraag, K.Sathish, "*A Novel Methodology for License Plate Detection Using Knn Classifier*", IAETSD JOURNAL FOR ADVANCED RESEARCH IN APPLIED SCIENCES, VOLUME 5, ISSUE 3, MAR/2018