The North African Journal of Scientific Publishing (NAJSP)
مجلة شمال إفريقيا للنشر العلمي (NAJSP)
E-ISSN: 2959-4820
Volume 2, Issue 3, July – September 2024, Page No: 36-42
Website: https://najsp.com/index.php/home/index
SJIFactor 2023: 3.733 0.63 :2023 (AIF) معامل التأثير العربي ISI 2023: 0.383

Automatic Libyan Vehicles Plate Recognition Based on K-Nearest Neighbor (KNN)

Nouri Ali [1]*, Suleiman Omar [2], Iman Namroud [3], Mohsen Ibrahim [4]
[1] Higher Institute of Medical Sciences and Technologies, Bani Walid, Libya
[2,4] Higher Institute of Engineering Technology, Bani Walid, Libya
[3] Collage of Electronics Engineering, Bani Walid, Libya

التعرف التلقائي على لوحات المركبات الليبية بناءً على (KNN)

نوري علي*[1]، سليمان عمر[2]، إيمان نمرود[3]، محسن إبراهيم[4]
[1] المعهد العالي للتقنيات الطبية، بني وليد، ليبيا
[4،2] المعهد العالي للتقنيات الهندسية، بني وليد، ليبيا
[3] كلية التقنية الإلكترونية، بني وليد، ليبيا

*Corresponding author: nouri.alhashi@imst.edu.ly

Received: May 18, 2024	Accepted: July 10, 2024	Published: July 19, 2024

Abstract:
There are numerous uses for license plate recognition and identification (LPR/I) in traffic systems, including border and customs checkpoints, red light enforcement, and highway electronic toll collection. It has been widely utilized in port, airport gate surveillance, highway and bridge charges and All of that drives our efforts to develop an effective LPR/I system tailored for Libyan license plate recognition. Using the current, legally needed license plate, automated license plate reading is a very practical and helpful method. This paper discusses the use of computer vision for real-time license plate recognition and vehicle detection. For detection and recognition, respectively, the primary focus will be on integrating vertical edge matching with k-nearest neighbor. Additionally, to effectively localize, regions must be further enhanced by removing non-text regions while keeping candidate text regions. However, morphology is necessary to distinguish some combined features. And that is a decent, logical system with k-nearest to recognize the character based on it. The two main components of the suggested method are the extraction of the plate region, segmentation of the characters and digits, and KNN-based plate character identification. In this work text categorization involves two processes: the training process and classification process Initially, the system is trained to comprehend the appearance of each category using a set of documents that have already been categorized. Second, the classifier classifies newly received documents by using the training 'model'. By using training documents with known categories, the K-Nearest neighbor classification approach locates the new sample document's closest neighbor among all of them. The new document category can be found thanks to these neighbors. Statistical template matching is used for the recognition of plate characters. On actual photos, the suggested algorithm's performance has been evaluated. We observed that our system performs better in car license plate recognition based on the trial findings while testing, and we obtained a strong result. The suggested technique verified which extraction rates are high by using different vertical edge matching with the k-nearest natural images.

Keywords: License Plate Recognition, Edge-Detection, K-Nearest Neighbor, Template Matching, Plate region extraction, Segmentation, Morphology, smearing.

الملخص
هناك العديد من الاستخدامات للتعرف على لوحة الترخيص وتحديد هويتها (LPR) في أنظمة المرور، بما في ذلك نقاط التفتيش الحدودية والجمركية، وجمع رسوم المرور الإلكترونية على الطرق السريعة. لقد تم استخدامه على نطاق واسع في مراقبة الموانئ وبوابات المطارات ورسوم الطرق السريعة والجسور وكل ذلك يدفع جهودنا لتطوير نظام LPR فعال مصمم خصيصًا للتعرف على لوحة الترخيص الليبية. باستخدام لوحة الترخيص الحالية والمطلوبة قانونًا، تعد القراءة الآلية للوحة الترخيص طريقة عملية ومفيدة للغاية. تناقش هذه الورقة استخدام

الرؤية الحاسوبية للتعرف على لوحة الترخيص في الوقت الحقيقي، سيكون التركيز الأساسي على تموضع المنطقة المستهدفة مع استخدام KNN. بالإضافة إلى ذلك، يجب تحسين المناطق بشكل أكبر عن طريق إزالة المناطق غير المستهدفة مع الاحتفاظ بمناطق النص المرشحة. ومع ذلك، التشكل ضروري للتمييز بين بعض الميزات مجتمعة. وهذا نظام منطقي لائق مع KNN. المكونان الرئيسيان للطريقة المقترحة هما استخراج منطقة اللوحة، وتجزئة الأحرف والأرقام، وتحديد أحرف اللوحة المستندة إلى K-NN. يتضمن تصنيف النص في هذا العمل عمليتين: عملية التدريب وعملية التصنيف. في البداية، يتم تدريب النظام على فهم مظهر كل فئة باستخدام مجموعة من المستندات التي تم تصنيفها بالفعل. ثانيًا، يقوم المصنف بتصنيف المستندات المستلمة حديثًا باستخدام "نموذج" التدريب. باستخدام مستندات التدريب ذات الفئات المعروفة، يقوم أسلوب تصنيف الجوار KNN بتحديد موقع أقرب جار لمستند العينة الجديد من بين كل هذه المستندات. يمكن العثور على فئة المستند الجديدة بفضل هؤلاء الجيران. ويتم استخدام مطابقة القالب الإحصائي للتعرف على أحرف اللوحة. في الصور الفعلية، تم تقييم أداء الخوارزمية المقترحة. لاحظنا أن أداء نظامنا أفضل في التعرف على لوحة ترخيص السيارة بناءً على نتائج التجربة.. أثناء الاختبار، حصلنا على نتيجة قوية. تحققت التقنية المقترحة من معدلات الاستخراج المرتفعة باستخدام حواف رأسية مختلفة تتوافق مع الصور الطبيعية الأقرب إلى

الكلمات المفتاحية: التعرف على لوحة الترخيص، KNN، مطابقة القالب استخراج منطقة اللوحة التجزئة التشكل، التلطيخ.

1 Introduction

There are variations in license plate standards across the globe. To sustain law enforcement on the highways, real-time LPR is an important type of automatic vehicle identification that not only recognizes but also distinguishes in automatic monitoring of traffic laws, access control systems, and toll road violations [1]. The majority of researchers face the issue of having to work quickly because they are anticipated to be used in real-time contexts [2]. Certain LPR techniques result in color photographs that differ significantly depending on the lighting circumstances in which they were shot, and the vehicle's color may match that of the background [3]. In recent decades, the varying forms, styles, colors, and sizes of license plates across many nations, together with their practical uses, have drawn numerous researchers from various fields to create systems for vehicle identification [4]. Therefore, vertical edge matching based algorithm to extract LPR from image [3]. Template matching, which uses an efficient algorithm for character recognition, matches numbers and characters inside a database. The best similarity is determined by comparing the character image with those in the database [1]. K-Nearest Neighbor is considered to be the best method for text categorization. Certain initial files, also referred to as training documents, are necessary for the KNN classification. Before using the approach, certain documents or LPR categories are known. A newly created document that lacks a category is compared to all of these training materials based on the phrases they share. Ultimately, the documents that the new document most closely resembles are identified, and a category is allocated by that determination [5]. Generally, LPR is composed of many stages; pre-processing, license plate localization and detection, character segmentation and character recognition, and each part may contain several sub-stages.

2 Proposed Method

This system's architecture is solely dependent on image processing methods and algorithms. There are numerous stages in this system. Initially, a picture of a car is taken. Next, the algorithm makes an effort to locate the license plate locally.

2.1 Image pre-processing

Pre-processing involves attempting to improve the image's suitability for the subsequent steps, which include grayscale processing, image complementing, filtering, and binarization [7]. To facilitate the identification of the target's outline, the input image was converted to grayscale format. Firstly, as in figure 1 shows picture is often in the RGB colour space we first convert the 3 colour components to intensity components (grey images) [7]. Then employ a weighted median filter to reduce noise, after the filtering step, a great part of the noise will be have removed while the edges in the image are still preserved. Additionally, we make an effort to expose the image by adjusting its contrast and brightness. Therefore, in order to distribute light and obtain an acceptable image, we utilize histogram equalization.

Figure 1 Original image.

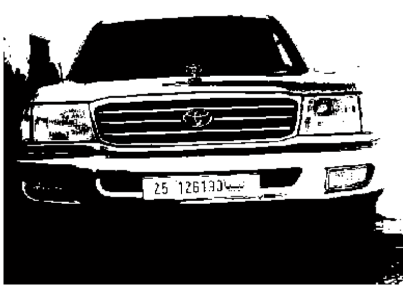


Figure 2 binary image.

2.2 Edge Detection

An edge picture that corresponds to a given grayscale image can be obtained using Prewitt or Sobel masks, As seen in fig. 2 above, an image with a lot of edges—both horizontal and vertical—usually has more horizontal lines than vertical lines. To be able to isolate the plate characters, the segmentation process involves two main setup, vertical and horizontal projection. In vertical projection the character-related columns of the plate picture remain and other ways while being projected horizontally.it crops any extra rows which not belong to character otherwise, with image is processed along vertical and horizontal scan White pixels are turned to black if the number of white pixels is fewer than the intended threshold or more than any other desired threshold [2]. Threshold summation values are between 240-350 for each row and 190-350 for each column as shown in figure 3 and figure 4.

Figure.3 horizontal edge image.

Figure.4 vertical edge image.

Both vertical and horizontal projection are accomplished by tracing the ones in the plate image raw wise for vertical projection, and then column wise for horizontal projection as shown in figure 5.

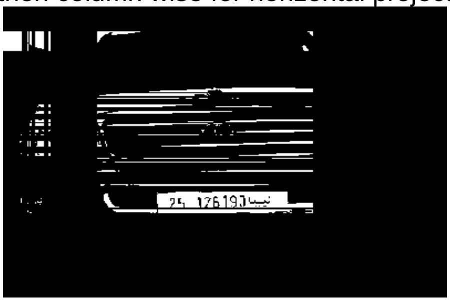


Figure 4 entire image (vertical & horizontal).

2.3 Plate Region Localization

In computer vision, finding an object in an image is one of the most common challenges and is the first step towards solving the recognition problem. Based on the threshold value, a grayscale image is converted to black and white, as shown in figure 5

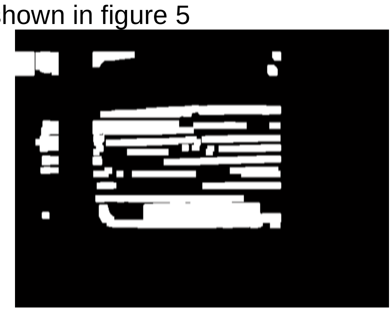


Figure 5 Localization process.

2.4 Plate region extraction

To identify the location of the plate, the image is subjected to plate region extraction. There might be more than one possible zone for the location of the plate, though. Using smearing and filtering operations, several criterion tests are conducted to the image in order to identify the precise region and remove the remaining regions. The processed image after this stage is as shown in figure6

Figure 6 Extraction process.

By applying some calculate averages of rows and columns with specific threshold, the following image as shown in Figure 7.

Figure 7 Determine process.

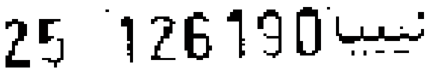


Figure 8 Region plate.

2.5 Plate segmentation

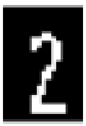
In this phase, we want to use the recovered plate picture from the previous phase to build sub-images. Typically, once the image has been converted to a binary image, we work with the numbers zero and one to represent the background and foreground. Operational Length Coding Algorithm. The RUN mean takes the pixels in same the row in blocks. That means, every block of RUN is represented in horizontal projection, and every horizontal projection is calculated from run length code. The equivalency table represents the labeling algorithm. Most labeling algorithms have focused on resolving the equivalency table, which is easy to build and uses less memory. The process uses a run-length encoding representation. Conversion of the original binary image to run-length encoded format is easily parallelized by processing multiple rows in parallel. The run-length encoded format is much more compact than the binary image (individual runs have a single label), Consequently, compared to the traditional technique, the sequential label propagation stage is substantially faster. The algorithm's specifics are shown below. The phases of our execution are as follows:

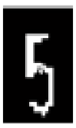
1. Pixels are converted to runs in parallel by rows.
2. Initial labeling and propagation of labels.
3. Equivalence table resolution.
4. Translating run labels to connected components.

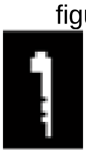
We have been used 8-neighbors where they share at least one corner, the positions are [i+1,j] ,[i-1,j] ,[i,j+1] , [i,j-1], [i+1,j+1], [i+1,j-1] , [i-1,j+1] and [i-1,j-1] as shown in fig 9 below

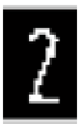
	[i,j]	

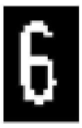
Figure 9 8-Neighbhorhoods for rectangular image tessellation pixel [i,j] is located in center of each figure.

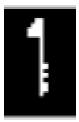

(a)

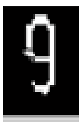

(b)

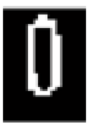

(c)


(d)


(e)


(f)


(g)


(h)

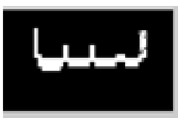

(i)

Figure 10 Objects which extracted.

2.6 K-Nearest Neighbour Classification

Because license plate characters all have the same font, character recognition is simple and dependable when template matching is used. This method is also more noise-tolerant. The personalities have become more relatable. Normalization is the process of refining the characters so that entire character sides are free of unnecessary white space into blocks containing no extra white space [5]. Next, each character is proportionately fitted. A fitting strategy is required for the template to match the characters in the database, input images must be equal-sized with the database characters. Here the characters are fit to 36*18. The extracted cut from plate and the characters on database are equal-sized. Template matching is the next stage. With the exception of word Libya "ليبيا" dealing with it as one block [2]. Template Matching is an effective algorithm for recognition of characters. The character image is compared with the ones in the database and best similarity is measured.

The k-nearest neighbor algorithm is one of the most basic classification techniques used in data mining and machine learning [6]. Because of its simplicity and usefulness, this classification approach is the most widely used. Doesn't require a fitted model and has demonstrated better effectiveness in identifying various kinds of data [5]. K-NN algorithm is a non-parametric method used for classification and regression, in both cases, the input consists of the k closest training in the feature space [8]. The output depends on whether k-NN is used for classification and regression.

The following steps describe how the technique works. K-NN has been used to estimate the numbers and characters of license plates.

1. Data collection and sample creation. The stage that creates the training dataset for detection and recognition required for Template Matching and KNN is the data collection procedure. where the dataset and initial structuring.
2. Define libraries and trained data into the main program used in the system is Python which deal with matrixes 36*18.
3. Splitting dataset into a training dataset and test dataset. In the training phase, stores all the training data set in set according to the selected features. In the test phase: computes the distance between the new feature vector and all the stored features (training data).
4. Extracting feature vectors:

The features mentioned above are used to train the KNN classifier initially. To verify the classifier's accuracy, test images are used to test the images. We need to make sure that both sets training set and test set are mutually disjoint. The training and test sets must be mutually disjoint. All of the images from the training set are mapped into D-dimensional space during the training phase, where D is the feature vector's dimension. Now given are unknown image, we map it into D-dimensional vector space.

5. Evaluating classifier of test dataset: Now, we found the labels of the k closest training photos, compare the distance between a test image and all training samples, and then use majority vote to assign the label for the test image.

3 Experiment Results and Discussion

For each feature, several training samples of the training data are used in the experiment. After pre-processing the training data and creating a dictionary of words, it will perform information gain to identify the features most pertinent to the current classification task, calculate weights for each feature in the dictionary, and create a k-NN model that the classifier can use to further classify newly incoming documents. Finally, we can measure the accuracy of k-NN classifier, the performance of k-Nearest Neighbor classifier is measured in terms. The query photos for the system evaluation were selected from a large number of images in the database collection that represented different anatomical locations. Precision (P) is used to represent the retrieval accuracy which is the percentage of similar images retrieved with respect to the number of retrieved images and Recall(R) is the percentage of relevant images among all possible relevant images. Precision and Recall is defined as:

$$P = \frac{\text{Number of relevant images retrieved}}{\text{Total number of iamegs retrieved}}$$

$$R = \frac{\text{Number of relevant images retrrieved}}{\text{Total number of relevant image in database}}$$

The aforementioned equations lead us to the conclusion that retrieval accuracy is measured by precision and recall. Table 1 shows Retrieval Performance of Template Matching, which Method obtained by varying the images of gives the screen shot of the retrieval of a sample query image. That shows Average Precision and Recall graph.

Table 1 Results of test.

Classifier K-NN	Total number of images	Feature	Average Precision	Average Recall
K=7	10	Template Matching	92	48

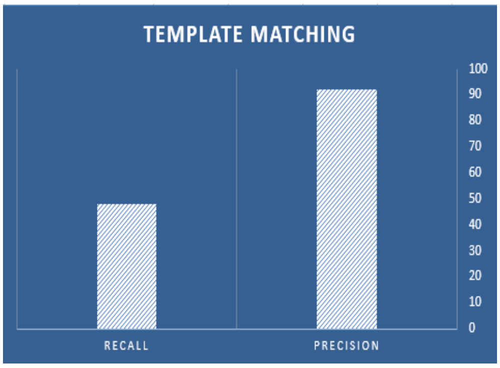


Figure 19 Performance evaluation.

Figure 20 Performance evaluation.

4: conclusion

In this paper, we demonstrated application software intended for license plate identification on automobiles. Firstly, we extracted the plate location, Next, we used segmentation to separate each of the plate characters separately. then lastly used template matching to recognize plate characters via correlation. This method is intended to identify Libyan license plates, and it has undergone extensive testing on a vast array of photos. Finally, it is proved the characters % 92 for Precision and % 48 for Recall accurate. This system can be redesigned for multinational car license plates in future studies. Thus, after the implementation of KNN using the Template Matching feature, it is found that the template matching feature performs the best with an accuracy of 92%.

References

[1] Serkan.Ozbay and Ergun Ercelebi, “Automatic Vehicle Identification by Plate Recognition".Tranaction on Engineering,Computing and Technology v9 November 2005 ISSN 1305-5313.

[2] Iman Namroud,Khlood Belhasan and Moammer Adbullaziz,"Libyan License Plate Recogntion",LICEET 2018,Troploi Libya.

[3] Mei Yu and Yong Deak Kim,"An Approach to Korean License Plate Recognition Based on Vertical Edge Matching",Ajou University Suwon,442-749,Korea,0-7803-6583-6/00,2000

[4] Areel Abugrain and Ali Ganoun,"Libyan Vehicle License Plate Localization System using Neural Networks",Libya International Conference on Electrical Engineering and Technologies,Tripoli-Radisson blu-4-6/04.2018.

[5] A.Kousar Nikhath,K.Subrahmanyam and R.Vasavi,"Building a K-Nearest Neighbor Classifier for Text Categorization",International Journal of Computer Sciences and Information Technologies,Vol 7,2016,254-256,ISSN :0975-9646
[6] Ayushi Dalmia,"Handwritten Digit Recognition using K-Nearest Neighbour",201307565.
[7] P.A.Charde.S.D.Lokhande,"Classification using K-Nearest Neighbor for Brain Image Retrieval",International Journal of Scientific Reserch,Volume 4,Issue 8,Auguest-2013.
[8] Mohammed Elborawi and Ali Ganoun,"Real TimevEHICLE Detection and License Plate Recognition Sytem", Libyan International Conference on Electrical Engineering and Technologies, Tripoli-Radisson blu-4-6/04/2018
[9] Ferdose Abdelsalam Almarimi and Moammer Hassan Dhaw,"Human Iris Recognition using Corellation Matrix", Libyan International Conference on Electrical Engineering and Technologies, Tripoli-Radisson blu-4-6/04/2018
[10] M.Seetha,K.V.N.Sunithaand G.Malini Devi,"Performance Assessment of Neural Network and k-Nearest Neighbour Classification with Random Subwindows",International Jouranl of Machine Learning and Computing, Vol 2,No.6,December 2012.
[11] Norhidayu Binit Abdul Hamid and Nilam Nur Binti Amir sjarif,"Handwriten Recognition using SVM,KNN AND Neural Network",
[12] Carlos F.Nava-Duenas and Felix F.Gonzalez-Navarro,"OCR for Unreadable Characters on PCBs using GSC Algorithim and KNN Classifer".
[13] Weiqiang Ren,Yinan Yu,Junge Zhang and Kaiqi Huang,"Learing Convolutional NonLinear Features for K-Nearest Neighbor Image Classification".
[14] Oliviu Matei ,Honriu Valean and Petrica C.Pop,"Optical Character Recognition in Real Environments using Neural Networks and K-Nearest Neighbor,ReseachGate.
[15] Gita Sinha,Mrs.Rajneesh Rani and Prof.Ranu Dhir,"Handwritten Gurumukhi Character Recognition using K-NN and SVM Classifier",International Journal of Advanced Research in Computer Sciences and Software Engineering",Volume 2,Issue 6,June 2012.
[16] Khalid Alhatib,Hassan Najadat,Ismail Hmeidi and Mohammed K. Ali Shatnawi,"Stock Price Prediction using K-nearest Neighbor Algorithm", International Journal of Business,Humanies and Technology",Vol 3, No. 3 March 2013.
[17] Shivani J,Surya N,"Analysis of Performance of KNN Classifier in Recognition of Handwritten Digits",International Journal of Advanced Research in Engineering Sciences & Technology",e-ISSN:2393-9877,P-ISSN:2394-2444,Volume 5,Issue 3,March-2018.
[18] Maryan Kuhkan,"Method to Improve the Accuracy of K-Nearest Neighbor Algorithm",International Journal of Computer Engineering and Information Technology,Vol 8,No 6,June 2016,90-95,E-ISSN 2412-8856.
[19] Krishna A Patel and Vinit Gupta,"Review on Handwritten Digit Recognizing System",International Journal of Advanced Reserch in Computer Sciences and Management Studies",Volume 3,Issue 4,April 2015,ISSN:2321-7782.
[20] Hassan Shee Khamis,Kipruto W.Cheruiyot,Stephen Kimani,"Application of K-Nearest Neighbour Classification in Medical Data Mining",International Journal of Information and Communication Technology Resech,Volume 4, No 4,April 2014,ISSN:2223-4985.
[21] Gulen Toker,Oznur Kirmemis,"Text Categorization using K-nearest Neighbor Classification".
[22] Jingwen Sun,Weixing Du,Niancai Shi,"A survey of K-NN Algorithm,Information Engineering and Applied Computing 2018.