

RESEARCH

Open Access



Performance enhancement method for multiple license plate recognition in challenging environments

Khurram Khan^{1*}, Abid Imran², Hafiz Zia Ur Rehman³, Adnan Fazil¹, Muhammad Zakwan¹ and Zahid Mahmood⁴

*Correspondence:

khurram.jadoon@mail.au.edu.pk

¹ Department of Avionics
Engineering, Air University,
Islamabad, Pakistan

Full list of author information is
available at the end of the article

Abstract

Multiple-license plate recognition is gaining popularity in the Intelligent Transport System (ITS) applications for security monitoring and surveillance. Advancements in acquisition devices have increased the availability of high definition (HD) images, which can capture images of multiple vehicles. Since license plate (LP) occupies a relatively small portion of an image, therefore, detection of LP in an image is considered a challenging task. Moreover, the overall performance deteriorates when the aforementioned factor combines with varying illumination conditions, such as night, dusk, and rainy. As it is difficult to locate a small object in an entire image, this paper proposes a two-step approach for plate localization in challenging conditions. In the first step, the Faster-Region-based Convolutional Neural Network algorithm (Faster R-CNN) is used to detect all the vehicles in an image, which results in scaled information to locate plates. In the second step, morphological operations are employed to reduce non-plate regions. Meanwhile, geometric properties are used to localize plates in the HSI color space. This approach increases accuracy and reduces processing time. For character recognition, the look-up table (LUT) classifier using adaptive boosting with modified census transform (MCT) as a feature extractor is used. Both proposed plate detection and character recognition methods have significantly outperformed conventional approaches in terms of precision and recall for multiple plate recognition.

Keywords: Deep convolutional neural network, Automatic number plate recognition, Modified census transform, Image morphology

1 Introduction

With the ever-increasing traffic situations in modern cities, the demand for the ITS is also increasing rapidly. License Plate Recognition (LPR) is a crucial component of an ITS, which is used to identify vehicles based on their number plates. The LPR includes license plate localization to find the location of the plate in an image, followed by segmentation and recognition of alphanumeric characters of the localized plate. Most of the LPR methods are capable of recognizing a single vehicle in an image [1].

In recent years, there is a considerable increase in problems of traffic congestion, security monitoring, and over-speeding in modern cities [2]. The afore-described phenomenon has increased the demand for identifying multiple vehicles in an image. A

high-resolution camera can monitor multiple lanes containing several vehicles. However, recognizing multiple vehicles becomes challenging as some plates will have a smaller size or low resolution, for instance, based on distance from the camera, different background colors, distortions, and different contrast as shown in Fig. 1. Moreover, the HD images also increase computational cost [3].

Several methods have been proposed for license plate recognition over the years, such as template matching [4], artificial neural networks [5], adaptive boosting [6], and Support Vector Machines (SVM) [7]. Most of these methods performed well in constrained environments, such as uniform illumination and fixed plates size. In the recent past, Deep Learning (DL) methods have been widely used as powerful tools for image recognition applications [8]. Henceforth researchers have also developed various Convolutional Neural Network (CNN) architectures for plate recognition [9]. Cascaded CNN [10] has a high computational cost and multi-CNN [11] struggles on varying parameters of size and angle. Faster-RCNN has high speed but performance deteriorates on low-resolution plates and non-uniform illuminations [12]. Therefore, it can be concluded that existing approaches have yet to achieve a satisfactory level when exposed to unconstrained environments, such as varying illuminations conditions, colored background plates, and variations in plate, and font sizes. Therefore, this paper proposes a novel MLPR system that is capable to handle all the aforementioned problems. The major contributions of this paper are as follows:

- The proposed technique improves the accuracy of plate detection in challenging environments that have non-uniform illumination and low resolution (based on distance from the camera). Our proposed MLPR technique divides plate detection problems into vehicle detection and plate localization, which results in scaled information for plate localization and helps to remove background noise and clutters.
- The proposed plate recognition algorithm does not put a restriction for uniform light conditions, low resolution, or angular plates. Moreover, the character recognition part is robust to varying illumination, low resolution, different orientations, and multiple fonts. Experimental results have shown that the character size of 6×9 pixels is also recognized effectively.

In this paper, Faster-RCNN is used for vehicle detection followed by plate localization using morphological operations in the HSI color space. Geometric properties of area



Fig. 1 Korean license plates

and aspect ratio of connected pixels are used for character segmentation. Moreover, this paper uses texture-based feature extraction method MCT, which is robust to illumination changes and low resolution [13], with lookup table classifier in boosting framework for character recognition.

The remainder of this paper is organized as follows. Related work is reviewed in Sect. 2. Section 3 presents the proposed license plate detection and recognition method, detailed simulation results are presented in Sect. 4. Finally, the conclusion is given in Sect. 5.

2 Related work

This section briefly introduces the recent advances and published works in the plate recognition domain.

2.1 License plate detection

Most of the existing work on plate detection target a single vehicle in an image. Therefore, the demand for multiple plate detection has increased considerably owing to an increase in multilane structure in modern cities. Edge detection methods consider an area with a higher density of characters as an LP. Combining this property with geometric properties of plates has been widely used to extract LPs. Vertical edge detection is more robust compared with horizontal edge detection, which provides inaccurate results owing to errors due to the car bumper area [3]. A fast and robust vertical edge detection method was proposed that increases the speed of detection by eliminating unwanted lines [14]. Yepez et al. [15] proposed a plate detection method based on only morphological operations. They developed an algorithm to select appropriate structuring element (SE) from a set of SEs by training these SEs on the whole dataset. This approach could not perform well for multiple license plate recognition, due to variations in the size of plates in an image.

In [16], the block processing method was proposed which detects the area of LP by finding the maximum edge magnitudes among the blocks. Connected component labeling (CCL) [17] was used in binary images to label groups of connected pixels and use attributes such as the height/width ratio, and area to localize the plate. In [18], a character-based approach was used to localize an LP by calculating the distance between the characters on the LP. Rizwan et al. [2] proposed a method for detecting Chinese plates by taking advantage of the chromatic component of the YD_bD_r space model and eliminates non-plate regions by using an average of energy map and edge information of the plate.

Kamal et al. combined AdaBoost [19] with Haar-like features in a cascaded manner for License Plate Detection (LPD) [6] and genetic algorithms [20] have been used to classify and identify plates based on color information by using geometric attributes CCL for localization. In [21], the authors proposed an entropy-based feature selection method, followed by SVM for classification for plate detection. The proposed method performed segmentation by identification of the luminance channel [22] and then used Otsu's thresholding for binary segmentation of that channel. This method was only able to produce reasonable results on a small number of images.

Recently, deep learning architectures have also been used for the task of the LPD [23]. The CNN has been used in a cascaded form, where the first CNN classifier searches for any text on the image, and 2nd classifier is used to reject false positives, i.e., any other text from the text on LP [24]. Xiang et al. [25] proposed CNN-based network that extracts low and high features at several stages to distinguish details of plate and background followed by three-loss layer architecture for accurate plate detection. To enhance efficiency, researchers in [26] used the advanced structure of Faster R-CNN [27], which directly detects the plates in an end-to-end manner. A modified YOLO [28] was used for license plate localization, which had the capability of detecting license plates that had different variations like rotation, skewness, and different orientations, this method had high computational complexity.

Faster-RCNN with VGG 16 as a feature extraction method, without utilizing the fully connected layers, was used for LPL [29]. Moreover, DL-based Image enhancement [30] and denoising [31] techniques can be applied to improve the overall license plate detection accuracy. Moreover, LPD can be related to scene text detection as a special case, several such methods have been presented for text/ number detection [32]. Xue et al. [33] technique identify dense text boundary points to detect scene text, which characteristically comprehends the shape and location of text lines in contrast to other methods using segmentation-based techniques. These methods require strong contextual information. In addition to plate characters not having a strong relation to context semantic information, issues like varying light conditions combined with low resolution, and angle further deteriorates overall performance of such techniques.

2.2 Character segmentation

Character segmentation is a key step used to isolate characters for recognition. Most popular methods used geometric properties of area and aspect ratio [34], horizontal and vertical projection methods of characters were used to segment the plates [35], and also multiple features were combined to segment the characters of LP [36].

A Convolutional Neural Network (CNN) based two-stage process is proposed in [37] to segment and recognize characters (0–9, A–Z). Tarigan et al. [38] proposed an LP segmentation technique that consists of horizontal character segmentation, connected component labeling, verification, and scaling.

2.3 Character recognition

Recently, many methods and classification techniques have been proposed for recognition [23]. The template-matching method [4], calculates the correlation between character and templates and the maximum correlation valued template is considered as a character. However, it has shown poor performance in variable character size, noise, and rotation. Multilayer NN is trained to recognize characters [5]. The multistage classifier was used to recognize characters with lower case, upper case, digits, and two-line plates. This technique's performance deteriorates with varying illumination and small size of characters [39].

In [40], CNN and bi-directional long short memory (BLSTM) are combined for plate recognition. The CNN was used as feature extraction due to its high discrimination ability and BLSTM is capable of extracting context information from past information

followed by dense cluster voting (DCV) for classification. Bulan et al. [41] proposed a segmentation and annotation free method for plate recognition. They proposed a two-stage classifier, which first used a winnows classifier for candidate region extraction followed by CNN for plate classification. For optical character recognition, a segmentation-free approach using hidden Markov models (HMMs) was proposed. In [42], a research group developed a robust ALPR technique in unconstrained environments. In [43], researchers presented a novel architecture for Chinese LPR by cascading CNN and extremal learning machines (ELMs). The CNN is applied for feature extraction and the ELM is used as a classifier, which yields encouraging results with short training time. In [10], a cascaded recurrent neural network (RNN) method integrated with a short-term memory (STM) is proposed to recognize the sequential features. These features are extracted from the whole license plate via CNN. Proposed License Plate Detection and Recognition Method. In addition, adapting encoder–decoder architecture character recognition can be regarded as scene text recognition (STR) in which the encoder extracts features followed by character sequence decoding. The RNN with wide applications in natural language processing has been used a lot in the STR [44]. However, one major drawback of RNN is processing data in order. YU et al. [45] presented semantic reasoning module, which utilizes parallel transmission to mitigate the limitation of one-way transmission of context. Similarly, in [46], researchers utilized two transformers, one for an image to character mapping and the second for the character to word, respectively.

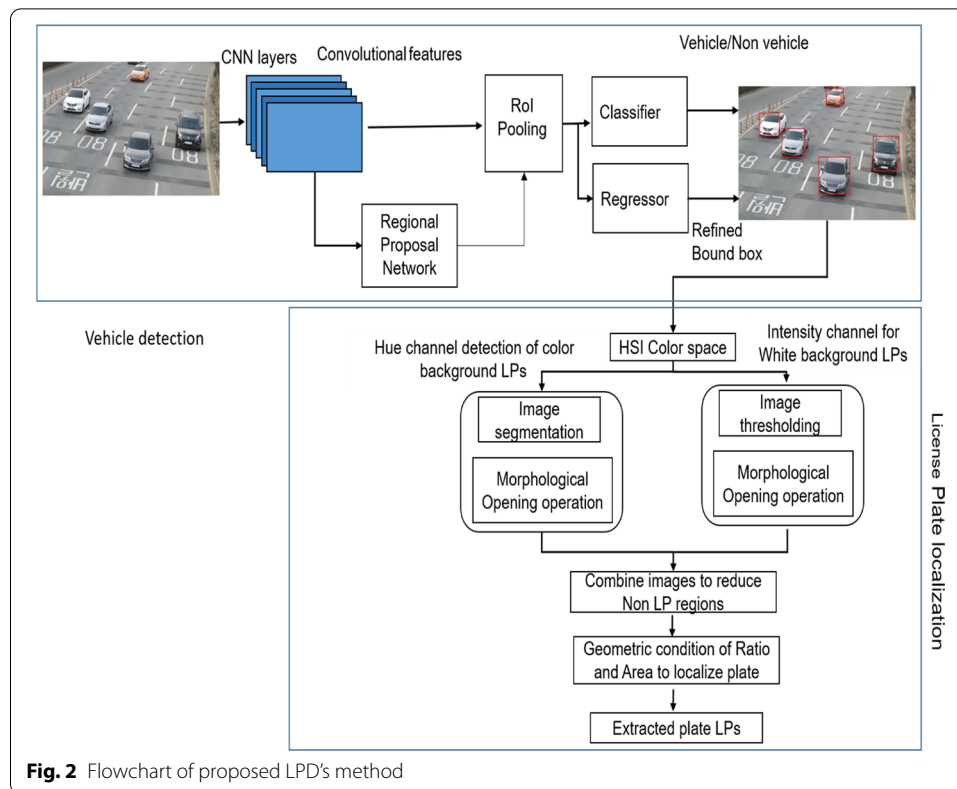
3 Proposed license plate detection and recognition method

This section describes the architecture of the proposed system. Figure 2 shows the overall architecture of the license plate detection method. For readers' easy understanding, we divide our developed method into the following interconnected steps.

3.1 Vehicle detection

Object detection is becoming a complex problem with an increase in applications, such as multiple object tracking [47], and self-driving cars. Many handcrafted methods, such as the HoG and the Haar features [47] and deep learning (R-CNN [48], Faster-RCNN [27], YOLO [49] methods were proposed recently. However, some have slow processing speed and others have a low accuracy rate. Faster-RCNN has shown the best detection rates among deep learning object detectors with real-time processing capabilities. However, Faster-RCNN performance deteriorates for small object detection, for instance, the LP localization in our case. Therefore, this paper uses Faster R-CNN for vehicle detection that helps in providing relevant and scaled information in an image. Moreover, Faster-RCNN shows excellent results for vehicle detection as vehicle size is large as compared to the plate's size in multiple license plate detection cases.

Faster-RCNN is divided into two parts: region proposal network (RPN) that generates a proposal for vehicles region, followed by fast R-CNN for vehicle/non-vehicle classification and to efficiently refine the proposal and detect the vehicles. To generate feature maps of the input image, we employ a pre-trained VGG-16 [50] model consisting of 13 convolutional, 5 max-pooling, and FC layers. The feature maps are fed to RPN which scans each map by sliding window and generates proposals with bound boxes for vehicle region. For multiple vehicle detection scenarios, the network has to detect vehicles



of multiple scales and aspect ratios as the distance between vehicle and camera varies. To deal with variable scales, anchors in RPN are introduced, which uses three scales (128×128 , 256×256 and 512×512) and 3 aspect ratios ($1 : 1$, $1 : 2$, $2 : 1$), which results in 9 anchors at each location. As the size of each region proposal is different from each other, it is difficult to make efficient architecture for different sizes. Region of interest (RoI) pooling simplifies the problem and extracts fixed-size feature representation. The features from RoI pooling are flattened into vectors. Eventually, these vectors are fed into two fully connected layers, the one for vehicle/non-vehicle classification based on each RoI SoftMax probability, and the other for predicting the rectangular coordinates for each region. In this method, faster-RCNN architecture is trained using stochastic gradient descent with momentum (SGDM) that minimizes the error and quickly updates the weights. SGD uses only one sample data set from the training data to update weights while the gradient descent (GD) method must consider all training datasets to update the weights/parameters. The initial learning rate of 0.001 was used for training VGG 16 parameters and 0.0001 for the remaining parameters for 50 k iterations. One image was randomly sampled per batch for training. Each image was resized to 600 and 1400 for shorter and longer sides, respectively. Figure 3 shows the results of vehicles detected in an image using faster-RCNN.

3.2 License plate localization

After successful vehicle detection, the next step is to locate the LP by using morphological operations in the HSI color space. This color space is known to be closely

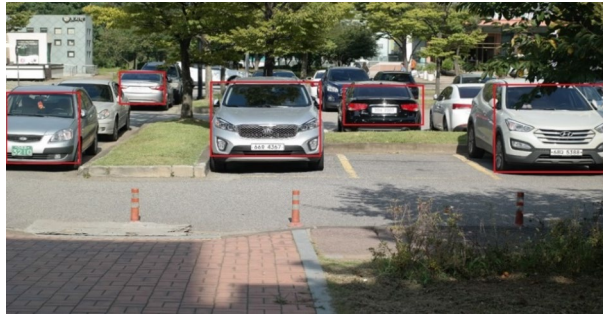


Fig. 3 Test image with vehicle detection

related to the color visualization of human beings [51]. The vehicle area is converted to HSI color space, which separates the color information from the intensity [51]. The current approach uses hue information to determine the colored background plates by defining specific criteria, as our main aim is to find yellow–green and orange plates. Based on our experiment, the following criteria proved sufficient for our requirement.

If Hue is between ($H > 40^\circ$ & $H < 60^\circ$)

Background Color is orange

Else if ($H > 80^\circ$ & $H < 120^\circ$)

Background Color is green

Else if ($H > 60^\circ$ & $H < 80^\circ$)

Background Color is yellow

And if ($H > 180^\circ$ & $H < 240^\circ$)

Background Color is Blue

For readers' information, we state here that there exist several other color spaces, which are based on the tristimulus values, such as CMY, HSI, or YCbCr [51]. Since, we are using the HSI color space in this stage, therefore, we empirically observe that the equations that define HSI are formed in a way that rotates the entire nominal RGB color cube and scales it to fit within a larger color cube. Although these conversions require high computations to correctly determine and interpret HSI signals. However, in our case, we observe that the use of hue to decide background color does not overfit to the investigated data. In addition, white background plates and monochrome images are located using intensity information of HSI color space. Figure 4a shows the binary image results intensity channel and Fig. 4b shows the segmentation result of the Hue channel.

After segmentation and binarization, the candidate's area contains regions of connected pixels. These connected components are labeled using the '4-connectivity labeling' method so that each pixel in the connected region has the same label. Edge detection methods have matrix multiplications that increase computational cost. Therefore, morphological operations have been used instead of edge detection



to reduce computational complexity. After binarization, most of the large connected groups of background are eliminated. However, there is still undesirable information in an image, which can affect the accuracy of plate detection as can be seen from Fig. 4. To minimize these areas and to enhance the pixel area of the plate, a morphological operations-based filtering approach consisting of open and close operations is designed. Morphological open and close operations are shown in Eqs. (1) and (2), respectively.

$$A \cdot SE = (A \ominus SE) \oplus SE, \quad (1)$$

$$A \bullet SE = (A \oplus SE) \ominus SE. \quad (2)$$

In multiple plate detection, the size of LPs depends upon the distance of a car from the camera.

Thus, having more than one SE for one task can increase the computational load. After testing and verifying on several test images, an optimum SE was selected. Figure 5a, b shows the effect of morphological operations on both binary images. Since



Fig. 6 Detection results of the proposed method



Fig. 7 Character segmentation

most of the non-LP regions are removed. Finally, we apply two geometric conditions of area and aspect ratio to locate the license plate. In multiple license plate detection, plate size will vary depending on the distance from the camera. Therefore, having multiple area and aspect ratio values is not an optimum solution. The experiment on a large number of test images was performed to find optimum values. Therefore, area values between 1500 and 4000 pixels and an aspect ratio of 0.2–0.6 are used in the proposed method for plate localization. A similar process is carried on the remaining detected vehicles. Figure 6 shows the overall result of detecting multiple license plates by the proposed methods.

3.3 Character segmentation

The recognition of localized LP now proceeds to the segmentation step. This is a crucial step as recognition totally depends on how well the characters are separated from one another. In this study, pixel connectivity in binary images is used for segmenting characters [52].

First, the LP regions are converted to binary values by using Otsu's threshold method [35]. Next, a morphological thinning is performed to reduce the joining between the LP boundaries and text and in between characters that can negatively impact the process. The connected components are labeled based on pixel connectivity. The labeled pixels are considered and pixels having the same area and aspect ratio are detected as characters. Figure 7 shows segmented characters of some number plates with varying light conditions, different backgrounds, and multiple sizes.

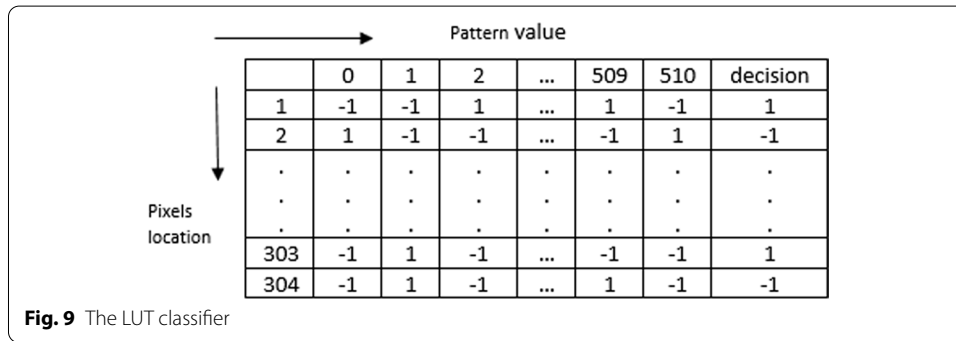
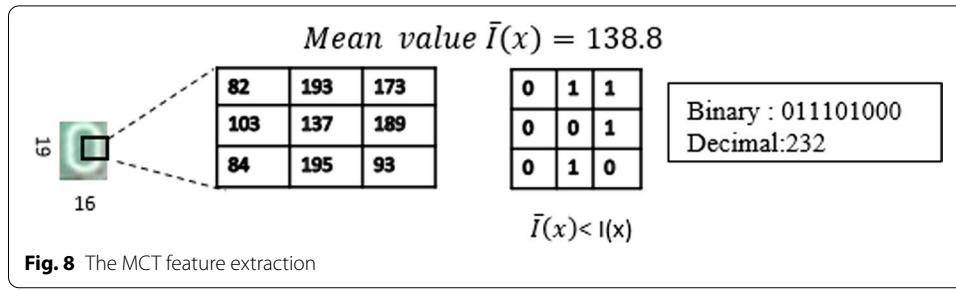
3.4 Character recognition

In multiple plate detection scenarios, the size of plates will vary depending on the distance from the camera. Therefore, the characters isolated during the segmentation step will also have variable sizes. Therefore, techniques, such as template matching do not perform well due to their requirement of fixed size. The resolution of characters plays a crucial role in the identification of characters. Moreover, conventional approaches do not perform well in challenging environments, and various illumination conditions, i.e., rainy time, dusk time, cloudy, and underground parking images.

For character recognition, the AdaBoost with modified census transform (MCT) [13] as a feature extractor is used with a lookup table (LUT) classifier. The LUT is efficient in multi-Gaussian samples classification, whose sensitivity to a fixed number of bins is suitable for the character recognition process. Table 1 shows the algorithm for character recognition. Texture-based analysis plays a vital role in vision-based applications, focusing mainly on how to derive texture features by taking advantage of neighborhood properties. Local binary pattern (LBP) computes a local representation of texture by comparison of center pixel to its neighboring pixels in a defined mask. However, the LBP features have shown poor results when the center pixel value is changed due to varying illumination conditions. Therefore, in our proposed method, MCT features are used and provided excellent results for texture description in the character recognition process in changing light conditions. Figure 8 shows a calculation of the MCT feature with a 3×3 window from a segmented character. MCT features first compute the mean intensity value of the 3×3 window around that specific pixel. For each pixel in the window, MCT assigns “1” if the current pixel value is higher than the mean value and it assigns “0” otherwise. This binary value is converted to decimal to obtain the feature value. This

Table 1 Character recognition with Adaboost algorithm

<ul style="list-style-type: none"> • $\Gamma(x) = \bigotimes_{q \in N'(x)} C(\bar{I}(x), I(x))$ MCT feature at pixel x, ($N'(x)$ neighboring pixels, $\bar{I}(x)$ average of neighboring pixels and C checks if $(\bar{I}(x) < I(x))$ and returns 1 if true and otherwise 0. • Boosting Round: $i = 1, \dots, T$ • Choose a weak classifier $h_t(x, \Gamma(x))$ at pixel x, which is chosen using $x = \operatorname{argmin}_p \varepsilon_t(x)$. Where $\varepsilon_t(x) = \sum_{i h_t(x, \Gamma(x)) \neq C_i} D_t(i)$, C_i represents class if sample i (positive or negative) $D_t(i)$ is weight of sample i at each iteration of t like $D_t(i) \sum_i D_t(i) = 1$ • Weight computation from error $\alpha_t = \frac{1}{2} \ln \left(\frac{1 - \varepsilon_t(p)}{\varepsilon_t(p)} \right)$ Update the weights $D_{t+1}(i) = \frac{D_t(i) e^{-\alpha_t C_i h_t(x, \Gamma(x))}}{\sum_i D_t(i)}$ • Local index LUT are generated $L^{pos}(x, k) = \sum_{i \in S_{pos}} D_t(i) I(\Gamma(x) = k)$ • Final strong classifier is combination of weak classifiers $H(\Gamma) = \sum_{t=1}^T \alpha_t h_t(x, \Gamma(x))$



integer value represents an independent local pattern. Therefore, a 3×3 kernel can have a total of 511 feature values.

Next, the LUT classifier is used for the classification of the MCT features at every pixel location of the character to produce 511 bin feature indices. 511 bin histogram is created $\Gamma(x)$ for all samples in the training set. LUT assigns +1 if positive samples are greater and -1 otherwise, as shown in (3):

$$h(x, k) = \begin{cases} +1 & \text{if } L^{\text{pos}}(x, k) > L^{\text{neg}}(x, k) \\ -1 & \text{otherwise} \end{cases}. \quad (3)$$

Figure 9 shows an example of the LUT classifier, where rows represent the pattern value and columns represent the weak classifier candidates. AdaBoost is an iterative method that sequentially selects a weak classifier pixel location with minimum weighted error in every iteration of learning. Finally, a strong classifier is constructed from the sum of all weak pixel classifiers as shown in (4):

$$H(\Gamma) = \sum_{t=1}^T \alpha_t h_t(x, \Gamma(x)). \quad (4)$$

As character recognition is a multiclass problem, we use one against all classification techniques to construct $k=50$ classifiers for 50 classes. Each classifier is trained by taking positive examples from one class and negative examples from the remaining classes. The output of the multi-class classifier is activated for class having maximum output among all binary classifiers. For outputs of multiple binary classifiers, a multi-classifier generates a vector output S as shown in (5):

$$t = \arg \max_{i=1 \dots k} (H(\Gamma)_i), \quad (5)$$

$$S_i = \begin{cases} +1 & \text{if } i = t, i = 1 \dots k \\ -1 & \text{otherwise} \end{cases}.$$

4 Experimental results and discussion

This section provides experimental results on the published dataset as well as private multiple license plate datasets to describe the effectiveness of the proposed method.

a) Benchmarks

Peking University dataset PKU This dataset was presented and collected by Yuan et al. [53]. It comprises 5 groups. Group 1 to group 4 contain images with a single license plate of cars/trucks on highways and roads in varying illumination conditions. However, group 5 contains 1152 images with multiple vehicles under diverse environmental conditions comprising low illumination, low resolution (size), and varying contrast.

Application-oriented license plate dataset (AOLP) [54] Contains a total of 2049 Taiwanese license plates images. Based on diverse applications this dataset is divided into three categories: AC (access control, 681 images), LE (law enforcement, 757 images), and RP (road patrol, 611 images). Specifically, AC contains images passing taken by a stationary camera moving at low speed, LE contains images of cars taken by the roadside camera moving at variable speeds, RP is the most challenging part of this dataset that contains images captured by law enforcement vehicles.

Media Lab dataset [55] Contains 706 images of Greek license plates under-constrained as well as the unconstrained environment. This dataset is divided into different groups based on various conditions such as blurred, color, and grayscale images, LPs with a close view, and shadows and images having more than one vehicle.

Xu et al. presented the CCPD [56] dataset, which is currently the largest Chinese vehicle dataset available, Table 2 describes the details of the dataset.

As discussed, the Subset AC and LE of the AOLP dataset are relatively less challenging than the RP subset as it contains images with blurriness and distortion introduced due to the motion of the camera. As can be seen from Table 3 that our method outperformed [57] and robust attention [58] in all three subsets of this dataset. However, Table 2 represents that we achieved comparable accuracy to the [59] which has shown the best possible accuracy to date. Thus proposed method performance on the AOLP dataset represents that it is effective in challenging conditions of distortion, blurriness, and rotation.

As already mentioned, in the PKU dataset the most relevant subgroup to our paper application is G5, which contains multiple license plate images in challenging conditions. As can be seen from Table 4, the proposed method achieves comparable accuracy

Table 2 Description of the CCPD dataset

Category	Conditions	Input image resolution (pixels)	No. of images
Base	Vehicles on roads with one license plate per image	720 × 1160	199,996
DB	Images of vehicle in extreme bright and low light conditions	720 × 1160	10,132
FN	Vehicle images were taken from far as well as closeup to cover various LP sizes	720 × 1160	20,976
Rotate	Vehicles images with angular views	720 × 1160	10,053
Blur	Blurriness is introduced due to the movement of the camera	720 × 1160	20,611
Weather	Vehicles images taken under diverse weather conditions	720 × 1160	9999
Challenge	Most challenging images include unconstrained illuminations and weather conditions along with images having multiple image processing problems (blurriness, size of LP, glare, etc.)	720 × 1160	50,003
Tilt	Vehicles images with angular views	720 × 1160	30,216

Table 3 Comparison on AOLP dataset

Method	AC	LE	RP
[57]	99.41	99.31	99.02
[58]	97.3	98.3	91.9
[60]	–	–	98.36
[59]	99.56	100	99.84
Ours	99.37	99.47	99.35

Table 4 Detection accuracy on the PKU dataset

Ref	G1	G2	G3	G4	G5
[61]	98.89	98.42	95.83	81.17	83.31
[53]	98.76	98.42	97.72	96.23	97.32
[62]	99.75	99.86	99.64	96.65	99.47
Ours	99.78	99.31	99.50	99.58	99.67

to all state of art methods in groups G1–G4 of the dataset. However, as the proposed algorithm outperforms all methods on the G5 subgroup which has huge vehicle diversity and plates with multiple orientations, etc., this further accentuates the superiority of the proposed scheme for multiple plate detection.

Table 5 depicts that performance on the CCPD dataset in terms of group-wise accuracy of the proposed method is comparatively better than other states of art algorithms. Luo et al. [63] is the only method performing better on groups of tilt and rotate as this method was explicitly presented to solve the tilted/rotated plates. The other state of art methods compared included [64] a text recognition approach and multi cascaded CNN [65] based approach and attention-based method which utilizes Xception CNN for feature extraction and recurrent neural network for decoding. As can be seen from the results proposed method accuracy is better for groups containing unconstrained environments like weather and challenge as state of art methods were unable to detect plates in extreme reflective glare and weather conditions.

Table 5 The LP detection accuracy (%) on each CCPD subset

Method	Base	Rotate	Tilt	FN	DB	Weather	Challenge
[65]	99.1	96.3	97.3	95.1	96.4	97.1	83.2
[63]	99.5	98.8	98.8	98.1	98.6	97.6	86.5
[64]	98.9	96.1	96.4	91.9	93.7	95.4	83.1
[58]	99.6	98.1	98.2	96.4	97.6	98.5	88.9
[56]	95.5	98.5	96.9	94.3	90.8	92.5	87.9
Ours	99.65	98.7	98.6	98.67	98.71	98.49	92.43

In the media lab dataset, the number of images is very less to effectively train the model, therefore we decided to choose the training model used by [59], which performs fourfold cross-validation. This model divides the subsets into four equal random parts followed by using three subsets for training and the fourth one for testing. It is evident from Table 6 that the proposed method outperformed the existing methods as this dataset contains relatively less challenging illumination and weather conditions compared to other datasets and contains good resolution images as well.

The private dataset contains a total of 4179 (resolution 1920×1080) images were taken, using 2000 for testing and 2179 for training purposes, in varying illumination conditions and environments, i.e., (night, day, dusk, cloudy weather, rainy weather, and parking).

Table 7 compares the results in terms of recall and precision ratio of detection method with existing methods when applied to images with multiple license plates. Figure 10 shows the results for challenging illuminating and weather conditions throughout the day.

Table 6 Detection accuracy on Media Lab dataset

Method accuracy	Accuracy
[55]	86.00
[66]	87.33
[57]	97.89
[59]	98.83
Proposed	99.19

Table 7 Comparison with existing methods

Method	Precision ratio	Recall ratio
[14]	80.17%	84.89%
[67]	85.45%	89.95%
[68]	85.45%	89.95%
[69]	86%	90.56%
[57]	87.05%	91.45%
[28]	89.19%	93.31%
[32]	90.47%	92.31%
Faster RCNN	92.87%	94.97%
[33]	93.87%	95.73%
Proposed	94.57%	96.72%



Fig. 10 Results of images taken: **a** sunny day, **b** night time, **c** dusk, **d** cloudy, **e** underground parking, and **f** rainy

The proposed method outperforms conventional methods in terms of both precision and recall. There were 5543 vehicles in 2000 images used for the testing process. The proposed method detected 5361 LPs correctly with an accuracy (recall rate) of 96.72%. Recall and precision are defined as below:

$$\text{Recall} = \frac{\text{No. of correctly detected LPs}}{\text{Total no of vehicles}} \quad (6)$$

$$\text{Precision} = \frac{\text{No. of correctly detected LPs}}{\text{Total detected (including False positives)}} \quad (7)$$

Figure 10 shows the results for challenging illuminating and weather conditions throughout the day.

The recall ratio of the proposed method is 13% higher than of the edge detection method since the edge method was unable to detect color background number plates. Precision is higher when compared with the AdaBoost method, as the AdaBoost method also detects headlights and text as a license plate. Li et al. [24] trained a 37 class CNN system for character detection in images followed by a CNN classifier as a false positive eliminator. This method also produced more false positives in real-world scenarios where images had text other than LP. The Yolo and SSD both underperformed on the private dataset as both were unable to detect plates in these images due to images containing small size license plates (based on distance from the camera). Similarly, the detection results of text detection techniques [32, 33] are less accurate as these methods performance deteriorates due to a combination of low resolution (size of plates) with other environmental factors such as varying illumination, angular plates, weather conditions, etc.

Character recognition performance is evaluated in Table 3. The proposed recognition method was tested on all the plates successfully detected. We compared the performance of the presented method with popular methods. First, a scale-invariant feature transform is used for feature extraction and a support vector machine is used for classification. The second method is a 3-layer multilayer neural network for character recognition. The third method is a traditional convolutional neural network having two convolution layers, 2 fully connected layers followed by SVM for classification.

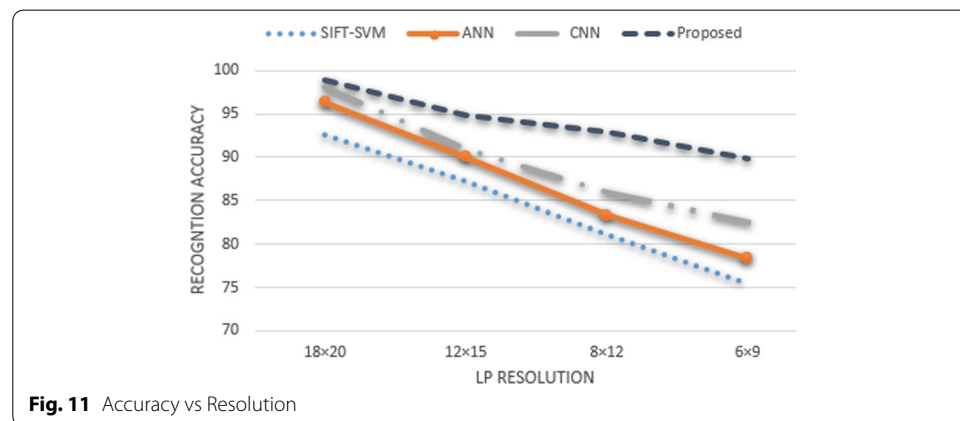
The SIFT and the SVM-based method was unable to classify characters due to partial occlusion in rainy images, the effect of vehicle headlights in basement images, and exposure to strong sunlight. ANN has the worst results as broken characters and two font characters on plates were unrecognizable by this method. As CNN can automatically learn features, it has performed better than both existing methods. However, its performance degraded for low-resolution (based on the distance of camera) characters. When compared to state-of-the-art methods in scene text recognition that included methods using RNN/transformers [44] and semantic reasoning network (SRN) [46]. The proposed method outperformed these methods as they were unable to perform well in an uncontained environment that includes occlusions (raindrops), very low resolution, characters with extreme reflective glare, and little semantic meanings. Moreover, the proposed method outperformed these methods in terms of accuracy in challenging conditions, such as varying illumination images as per the

Table 8 Character recognition

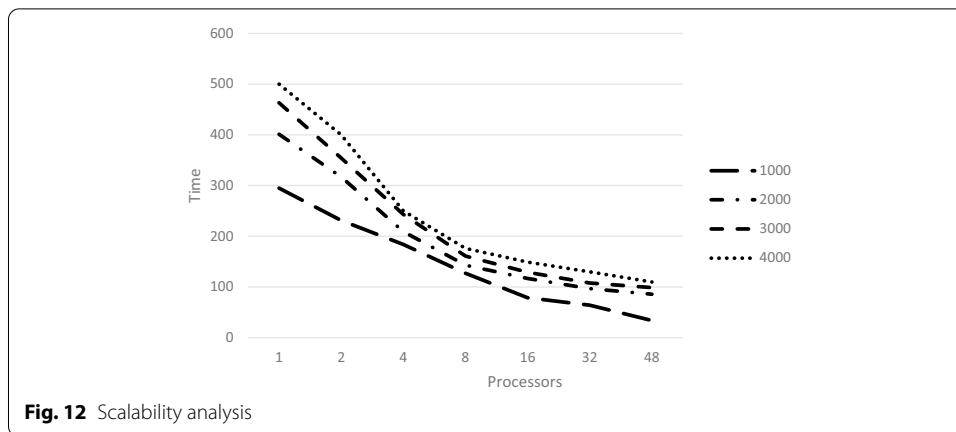
Method	Precision ratio	Recall ratio
SIFT-SVM	84.13%	87.97%
ANN	87.36%	91.32%
CNN	90.05%	94.67%
[44]	91.33%	95.89%
[46]	92.35%	96.98%
Proposed	94.57%	98.02%

results shown in Table 8. Figure 11 shows the comparison of character recognition performance of conventional approaches with the proposed method for low-resolution characters and demonstrates the superiority of the proposed scheme when compared with the benchmarks.

Table 9 lists the overall (detection + recognition) performance results obtained under different lighting conditions and weather conditions. Our dataset consists of images taken during night time, daytime, dusk, and cloudy weather. Images taken during cloudy and sunny days produce a better result due to consistent light conditions, except for cases when LPs are affected by the reflection of sunlight from surroundings. Images of a car parked in the basement also produced good results with exception to a reflection of other cars' headlights. Worst results were produced during dusk time owing to the quickly varying illumination during this time of day.

**Fig. 11** Accuracy vs Resolution**Table 9** Performance comparison for varying conditions

Weather	# of images	Total vehicles	Detection rate	Character recognition
Sunny	450	1193	97.98%	98.97%
Night	295	839	97.73%	98.31%
Dusk	397	1120	94.12%	96.43%
Cloudy	438	1175	97.87%	98.45%
Parking	205	589	96.43%	98.01%
Rainy	215	627	95.53%	97.91%
Total	2000	5543	96.72%	98.02%

**Table 10** Performance comparison of state-of-the-art detectors on decreasing LP size

Detectors	Resolution	DB	FN	Rotate	Tilt	DB	Weather	Challenge
Yolo	50 × 30	98.32	98.11	97.76	97.67	97.89	91.78	88.23
	35 × 20	97.11	96.89	96.13	96.03	96.54	90.25	86.12
	25 × 15	93.19	93.67	93.5	92.39	92.19	88.72	82.01
SSD	50 × 30	96.23	96.24	96.32	96.12	95.01	90.89	87.26
	35 × 20	95.32	95.14	95.15	95.21	94.76	90.16	86.23
	25 × 15	91.14	91.24	90.88	89.33	89.43	87.43	81.12
Faster RCNN	50 × 30	99.19	99.11	98.99	98.78	98.34	96.45	95.44
	35 × 20	98.31	97.11	96.87	96.57	96.18	94.34	91.08
	25 × 15	94.73	94.11	93.75	93.11	91.02	89.23	87.72

Moreover, dimming sunlight, street lights, and cars' headlights have a negative impact on the overall performance of the method, especially when vehicles are at a far distance from the camera. Results of images taken in rainy conditions are also encouraging. However, some characters were not recognized owing to images getting blurred due to water pouring down the windscreens of the cars containing the camera.

In terms of scalability, the proposed algorithm performance should be consistent and must not deteriorate drastically. To evaluate scalability, we perform analysis on execution time by increasing the number of plate images and processors, respectively. We observed that with a gradual increase in plate images from 1000 to 4000 images processing time increases, however, this increase is mitigated by increasing the number of processors as shown in Fig. 12.

4.1 Ablation study

A detailed analysis to demonstrate the effectiveness of the proposed method is presented in this section. A License plate occupying a small portion of the image loses critical information after several down-sampling stages applied by CNN based object detector. Hence, these methods were not able to achieve better accuracy. Table 10 shows the comparison of state of art object detectors on various LP sizes. It is evident from the table that with a decrease in plate size and an increase in challenging

Table 11 Effect of change in scale and aspect ratio on accuracy

Anchor scales	Aspect ratios	Accuracy
$128 \times 128, 256 \times 256$	1:1, 1:2	92.03
$128 \times 128, 256 \times 256$	2:1, 1:1, 1:2	93.45
$128 \times 128, 256 \times 256, 512 \times 512$	1:1, 1:2	96.05
$128 \times 128, 256 \times 256, 512 \times 512$	2:1, 1:1, 1:2	99.03

Table 12 Processing time

Processing steps	Average time (ms)
Vehicle detection	463
Plate detection	107
Plate recognition	91

conditions the accuracy levels of object detectors also reduce significantly. Thus based on these results and multiple license plate recognition involves small size plate and challenging conditions proposed method used Faster-RCNN was used as object detector and then license plate region is localized using Image processing techniques. To further improve the overall accuracy selection of an optimum number of scales and anchor boxes play a crucial role. Therefore, we conducted multiple experiments by changing scales and aspect ratios to evaluate the performance of Faster-RCNN on our dataset. Table 11 compares the results of different combinations of scales and aspect ratios. We can conclude that the default setting of the 3 anchors scale and 3 aspect ratio produce the best results in vehicle detection.

Furthermore, the same color of LP with different values of color components in the day and night hours makes it a difficult condition to handle particularly in extreme cases. However, the proposed algorithm resulted in the efficient handling of this difficulty. In this application, the HSI color model has been applied, in which any component of the color can be altered separately without disturbing others. This feature of the proposed model is very effective in dealing with adverse conditions such as extreme light conditions. This characteristic of the current model resulted in achieving high precision even in challenging conditions.

As it is obvious that this paper is targeting multiple license plate recognition, therefore computational performance will be not as much as efficient for images with a single license plate Table 12 shows the time consumption of each portion of the algorithm. The proposed approach requires only 570 ms to detect a license plate, which includes vehicle detection and plate detection in an image with a resolution of 1920×1080 pixels. Moreover, it almost takes 91 ms to recognize the characters of the license plate. It was observed that our two-step approach does not outperform other methods in terms of computational performance, however, it is sufficient enough to achieve the real-time processing speed required for ITS applications.

5 Conclusions and future work

In this paper, a multiple license plate recognition method, for high-resolution images, was presented, which works in challenging illumination conditions in real-time scenarios. The proposed technique divided plate detection into two steps. In the first step, faster-RCNN was used to detect all the vehicles in an image resulting in scaled information to locate plates. Meanwhile, morphological operations were used to reduce non-plate regions and geometric properties were used to localize plate HSI color space. Then, character recognition is executed by a LUT classifier using adaptive boosting with MCT as a feature extractor. Experimental results showed that the detection rate of the proposed method is much higher than existing methods, with an overall detection rate of 96.72% and a recognition rate of 98.02% in multiple LPs and varying illumination scenarios. The proposed algorithm might be suitable for real-time ITS applications [32].

Future work could focus on developing a parallel version of the developed algorithm. We believe that will further improve the execution time to recognize a license plate.

Abbreviations

ITS: Intelligent transport systems; RCNN: Region-based convolutional neural Network; AdaBoost: Adaptive boosting; MCT: Modified census transform; LUT: Lookup table; LDA: Linear discriminant analysis.

Acknowledgements

Not applicable.

Authors' contributions

All authors have equivalent contributions. All authors read and approved the final manuscript.

Author's Information

Khurram Khan received his Ph.D. degree in Electronic & Communication Engineering from Hanyang University South Korea. His research interests are image recognition, Computer vision, and Machine/deep learning. Abid Imran received his Ph.D. degree in Electronics & systems engineering from Hanyang University South Korea. His research interests are Bio-Inspired Manipulators, Robot Manipulation, and machine vision. Hafiz Zia ur Rehman received his Ph.D. degree in Mechatronics engineering from Hanyang University South Korea. His research interests are Medical Image processing and Machine /deep learning. Adnan Fazil received his Ph.D. degree in Electrical engineering from Air University. His research interests are image Communication, Wireless Communication. Muhammad Zakwan received his Ph.D. degree in Electronics engineering from KoC University Turkey. His research interests are optics and Photonics. Zahid Mehmood received his Ph.D. degree in Electronics Engineering from North Dakota University, USA. His research interests are Object Detection, Pattern Recognition, Image Enhancement, and Digital Image/Video Processing.

Funding

No funding.

Availability of data and materials

Some data can be shared on request.

Declarations

Competing interests

The authors declare that they have no conflict of interest.

Author details

¹Department of Avionics Engineering, Air University, Islamabad, Pakistan. ²Faculty of Mechanical Engineering, Ghulam Ishaq Khan Institute of Engineering and Technology, Topi, Pakistan. ³Department of Mechatronics Engineering, Air University, Islamabad, Pakistan. ⁴Department of Electrical and Computer Engineering, Comsats University, Abbottabad, Pakistan.

Received: 25 April 2021 Accepted: 2 September 2021

Published online: 17 September 2021

References

1. C.-N.E. Anagnostopoulos, License plate recognition: a brief tutorial. *IEEE Intell. Transp. Syst. Mag.* **6**(1), 59–67 (2014). <https://doi.org/10.1109/MITS.2013.2292652>
2. M.R. Asif, Q. Chun, S. Hussain, M.S. Fareed, Multiple licence plate detection for Chinese vehicles in dense traffic scenarios. *IET Intell. Transp. Syst.* **10**(8), 535–544 (2016). <https://doi.org/10.1049/iet-its.2016.0008>
3. Shen-Zheng Wang and Hsi-Jian Lee, "Detection and recognition of license plate characters with different appearances," in *Proceedings of the 2003 IEEE International Conference on Intelligent Transportation Systems*, (2003), vol. 2, pp. 979–984. <https://doi.org/10.1109/ITSC.2003.1252632>
4. M.A. Massoud, M. Sabee, M. Gergais, R. Bakhit, Automated new license plate recognition in Egypt. *Alex. Eng. J.* **52**(3), 319–326 (2013). <https://doi.org/10.1016/j.aej.2013.02.005>
5. T. Ibrahim, K. Kirami, License plate recognition system using artificial neural networks. *ETRI J.* **39**(2), 163–172 (2017). <https://doi.org/10.4218/etrij.17.0115.0766>
6. Md. Mostafa Kamal Sarke, Sook Yoon, and Dong Sun Park, A fast and robust license plate detection algorithm based on two-stage cascade AdaBoost, *8*(10): 3490–3507. <https://doi.org/10.3837/tiis.2014.10.012>
7. W.T. Ho, H.W. Lim, and Y.H. Tay, "Two-stage license plate detection using gentle adaboost and SIFT-SVM," in *2009 First Asian Conference on Intelligent Information and Database Systems*, (2009), pp. 109–114. <https://doi.org/10.1109/ACIDS.2009.25>
8. C. Yan, B. Shao, H. Zhao, R. Ning, Y. Zhang, F. Xu, 3D room layout estimation from a single RGB image. *IEEE Trans. Multimed.* **22**(11), 3014–3024 (2020). <https://doi.org/10.1109/TMM.2020.2967645>
9. "A Single Neural Network for Mixed Style License Plate Detection and Recognition | IEEE Journals & Magazine | IEEE Xplore," <https://ieeexplore.ieee.org/document/9337806> (accessed Jul. 16, 2021)
10. S. A. Radzi and M. Khalil-Hani, "Character recognition of license plate number using convolutional neural network," in *Visual Informatics: Sustaining Research and Innovations*, Berlin, Heidelberg, (2011), pp. 45–55. https://doi.org/10.1007/978-3-642-25191-7_6
11. C. Gerber, M. Chung, Number plate detection with a multi-convolutional neural network approach with optical character recognition for mobile devices. *J. Inf. Process. Syst.* **12**, 100–108 (2016). <https://doi.org/10.3745/JIPS.04.0022>
12. K. Khan, M.-R. Choi, Automatic license plate detection and recognition framework to enhance security applications. *J. Electron. Imaging* **28**(1), 013036 (2019). <https://doi.org/10.1117/1.JEI.28.1.013036>
13. S.-K. Park, D.-G. Sim, New MCT-based face recognition under varying lighting conditions. *Int. J. Control Autom. Syst.* **9**(3), 542–549 (2011). <https://doi.org/10.1007/s12555-011-0314-0>
14. A.M. Al-Ghaili, S. Mashohor, A.R. Ramli, A. Ismail, Vertical-edge-based car-license-plate detection method. *IEEE Trans. Veh. Technol.* **62**(1), 26–38 (2013). <https://doi.org/10.1109/TVT.2012.2222454>
15. J. Yopez, S.-B. Ko, Improved license plate localisation algorithm based on morphological operations. *IET Intell. Transp. Syst.* **12**(6), 542–549 (2018). <https://doi.org/10.1049/iet-its.2017.0224>
16. Hsi-Jian Lee, Si-Yuan Chen, and Shen-Zheng Wang, "Extraction and recognition of license plates of motorcycles and vehicles on highways," in *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, (2004), vol. 4, pp. 356–359. <https://doi.org/10.1109/ICPR.2004.1333776>
17. C.-N.E. Anagnostopoulos, I.E. Anagnostopoulos, I.D. Psoroulas, V. Loumos, E. Kayafas, License plate recognition from still images and video sequences: a survey. *IEEE Trans. Intell. Transp. Syst.* **9**(3), 377–391 (2008). <https://doi.org/10.1109/TITS.2008.922938>
18. B.K. Cho, S.H. Ryu, D.R. Shin, J.I. Jung, License plate extraction method for identification of vehicle violations at a railway level crossing. *Int. J. Automot. Technol.* **12**(2), 281–289 (2011). <https://doi.org/10.1007/s12239-011-0033-9>
19. M.U. Haq, A. Shahzad, Z. Mahmood, A.A. Shah, Boosting the face recognition performance of ensemble based LDA for pose, non-uniform illuminations, and low-resolution images. *KSII Trans. Internet Inf. Syst.* **13**(6), 3144–3164 (2019)
20. Sang Kyoon Kim, Dae Wook Kim, and Hang Joon Kim, "A recognition of vehicle license plate using a genetic algorithm based segmentation," in *Proceedings of 3rd IEEE International Conference on Image Processing*, (1996), vol. 2, pp. 661–664. <https://doi.org/10.1109/ICIP.1996.560964>
21. M.A. Khan, M. Sharif, M.Y. Javed, T. Akram, M. Yasmin, T. Saba, License number plate recognition system using entropy-based features selection approach with SVM. *IET Image Process.* **12**(2), 200–209 (2018). <https://doi.org/10.1049/iet-ipr.2017.0368>
22. Z. Mahmood, T. Ali, N. Muhammad, N. Bibi, I. Shahzad, S. Azmat, EAR: enhanced augmented reality system for sports entertainment applications. *KSII Trans. Internet Inf. Syst.* **11**(12), 6069–6091 (2017)
23. J. Shashirangana, H. Padmasiri, D. Meedeniya, C. Perera, Automated license plate recognition: a survey on methods and techniques. *IEEE Access* **9**, 11203–11225 (2021). <https://doi.org/10.1109/ACCESS.2020.3047929>
24. H. Li and C. Shen, "Reading car license plates using deep convolutional neural networks and LSTMs," *ArXiv160105610 Cs*, (2016), Accessed: Dec. 10, 2019. [Online]. <http://arxiv.org/abs/1601.05610>
25. H. Xiang, Y. Yuan, Y. Zhao, Z. Fu, License plate detection based on fully convolutional networks. *J. Electron. Imaging* **26**(5), 053027 (2017). <https://doi.org/10.1117/1.JEI.26.5.053027>
26. M.A. Rafique, W. Pedrycz, M. Jeon, Vehicle license plate detection using region-based convolutional neural networks. *Soft Comput.* (2018). <https://doi.org/10.1007/s00500-017-2696-2>
27. S. Ren, K. He, R. Girshick, J. Sun, Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(6), 1137–1149 (2017). <https://doi.org/10.1109/TPAMI.2016.2577031>
28. L. Xie, T. Ahmad, L. Jin, Y. Liu, S. Zhang, A new CNN-based method for multi-directional car license plate detection. *IEEE Trans. Intell. Transp. Syst.* **19**(2), 507–517 (2018). <https://doi.org/10.1109/TITS.2017.2784093>
29. H. Li, P. Wang, C. Shen, Toward end-to-end car license plate detection and recognition with deep neural networks. *IEEE Trans. Intell. Transp. Syst.* **20**(3), 1126–1136 (2019). <https://doi.org/10.1109/TITS.2018.2847291>
30. C. Yan, B. Gong, Y. Wei, Y. Gao, Deep multi-view enhancement hashing for image retrieval. *IEEE Trans. Pattern Anal. Mach. Intell.* **43**(4), 1445–1451 (2021). <https://doi.org/10.1109/TPAMI.2020.2975798>
31. C. Yan, Z. Li, Y. Zhang, Y. Liu, X. Ji, Y. Zhang, Depth image denoising using nuclear norm and learning graph model. *ACM Trans. Multimed. Comput. Commun. Appl.* **16**(4), 122:1–122:17 (2020). <https://doi.org/10.1145/3404374>

32. C. Xue, S. Lu, and W. Zhang, "MSR: multi-scale shape regression for scene text detection," (2019). <https://doi.org/10.24963/ijcai.2019/139>
33. M. Liao, Z. Wan, C. Yao, K. Chen, X. Bai, Real-time scene text detection with differentiable binarization. *Proc. AAAI Conf. Artif. Intell.* **34**(07), 11474–11481 (2020). <https://doi.org/10.1609/aaai.v34i07.6812>
34. K. K. Kim, K. I. Kim, J. B. Kim, and H. J. Kim, "Learning-based approach for license plate recognition," in *Neural Networks for Signal Processing X. Proceedings of the 2000 IEEE Signal Processing Society Workshop (Cat. No.00TH8501)*, (2000), vol. 2, pp. 614–623. <https://doi.org/10.1109/NNSP.2000.890140>
35. N. Otsu, A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **9**(1), 62–66 (1979). <https://doi.org/10.1109/TSMC.1979.4310076>
36. S. Nomura, K. Yamanaka, O. Katai, H. Kawakami, T. Shiose, A novel adaptive morphological approach for degraded character image segmentation. *Pattern Recognit.* **38**(11), 1961–1975 (2005). <https://doi.org/10.1016/j.patcog.2005.01.026>
37. Y. Zhang et al., Multi-kernel extreme learning machine for EEG classification in brain-computer interfaces. *Expert Syst. Appl.* **96**, 302–310 (2018). <https://doi.org/10.1016/j.eswa.2017.12.015>
38. J. Tarigan, Nadia, R. Diedan, Y. Suryana, Plate recognition using backpropagation neural network and genetic algorithm. *Procedia Comput. Sci.* **116**, 365–372 (2017). <https://doi.org/10.1016/j.procs.2017.10.068>
39. Y. Wen, Y. Lu, J. Yan, Z. Zhou, K.M. von Deneen, P. Shi, An algorithm for license plate recognition applied to intelligent transportation system. *IEEE Trans. Intell. Transp. Syst.* **12**(3), 830–845 (2011). <https://doi.org/10.1109/TITS.2011.2114346>
40. P. Shivakumara, D. Tang, M. Asadzadehkaljahi, T. Lu, U. Pal, M. Hossein Anisi, CNN-RNN based method for license plate recognition. *CAAI Trans. Intell. Technol.* **3**(3), 169–175 (2018). <https://doi.org/10.1049/trit.2018.1015>
41. O. Bulan, V. Kozitsky, P. Ramesh, M. Shreve, Segmentation- and annotation-free license plate recognition with deep localization and failure identification. *IEEE Trans. Intell. Transp. Syst.* **18**(9), 2351–2363 (2017). <https://doi.org/10.1109/TITS.2016.2639020>
42. S. Du, M. Ibrahim, M. Shehata, W. Badawy, Automatic License Plate Recognition (ALPR): a state-of-the-art review. *IEEE Trans. Circuits Syst. Video Technol.* **23**(2), 311–325 (2013). <https://doi.org/10.1109/TCSVT.2012.2203741>
43. Y. Yang, D. Li, Z. Duan, Chinese vehicle license plate recognition using kernel-based extreme learning machine with deep convolutional features. *IET Intell. Transp. Syst.* **12**(3), 213–219 (2018). <https://doi.org/10.1049/iet-its.2017.0136>
44. C. Xue, S. Lu, S. Bai, W. Zhang, and C. Wang, "I2C2W: image-to-character-to-word transformers for accurate scene text recognition," *ArXiv210508383 Cs*, (2021), Accessed: Jul. 15, 2021. [Online]. <http://arxiv.org/abs/2105.08383>
45. D. Yu et al., "Towards accurate scene text recognition with semantic reasoning networks," 2020, pp. 12113–12122. Accessed: Jul. 15, 2021. [Online]. https://openaccess.thecvf.com/content_CVPR_2020/html/Yu_Towards_Accurate_Scene_Text_Recognition_With_Semantic_Reasoning_Networks_CVPR_2020_paper.html
46. B. Su and S. Lu, "Accurate scene text recognition based on recurrent neural network," in *Computer Vision—ACCV 2014*, (Cham, 2015), pp. 35–48. https://doi.org/10.1007/978-3-319-16865-4_3
47. "IET Digital Library: towards a fully automated car parking system." <https://digital-library.theiet.org/content/journals/10.1049/iet-its.2018.5021> (accessed Jul. 19, 2021)
48. R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, (Columbus, 2014), pp. 580–587. <https://doi.org/10.1109/CVPR.2014.81>
49. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: unified, real-time object detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2016), pp. 779–788. <https://doi.org/10.1109/CVPR.2016.91>
50. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," presented at the ICLR, (2015)
51. R.C. Gonzalez, R.E. Woods, *Digital Image Processing*, 3rd edn. (Prentice-Hall Inc, Upper Saddle River, 2006)
52. B.-F. Wu, S.-P. Lin, C.-C. Chiu, Extracting characters from real vehicle licence plates out-of-doors. *IET Comput. Vis.* **1**(1), 2–10 (2007). <https://doi.org/10.1049/iet-cvi:20050132>
53. Y. Yuan, W. Zou, Y. Zhao, X. Wang, X. Hu, N. Komodakis, A robust and efficient approach to license plate detection. *IEEE Trans. Image Process.* **26**(3), 1102–1114 (2017). <https://doi.org/10.1109/TIP.2016.2631901>
54. G.-S. Hsu, J.-C. Chen, Y.-Z. Chung, Application-oriented license plate recognition. *IEEE Trans. Veh. Technol.* **62**(2), 552–561 (2013). <https://doi.org/10.1109/TVT.2012.2226218>
55. C.N.E. Anagnostopoulos, I.E. Anagnostopoulos, V. Loumos, E. Kayafas, A license plate-recognition algorithm for intelligent transportation system applications. *IEEE Trans. Intell. Transp. Syst.* **7**(3), 377–392 (2006). <https://doi.org/10.1109/TITS.2006.880641>
56. Z. Xu et al., "Towards end-to-end license plate detection and recognition: a large dataset and baseline," (2018), pp. 255–271. Accessed: Jul. 15, 2021. [Online]. https://openaccess.thecvf.com/content_ECCV_2018/html/Zhenbo_Xu_Towards_End-to-End_License_ECCV_2018_paper.html
57. J. Zhuang, S. Hou, Z. Wang, and Z.-J. Zha, "Towards human-level license plate recognition," in *Computer Vision—ECCV 2018*, (Cham, 2018), pp. 314–329. https://doi.org/10.1007/978-3-030-01219-9_19
58. L. Zhang, P. Wang, H. Li, Z. Li, C. Shen, Y. Zhang, A robust attentional framework for license plate recognition in the wild. *IEEE Trans. Intell. Transp. Syst.* (2020). <https://doi.org/10.1109/TITS.2020.3000072>
59. Y. Zhang, Z. Wang, J. Zhuang, Efficient license plate recognition via holistic position attention. *Proc. AAAI Conf. Artif. Intell.* **35**(4), 3438–3446 (2021)
60. Y. Lee, J. Lee, H. Ahn, and M. Jeon, "SNIDER: single noisy image denoising and rectification for improving license plate recognition," (2019), pp. 0–0. Accessed: Jul. 19, 2021. [Online]. https://openaccess.thecvf.com/content_ICCVW_2019/html/RLQ/Lee_SNIDER_Single_Noisy_Image_Denoising_and_Rectification_for_Improving_License_ICCVW_2019_paper.html
61. B. Li, B. Tian, Y. Li, D. Wen, Component-based license plate detection using conditional random field model. *IEEE Trans. Intell. Transp. Syst.* **14**(4), 1690–1699 (2013). <https://doi.org/10.1109/TITS.2013.2267054>

62. S. M. Silva and C. R. Jung, "License plate detection and recognition in unconstrained scenarios," in *Computer Vision—ECCV 2018*, (Cham, 2018), pp. 593–609. https://doi.org/10.1007/978-3-030-01258-8_36
63. C. Luo, L. Jin, Z. Sun, MORAN: a multi-object rectified attention network for scene text recognition. *Pattern Recognit.* **90**, 109–118 (2019). <https://doi.org/10.1016/j.patcog.2019.01.020>
64. T. Wang et al., Decoupled attention network for text recognition. *Proc. AAAI Conf. Artif. Intell.* **34**(07), 12216–12224 (2020). <https://doi.org/10.1609/aaai.v34i07.6903>
65. S. Zherzdev and A. Gruzdev, "LPRNet: license plate recognition via deep neural networks," *ArXiv180610447 Cs*, (2018), Accessed: Jul. 15, 2021. [Online]. <http://arxiv.org/abs/1806.10447>
66. "End-to-end system of license plate localization and recognition." <https://www.spiedigitallibrary.org/journals/Journal-of-Electronic-Imaging/volume-24/issue-2/023020/End-to-end-system-of-license-plate-localization-and-recognition/10.1117/1.JEI.24.2.023020.short?SSO=1> (accessed Jul. 15, 2021)
67. Z. Mahmood, T. Ali, S. Khattak, L. Hasan, S.U. Khan, Automatic player detection and identification for sports entertainment applications. *Pattern Anal. Appl.* **18**(4), 971–982 (2015). <https://doi.org/10.1007/s10044-014-0416-4>
68. W. Liu et al., "SSD: single shot multibox detector," in *Computer Vision—ECCV 2016*, (Cham, 2016), pp. 21–37. https://doi.org/10.1007/978-3-319-46448-0_2
69. J. Redmon and A. Farhadi, "YOLOv3: an Incremental Improvement," *ArXiv180402767 Cs*, (2018), Accessed: Jul. 15, 2021. [Online]. <http://arxiv.org/abs/1804.02767>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)