CAAI Transactions on Intelligence Technology

Special Issue: Selected Papers from The 4th Asian Conference on Pattern Recognition (ACPR 2017)

Symmetry features for license plate classification

ISSN 2468-2322
Received on 20th July 2018
Accepted on 21st July 2018
doi: 10.1049/trit.2018.1016
www.ietdl.org

Karpuravalli Srinivas Raghunandan[1], Palaiahnakote Shivakumara[2] ✉, Lolika Padmanabhan[3], Govindaraju Hemantha Kumar[1], Tong Lu[4], Umapada Pal[5]

[1]*Department of Studies in Computer Science, University of Mysore, Karnataka, India*
[2]*Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur, Malaysia*
[3]*PES Institute of Technology, Bangalore, Karnataka, India*
[4]*National Key Lab for Novel Software Technology, Nanjing University, Nanjing, People's Republic of China*
[5]*Computer Vision and Pattern Recognition Unit, Indian Statistical Institute, Kolkata, India*
✉ *E-mail: shiva@um.edu.my*

Abstract: Achieving high recognition rate for license plate images is challenging due to multi-type images. We present new symmetry features based on stroke width for classifying each input license image as private, taxi, cursive text, when they expand the symbols by writing and non-text such that an appropriate optical character recognition (OCR) can be chosen for enhancing recognition performance. The proposed method explores gradient vector flow (GVF) for defining symmetry features, namely, GVF opposite direction, stroke width distance, and stroke pixel direction. Stroke pixels in Canny and Sobel which satisfy the above symmetry features are called local candidate stroke pixels. Common stroke pixels of the local candidate stroke pixels are considered as the global candidate stroke pixels. Spatial distribution of stroke pixels in local and global symmetry are explored by generating a weighted proximity matrix to extract statistical features, namely, mean, standard deviation, median and standard deviation with respect the median. The feature matrix is finally fed to an support vector machine (SVM) classifier for classification. Experimental results on large datasets for classification show that the proposed method outperforms the existing methods. The usefulness and effectiveness of the proposed classification is demonstrated by conducting recognition experiments before and after classification.

1 Introduction

Automatic license plate recognition has become an integral part of smart city development in the current scenario for developing countries [1] such as Malaysia, India, and China. This is because license plate recognition plays a vital role in several real-time applications like traffic management, toll collection, traffic law enforcement, and parking lot access control [1, 2]. There are sophisticated methods developed for license plate recognition in the literature, which address the challenges of images with blur, affected illumination, low resolution or dirty, and images with noises (extra symbols) [2]. However, these methods do not focus much on causes created by background and foreground colour changes and the presence of both printed and cursive (handwritten) texts in one image. These challenges are common in the case of Malaysian license plate images, where we can see multi-type images, namely, taxi number plates which are supposed to contain numbers with a white colour background as shown in the left side of Fig. 1*a*, and private plates which should contain numbers with a black colour background as shown in the middle of Fig. 1*a*.

In reality, it is not true because we can see taxi plates containing the mixture of white colour as background as shown in the right of Fig. 1*a*. In the same way, private number plate images contain cursive texts (like handwritten texts) instead of printed texts along with numbers on the black colour background as shown in the right of Fig. 1*a*. These factors make the recognition problem more complex. To alleviate this problem, in this work, we prefer to propose a new classification method rather than developing a unified recognition method, which is not advisable. It can be seen from Figs. 1*b* and *c* that binarisation [3] and OCR [4] report poor results especially for cursive texts, to which binarisation and OCR give nothing before classification, while the same methods [3, 4] report good results for the same images after classification. This is the advantage of classification, which provides an option to tune parameters or choose and modify new methods according to the complexities. Therefore, we consider four types of images, namely, taxi, private, cursive and non-text (images with symbols, headlights etc.) as a four-class classification problem in this work.

2 Related work

Since the focus of the work is to classify the different types of license plate images, we review the methods on license plate recognition to understand the inability to recognise images affected by the above-mentioned challenges. We also review the methods which classify text types extracted from natural scene images and videos. This is because these methods consider texts extracted from the natural scene and video are the same as license plate images in terms of contrast and background variations. Furthermore, we review the state-of-the art binarisation methods, which work well for degraded images and complex background images.

For example, Saha *et al.* [5] proposed an Indian license plate recognition system. Safaei *et al.* [6] proposed a real-time search-free multiple license plate recognition through likelihood estimation of saliency. Panahi and Gholampour [7] proposed the accurate detection and recognition of dirty vehicle pate numbers for high speed applications. Bulan *et al.* [8] proposed segmentation and annotations free license plate recognition with deep localisation and failure identification. It is noticed from the above methods that these methods work well for images which have numbers with white background. Gou *et al.* [9] proposed vehicle license plate recognition based on extremal regions and restricted Boltzmann machines. Kim *et al.* [10] proposed effective character segmentation for license plate recognition under an illumination changing environment. It is noted from the above two methods that they consider images with the black colour

This is an open access article published by the IET, Chinese Association for Artificial Intelligence and Chongqing University of Technology under the Creative Commons Attribution-NoDerivs License (http://creativecommons.org/licenses/by-nd/3.0/)



Fig. 1 *Illustrating the effect of the proposed classification*

a Multi type input license plate images
b Binarisation and recognition results by the method [3] before classification
c Binarisation [3] and recognition results by the method [4] after classification

background. It is not clear that these methods work well for different background images and cursive number plate images. Recently, Polishetty *et al.* [1] proposed a next generation secure cloud-based deep learning license plate recognition for smart cities. This method explores deep learning at the pixel level. Despite the use of deep learning, the method targets images with a white background [11]. In summary, the above-methods work well for a particular type of image but not multi-type images.

There are methods for classifying texts in the literature. For example, Xu *et al.* [12] proposed graphics and scene text classification in the video. Shivakumara *et al.* [13] proposed the separation of the graphics and scene texts in video frames. Roy *et al.* [14] proposed new tampered features for scene and caption text classification in video frames. These methods focus on images affected by particular causes. Raghunandan *et al.* [15] proposed new sharpness features for image type classification based on textual information. Qin *et al.* [16] proposed a video scene text frame categorisation for text detection and recognition. However, these methods extracted specific features to address the issues.

There are methods for recognising texts of different backgrounds or degradations, and contrast variations through binarisation. Su *et al.* [17] proposed a robust document mage binarisation technique for degraded document images. Howe [3] proposed Laplacian energy for document binarisation. Roy *et al.* [18] proposed a Bayesian classifier for multi-oriented video text recognition. Miyaev *et al.* [19] proposed image binarisation for an end to end understanding in natural scene images. In the same way, the methods are developed by exploring deep learning for binarisation of images of different causes. Tensmeyer and Martinez [20] proposed document image binarisation with fully convolutional neural networks. The method formulates binarisation as a pixel classification learning task and applies fully convolutional neural network architecture. This architecture operates at multiple images scaled including full resolution. Javed and Shafait [21] proposed real-time document localisation in natural images by the recursive application of convolution neural network (CNN). The method is formulated as a key point detection problem. Four corner points of a document are jointly predicted using the deep convolutional neural network. Then the prediction is refined using the recursive application of CNN. The main target of the above two methods is degraded document images for binarisation but not images such as license plates, where one can expect a complex background with a degradation effect. In summary, from the above discussions, we can conclude that none of the methods is adequate to address the challenges posed by license plates especially with background colour changes and cursive text presence. However, the methods are good for images of particular types. Therefore, in order to take the advantage of the existing resources rather than developing a new recognition method which can work for different types of license plate images, it is good to classify images such that an appropriate method can be used to achieve a better recognition rate.

Hence, this study presents a new stroke width-based symmetry features for classifying four types of license plate images. Since our intention is to develop a method which does not depend much on the number of training samples, parameter setting etc., and the class which contains non-text images has no boundary limit, we propose new features which show the unique property of texts with the basic setting of a classifier, such that we can discard non-text images without recognising them. In this way, the proposed method extracts features for the classification of complex images and the way it contributes to live projects to improve license plate recognition is significant.

3 Proposed methodology

This method considers detected licensed plate images as the input for classification because there are sophisticated methods that detect license plates well, irrespective of background changes, scripts, and text types [2]. Inspired by the work in [22] where it is noted that stroke width distance and stroke pixels of components play a vital role in extracting characteristics that represent text components, we explore stroke width distance and stroke pixel information to derive new symmetry features for differentiating four types of images in this work. This is illustrated in Fig. 2, where we perform histogram operation on stroke width distances for images shown in Figs. 1*a* and 2*a*. It can be seen from Figs. 2*b* and *c* that the graphs of Taxi and Private appear smooth due to the common text patterns, the graphs of cursive appear less smooth due to the cursive text presence, and the graphs of non-text appear non-linear due to the absence of text pixels. These cues are explored for classifying multi-type license plate images. Therefore, we propose new symmetry features based on stroke width, which results in local candidate stroke pixel (CSP) and global-CSP (G-CSP).

As stroke width distance differs for multi-type images, the same would be reflected in the spatial distributions of L-CSP and G-CSP. Therefore, the proposed method estimates a proximity

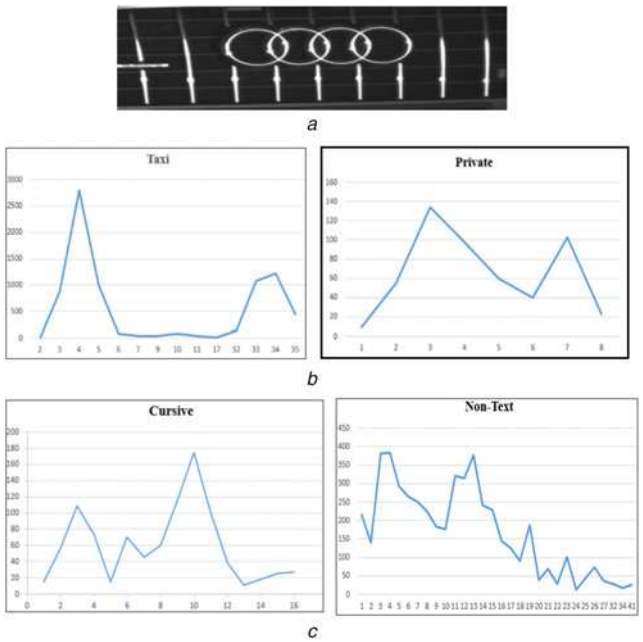


Fig. 2 *Histogram for stroke width distances of taxi, private, handwritten and non-text license plate images*

a Sample non-text image
b Stroke width distances for taxi and private
c Stroke width distances for cursive and non-text

This is an open access article published by the IET, Chinese Association for Artificial Intelligence and Chongqing University of Technology under the Creative Commons Attribution-NoDerivs License (http://creativecommons.org/licenses/by-nd/3.0/)

matrix for L-CSP and G-CSP. In addition, since background colour changes for taxi and private license plate images, intensity values of CSPs also change. To take the advantage of this cue, we multiply the grey value of each CSP to distance, which results in a weighted proximity matrix. For the weighted proximity matrix of local and global symmetries, we extract mean, standard deviation with respect mean, median and standard deviation with respect to the median as features. In total, the proposed method extracts 20 features from local and global weighed proximity matrices for classification with the help of a well-known SVM classifier with default parameter settings.

3.1 L-CSP and G-CSP pixels detection

For each input image, the proposed method finds stroke width distance by moving in the gradient direction, which is perpendicular to the edge direction until it reaches a white pixel [22]. This gives two stroke pixels, which are named as a stroke width pair. For each pair, the proposed method checks the direction given by the gradient vector flow (GVF), which is an improved version of the gradient direction and attracts towards edges, where there is a high force as defined in (1) [23]. If a stroke width pair has opposite directions to each other, the pair is said to be a satisfied GVF opposite arrow symmetry. Motivated by the observation in [23] where it is mentioned that pixels that represent texts satisfy the symmetry in Canny and Sobel edge images of the input image [22, 23], the symmetry is checked for the pixels in both Canny and Sobel edge images. The common pixels which satisfy the symmetry are considered as L-CSPs.

It is illustrated with sample images in Fig. 3, where Fig. 3*a* is a number plate image from a taxi, Fig. 3*b* shows the Canny and Sobel edge images of Fig. 3*a*, the GVF for Canny and Sobel can be seen in Fig. 3*c*, the GVF opposite direction arrows of stroke width pairs can be seen in the magnified portion in Fig. 3*d*, the results of symmetry for respective Canny and Sobel are shown in Fig. 3*e* where we can see background pixels are reduced, and the common pixels which satisfy the symmetry are shown in Fig. 3*f*, where one can judge that most background pixels are eliminated. This results in L-CSP. The effect of L-CSP on a taxi, private, cursive and non-text can be seen in Figs. 3*f* and *g*.

For a stroke width pair, the proposed method finds its mid-point by splitting the line between stroke pixels into two equal halves and its centroid by averaging X and Y coordinates of stroke pixels as shown in Fig. 4*a*, where we can see mid-point and centroid are computed for the CSP. If the mid-point and centroid of the stroke width pair meet at the same point, it is called centroid symmetry as defined in (2). If the pair satisfies centroid symmetry, the pair is classified as L-CSP. The effect of the L-CSP given by centroid symmetry is shown in Figs. 4*b* and *c*, respectively, for taxi, private, cursive and non-text images

$$\mathrm{GVF} = \int\int \mu(a_m^2 + a_n^2 + b_m^2 + b_n^2) + |\nabla f|^2 \, |\alpha - \nabla f|^2 \, \mathrm{d}m \, \mathrm{d}n, \quad (1)$$

where $\alpha\,(m, n) = (a(m, n), b(m, n))$ is the GVF field and $f(m, n)$ is the edge map of the input image

$$CenSym = \begin{cases} 1 & \text{if Midpoint}(x, y) = \text{Centroid}(x, y), \\ 0 & \text{else.} \end{cases} \quad (2)$$

Let P_1 and P_2 be the pixels of a stroke width pair. The proposed method finds the gradient direction of P_1 with which it finds another white pixel, say, P_2 in the edge image. The same is checked for P_2 as shown in Fig. 5*a*, where we can see directions from P_1 to P_2 and P_2 to P_1. If the gradient direction of P_1 leads to meet at P_2 and if the gradient direction of P_2 leads to meet at P_1, it is said to be direction symmetry as defined in (3). The pixels which satisfy the direction symmetry in both Canny and Sobel are considered as L-CSP as shown in Figs. 5*b* and *c* for respective taxi, private, cursive and non-text images. When the proposed method searches for white pixels along the gradient directions of

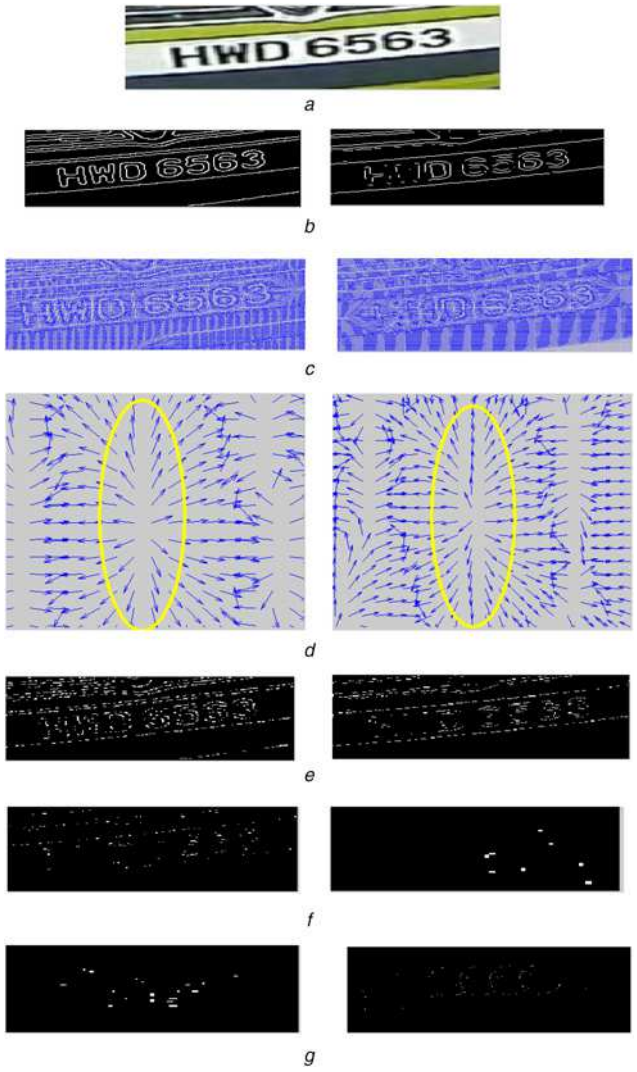


Fig. 3 *L-CSP detection using GVF opposite arrow symmetry*

a Input image
b Canny and Sobel edge images of the image in (a)
c GVF of the images in (b)
d Magnified portion for opposite arrow symmetry
e GVF opposite direction arrow symmetry results
f GVF local symmetry of the taxi in (a) and private image in Fig. 1*a*
g GVF local symmetry of cursive in Fig. 1*a* and non-text in Fig. 3*a*

P_1 and P_2, it records the distances as shown in Fig. 6*a*, where D1 and D2 denote the distances between stroke pixels. If the distance moving from P_1 to P_2 and P_2 to P_1 is the same, it is called as distance symmetry as defined in (4). The pixels which satisfy the distance symmetry in Canny and Sobel are considered as L-CSP as shown in Figs. 6*b* and *c* for four-type images. To strengthen the discriminative power of the symmetry features, furthermore, the proposed method performs intersection operation for the four L-CSP images, which results in G-CSP as shown in Figs. 7*a* and *b* for a taxi, private, cursive and non-text. It is observed from Fig. 7 that only true CSPs are retained by eliminating non-significant pixels

$$\mathrm{DirSym}(X, Y) = \begin{cases} 1 & \text{if } T_1(x, y) = T_2(x, y), \\ 0 & \text{else,} \end{cases} \quad (3)$$

where X and Y are CSPs, $T_1(x, y) = \omega(P_1 \to P_2)$, $T_2(x, y) = \omega(P_2 \to P_1)$, and ω represents the direction

$$\mathrm{DistSym}(X, Y) = \begin{cases} 1 & \text{if } d(P_1, P_2) = d(P_2, P_1), \\ 0 & \text{else,} \end{cases} \quad (4)$$

CAAI Trans. Intell. Technol., 2018, Vol. 3, Iss. 3, pp. 176–183
This is an open access article published by the IET, Chinese Association for Artificial Intelligence and Chongqing University of Technology under the Creative Commons Attribution-NoDerivs License (http://creativecommons.org/licenses/by-nd/3.0/)

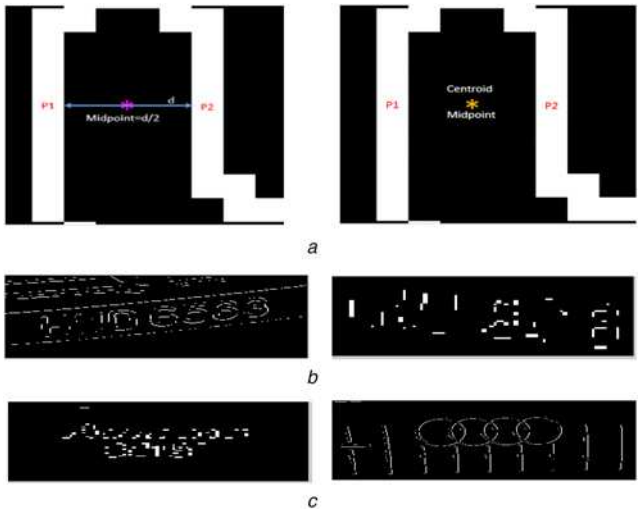


Fig. 4 *Centroid symmetry for L-CSP detection*

a Finding mid-point and centroid
b Centroid local symmetry of taxi and private images
c Centroid local symmetry of cursive and non-text images

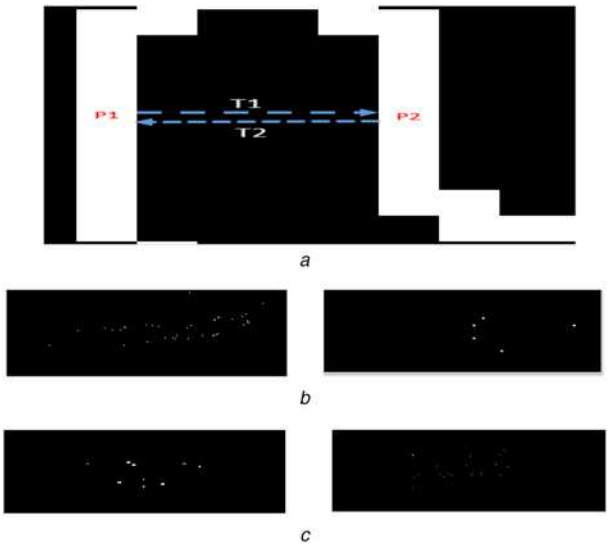


Fig. 5 *Direction symmetry for L-CSP detection*

a Directions between two stroke pixels
b Direction local symmetry of taxi and private images
c Direction local symmetry of cursive and non-text images

where X and Y are CSPs and $d(.,.)$ denotes the distance between two CSPs.

3.2 Weighted proximity matrix for spatial feature extraction

It is clear from the L-CSP and G-CSP of the taxi, private, cursive and non-text type images, which we discussed in the previous section that the CSP pixels in the symmetry images have different spatial distributions. To extract such spatial features, the proposed method computes the centroids for L-CSP images as the shown sample result in Fig. 8*a*, where the centroid (C) for L-CSP of the GVF opposite arrow direction symmetry is computed. Then, it estimates the distance between the centroid and the CSP as shown in Fig. 8*b*. The intensity value of the CSP pixels extracted from the grey image of the input one is considered as weight as shown in Fig. 8*c*. The weight is multiplied by the distance of the respective CSP as defined in (5). This results in a weighted proximity matrix. The proposed method computes mean, standard deviation with

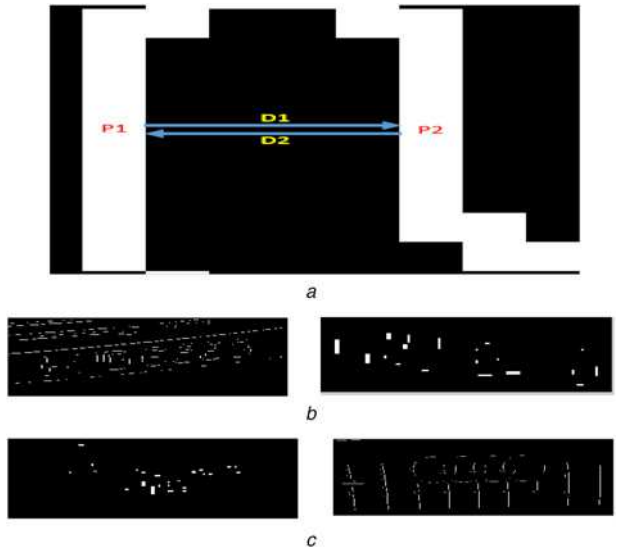


Fig. 6 *Distance symmetry for L-CSP detection*

a Distances between two stroke pixels
b Distance local symmetry of taxi and private images
c Distance local symmetry of cursive and non-text images

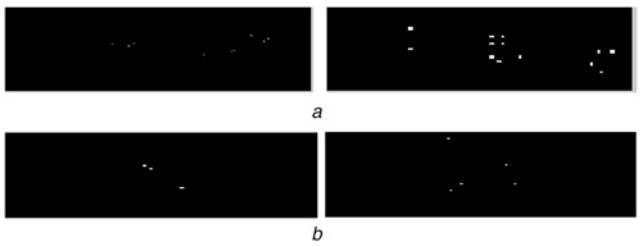


Fig. 7 *G-CSP detection*

a Global symmetry for taxi and private images
b Global symmetry for cursive and non-text images

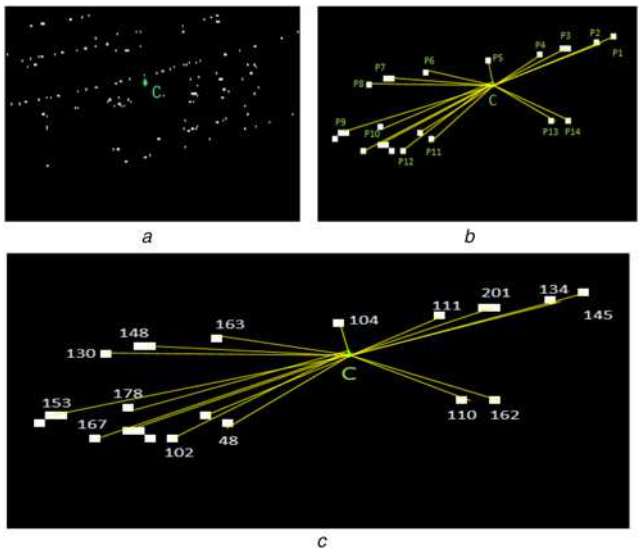


Fig. 8 *Weighted proximity for spatial feature extraction. Note that for (b) and (c), the points are chosen manually from (a) for the purpose of illustration*

a Centroid
b Distance between centroid and pixels
c Weights (grey values) for the pixels in (b)

This is an open access article published by the IET, Chinese Association for Artificial Intelligence and Chongqing University of Technology under the Creative Commons Attribution-NoDerivs License (http://creativecommons.org/licenses/by-nd/3.0/)

respect to mean, median and standard deviation with respect median for L-CSP and G-CSP images. In total, 20 features are extracted, which include four features from four local symmetries ($4 \times 4 = 16$ features) and four more from the global symmetry images. Then we feed the feature matrix to the standard SVM classifier with the radial bias function (RBF) kernel for the final classification. We use 10-fold cross validation for training and testing with basic settings

$$\mathrm{WDPM}(x, y) = \sum_{i=1}^{n} \sum_{j=1}^{m} \mathrm{GW}(i, j) \times \mathrm{CDPS}(x, y), \tag{5}$$

where WDPM is the weighted distance proximity matrix, GW is the weight from the grey scale image, and CDPS denotes the centroid distance proximity matrix.

4 Experimental results

We consider the dataset provided by the research institute funded by the Malaysian Government, where it is a live project to evaluate the proposed method. The dataset includes 335 taxi images, 1389 private images, 223 cursive and 1140 non-text images. In total, there are 3087 images for experimentation. In addition, the images of each class include samples with poor resolution and blur affected by illumination. Therefore, we believe that the considered dataset is good enough to evaluate robustness, usefulness, and effectiveness of the proposed method.

To measure the performance of the proposed and existing methods, we consider the classification rate with a confusion matrix for experiments on classification, and the recognition rate at the character level for recognition experiments. Since there is no ground truth for the dataset, we count manually for calculating measures. For recognition experiments, we consider prior to classification and after classification to show the advantage of the proposed classification. Prior to classification includes the data of four-type images for experiments using different binarisation methods. After classification, includes the data of individual class for experiments using the same binarisation methods. Besides, the same experimentation set up is repeated for each existing classification method to give comparative study.

For comparison, we implement two latest classification methods, namely, Raghunandan *et al.*'s method [15] and Qin *et al.*'s method [16]. Raghunandan *et al.*'s method [15], which is developed for the classification of video, natural, born digital text types and text captured by mobile at the text line level. Since license plate number is the same as the text in video text types and it is state-of-the-art, we use it for a comparative study. Qin *et al.*'s method [16] developed for the classification of scene images. The main reason to choose these two existing methods for a comparative study is that both the methods have the same objective and use text properties for classification.

Similarly, we implement the state-of-the-art binarisation methods for recognition experiments before and after classification, namely, Howe's method [3], which focuses on printed and handwritten document images with colour bleeding effect, Su *et al.*'s method [17], which focuses on degraded document images, Milyae *et al.*'s method [19] focuses on natural scene images, and Roy *et al.*'s method [18] focuses on both natural scene and video images. Since the codes of the methods, namely, Tensmeyer and Martinez [20] who explore deep learning for binarising degraded document images, and Javed and Shafait [21] who use deep learning for localising natural document images, are available publicly, we use the same codes for conducting experiments in this work. We follow almost the same set up and instructions. The reason to consider these different methods for experiments prior to classification and after classification is that each method has its own objective and addresses challenges as our data. Besides, our intention is to show the advantage of the classification method through recognition experiments after classification. Therefore, this work focuses on improvement after classification compared with before classification rather than achieving the high recognition rate.



Fig. 9 *Sample images for successful and unsuccessful classification of the proposed method*

a Sample successful classification of the proposed method
b Sample misclassification of the proposed method

4.1 Evaluating classification method

Sample successful and unsuccessful classification images of the proposed method are shown in Figs. 9*a* and *b*, respectively. Fig. 9 shows that the proposed method works well for images with blur, low contrast, and illumination effect. However, the proposed method fails to classify images with a complex background and too distorted images correctly. Therefore, there is a scope for improvement and extension of the proposed method in the near

Table 1 Confusion matrix of the proposed method with weight, without the weight and combined (T, taxi; P, private; C, cursive; N, non-text license plate images)

Class	Without weight				With weight				Proposed (combined)			
	T	P	C	N	T	P	C	N	T	P	C	N
T	**76.4**	13.1	5.7	4.6	**71.2**	12.4	9.3	6.9	**78.6**	11.4	6.7	3.2
P	6.2	**82.3**	5.0	6.3	9.2	**75.3**	8.3	7.0	3.1	**85.2**	4.3	3.2
C	3.3	12.3	**77.6**	6.6	7.2	15.6	**73.8**	3.3	5.7	9.2	**79.1**	5.8
N	2.2	5.2	9.2	**83.2**	8.4	5.7	7.1	**78.6**	6.5	4.2	3.3	**85.8**

CAAI Trans. Intell. Technol., 2018, Vol. 3, Iss. 3, pp. 176–183
This is an open access article published by the IET, Chinese Association for Artificial Intelligence and Chongqing University of Technology under the Creative Commons Attribution-NoDerivs License (http://creativecommons.org/licenses/by-nd/3.0/)

future. In this work, estimating a weighted proximity matrix for local and global symmetry pixels is the key step for feature extraction. Therefore, to analyse the contribution of this step, we conduct experiments without weight, with weight and combined. The results are reported in Table 1, where we can see that the combined gives better results than the individual.

To show the proposed method is effective, we compare the proposed method with two existing methods as discussed in the previous section. The results of the existing methods are reported in Table 2. When we compare the proposed method (combined) reported in Table 1 and the existing methods reported in Table 2, the proposed method gives better results than the existing methods. Qin *et al.*'s method [16] is developed at the image level and requires objects and text components, hence it scores poor results compared with the proposed method. Raghunandan *et al.*'s method [15] considers quality features for the classification of different video text type images. However, the quality does not differ from one type to another for license plate images. As a result, the existing methods report poor results compared with the proposed method. On the other hand, the proposed method extracts feature based on the spatial distribution of CSPs, which makes a difference compared with the existing methods.

To show that the proposed features are robust to rotation and scaling, we rotate images in respective classes with different

Table 2 Confusion matrix of the existing methods (T, taxi; P, private; C, cursive, N, non-text license plate images)

Class	Raghunandan *et al.* [15]				Qin *et al.* [16]			
	T	P	C	N	T	P	C	N
T	**48.2**	15.4	6.1	30.1	**58.3**	18.1	13.6	9.9
P	24.3	**47.1**	18.4	10.0	15.7	**50.8**	6.0	27.3
C	7.5	28.1	**43.7**	20.5	16.5	12.0	**63.1**	8.1
N	19.2	15.2	21.1	**44.3**	14.2	6.1	12.2	**67.3**

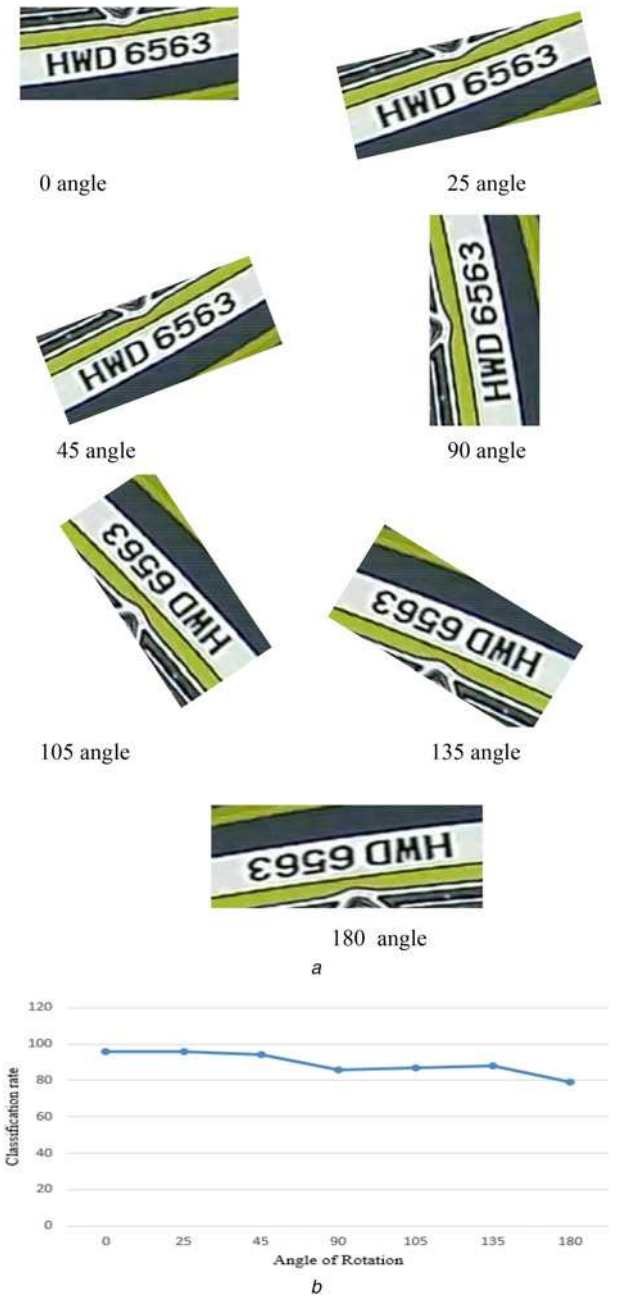


Fig. 10 *Proposed method for different orientations*

a Image of different orientations (in degrees)
b Classification rate of the proposed method for different orientations

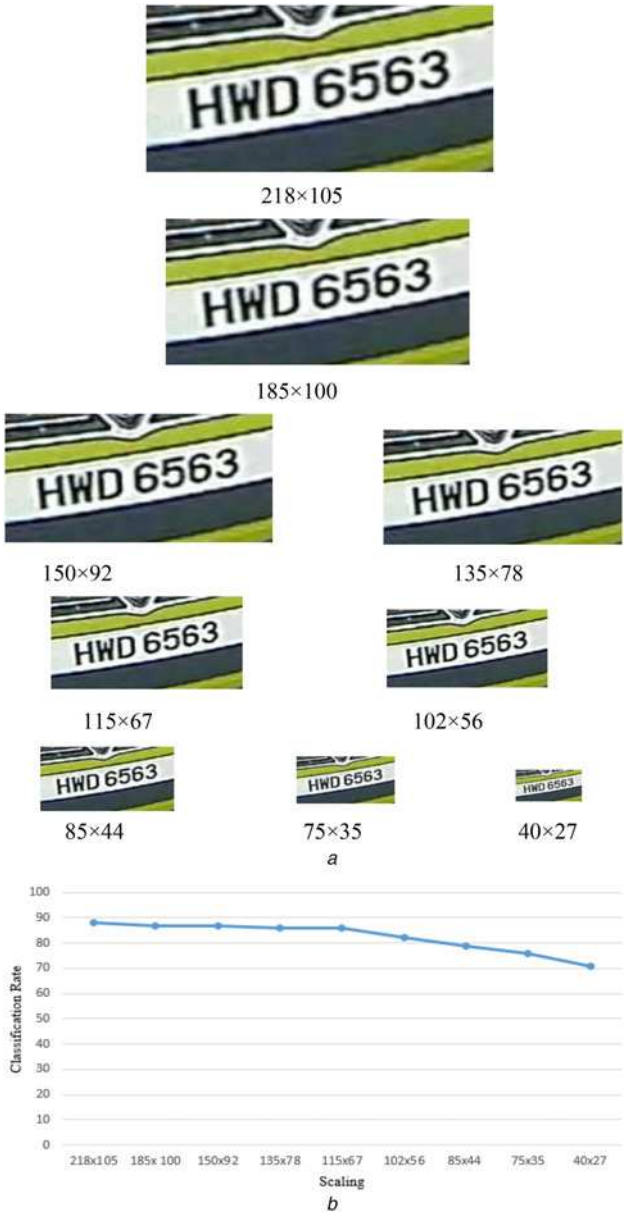


Fig. 11 *Proposed method for different scaling*

a Image of different scaling
b Classification rate of the proposed method for different scaling

This is an open access article published by the IET, Chinese Association for Artificial Intelligence and Chongqing University of Technology under the Creative Commons Attribution-NoDerivs License (http://creativecommons.org/licenses/by-nd/3.0/)

Table 3 Recognition rate of the binarisation methods for before and after classification on each classification methods (T, taxi; P, private; C, cursive; N, non-text license plate images)

Binarisation methods	Before classification	After classification								
		Proposed			Raghunandan *et al.* [15]			Qin *et al.* [16]		
	T+P+C	T	P	C	T	P	C	T	P	C
Roy *et al.* [18]	57.34	76.2	78.6	73.1	63.1	65.1	60.4	70.5	68.2	66.2
Howe [3]	52.12	70.4	74.2	71.4	67.1	61.4	66.4	68.7	63.8	68.4
Su *et al.* [17]	48.97	68.3	72.7	69.0	58.7	55.7	52.7	63.1	59.4	54.1
Milyaev *et al.* [19]	43.12	64.7	71.3	68.0	56.1	51.7	52.0	61.3	55.5	63.8
Tensmeyer [20]	65.24	85.7	79.3	86.2	78.4	71.3	74.6	80.7	78.3	83.7
Javed [21]	62.73	80.4	84.6	82.7	71.5	75.6	72.8	73.7	72.8	79.6

orientations as shown in Fig. 10*a*, where we can see the sample images of different orientations. The proposed method calculates the classification rate for different rotated images. Fig. 10*b* shows the different orientations versus classification rate of the proposed method. It is observed from Fig. 10*b* that the classification rate of the proposed method is almost the same for different orientations. This shows that the proposed features are invariant to rotation. However, there are small deviations in the classification rate for different orientations as shown in Fig. 10*b*. This is due to the distortion created during the different rotations. In the same way, the proposed features are tested on the images of different scaling as shown in Fig. 11*a*, where one can see differently scaled images. The classification rates for the different scaling of classes are shown in Fig. 11*b*, where it is noted that when the images are scaled down to a small size, the classification rate gradually decreases. This is valid because when the image is scaled down to a small size, the image loses quality. However, as per our knowledge, such small-sized images are seldom practically. Therefore, the proposed method is robust to scaling. Furthermore, the proposed symmetry features, namely, GVF opposite arrow, centroid, direction and distance for extracting features are independent of rotation and scaling.

4.2 Validating the proposed classification method

As mentioned in the experimental section, we conduct recognition experiments using several binarisation methods prior to classification and after classification for each existing classification method as reported in Table 3. Table 3 shows that the binarisation methods report better recognition rates after classification compared with prior classification. This is valid because we tune the parameters of the binarisation methods according to the complexity of input classes after classification. For example, parameter c used in Howe *et al.*'s method [3] to control the Laplacian energy is set as 27.5 as a default value. The same default value has been used prior to classification. However, after tuning, we set 50, 72 and 147 for a taxi, private and cursive, respectively. In the same way, we set different window sizes for Milyae *et al.*'s [19] and Roy *et al.*'s [18] methods to achieve better results after classification. For the methods, which explore deep learning [20, 21], we noted that the size of the kernel is one of the key parameters for assessing the performances of the methods. Therefore, we use default kernel sizes of 9×9 and 5×5 of the methods [20, 21], respectively, for calculating the recognition rate of prior to classification. For after classification, we tune the kernel size to 3×3 according to our experimental analysis of different classes. It is noted that the same kernel size of 3×3 works well for different classes as it helps in extracting local information compared with default window size. This is the advantage of the classification method. The results of the methods [20, 21] before classification and after classification are reported in Table 3. Table 3 shows that all the binarisation methods score better results after classification compared with before classification. It is also observed that the methods which use deep learning report the highest recognition rate for both before classification and after classification compared with the classical methods. This is expected because of the powerful deep learning tool. Interestingly, when we compare the recognition rates of all the binarisation methods before classification and after classification, there is a significant difference. This shows that the classification is effective and useful.

However, it is noted from the results of Su *et al.*'s method [17] for before classification and after classification, there is not much improvement compared with the other methods for after classification because the code does not allow us to modify parameters. At the same time, when we compare recognition rates of binarisation methods for classification methods, the recognition rates of all the binarisation methods give better results for the proposed classification method compared with the other existing classification methods. This shows that the proposed method classifies accurately compared with the existing methods. Note that since the non-text class does not contain text information, we do not consider the class for recognition experiments.

5 Conclusion and future work

In this work, we have proposed a new symmetry feature based on the stroke width of character components to classify multi-type license plate images to enhance the performance of the recognition systems. The common pixels which satisfy in both Sobel and Canny result in L-CSP and the common pixels of local symmetry result in the G-CSP. The proposed method extracts spatial features for both local and global symmetry images through a weighted proximity matrix. Furthermore, the feature matrix is fed to an SVM classifier for classification. Experimental results of the proposed and existing methods on classification show that the proposed method outperforms the existing methods in terms of the classification rate. The recognition results of the different binarisation methods show that the proposed classification is useful and effective to improve recognition performance especially when the image of the same class is affected by different causes. The proposed method reports misclassification for images with a complex background, thus we plan to extend the proposed method to overcome this problem in the near future. In addition, we plan to expand the proposed method for more classes affected by different poses.

6 References

[1] Polishetty, R., Roopaei, M., Rad, P.: 'A next generation secure cloud based deep learning license plate recognition for smart cities'. Proc. Int. Conf. on Machine Learning and Applications, USA, 2016, pp. 286–294

[2] Du, S., Ibrahim, M., Shehata, M., *et al.*: 'Automatic license plate recognition (ALPR): A state-of-the-art review', *IEEE Trans. Circuits Syst. Video Technol.*, 2013, **23**, pp. 311–325

[3] Howe, N.R.: 'Document binarization with automatic parameter tuning', *Int. J. Document Analysis Recognition*, 2013, **16**, pp. 247–258

[4] Tesseract. http://code.google.com/p/tesseract-ocr/

[5] Saha, S., Basu, S., Nasipuri, M.: 'iLPR: an Indian license plate recognition system', *Multimed. Tools Appl.*, 2015, **74**, pp. 10621–10656

[6] Safaei, A., Tang, H.L., Sanei, S.: 'Real time search free multiple license plate recognition via likelihood estimation of saliency', *Comput. Electr. Eng.*, 2016, **56**, pp. 15–29

This is an open access article published by the IET, Chinese Association for Artificial Intelligence and Chongqing University of Technology under the Creative Commons Attribution-NoDerivs License (http://creativecommons.org/licenses/by-nd/3.0/)

[7] Panahi, R., Gholampour, I.: 'Accurate detection and recognition of dirty vehicle plate numbers for high speed applications', *IEEE Trans. Intell. Transp. Syst.*, 2016, **18**, pp. 1–13
[8] Bulan, B., Kozitsky, V., Ramesh, P., *et al.*: 'Segmentation and annotation free license plate recognition with deep localization and failure identification', *IEEE Trans. Intell. Transp. Syst.*, 2017, **18**, pp. 1–13
[9] Guo, C., Wang, K., Yao, Y., *et al.*: 'Vehicle license plate recognition based on extremal regions and restricted Boltzmann machines', *IEEE Trans. Intell. Transp. Syst.*, 2016, **17**, pp. 1096–1107
[10] Kim, D., Song, T., Lee, Y., *et al.*: 'Effective character segmentation for license plate recognition under illumination changing environment'. Proc. Int. Conf. on Consumer Electronics, USA, 2016, pp. 532–533
[11] Liu, W., Wang, Z., Liu, X., *et al.*: 'A survey of deep neural network architecture and their applications', *Neurocomputing*, 2017, **234**, pp. 11–26
[12] Xu, J., Shivakumara, P., Lu, T., *et al.*: 'Graphics and scene text classification in video'. Proc. Int. Conf. on Pattern Recognition, Sweden, 2016, pp. 4714–4719
[13] Shivakumara, P., Kumar, N.V., Guru, D.S., *et al.*: 'Separation of graphics (superimposed) and scene text in video'. Proc. Document Analysis Systems, France, 2014, pp. 344–348
[14] Roy, S., Shivakumara, P., Pal, U., *et al.*: 'New tampered features for scene and caption text classification in video frame'. Proc. Int. Conf. on Frontiers in Handwriting Recognition, China, 2016, pp. 36–41
[15] Raghunandan, K.S., Shivakumara, P., Hemantha Kumar, G., *et al.*: 'New sharpness features for image type classification based on textual information'. Proc. Document Analysis Systems, China, 2016, pp. 204–209
[16] Qin, L., Shivakumara, P., Lu, T., *et al.*: 'Video scene text frames categorization for text detection and recognition'. Proc. Int. Conf. on Pattern Recognition, Mexico, 2016, pp. 3875–3880
[17] Su, B., Lu, S., Tan, C.L.: 'Robust document image binarization technique for degraded document images', *IEEE Trans. Image Process.*, 2013, pp. 1408–1417
[18] Roy, S., Shivakumara, P., Roy, P.P., *et al.*: 'Bayesian classifier for multi-oriented video text recognition system', *Expert Syst. Appl.*, 2015, **42**, pp. 5554–5566
[19] Milyae, S., Barinova, O., Novikova, T., *et al.*: 'Image Binarization for end-to-end text understanding in natural images'. Proc. Int. Conf. on Document Analysis and Recognition, USA, 2013, pp. 128–132
[20] Tensmeyer, C., Martinzez, T.: 'Document image binarization with fully convolutional neural networks'. Proc. Int. Conf. on Document Analysis and Recognition, Japan, 2017, pp. 99–104
[21] Javed, K., Shafait, F.: 'Real time document binarization in natural images by recursive application of a CNN'. Proc. Int. Conf. on Document Analysis and Recognition, Japan, 2017, pp. 105–110
[22] Epshtein, B., Ofek, E., Wexler, Y.: 'Detecting text in natural scenes with stroke width transform'. Proc. Computer Vision and Pattern Recognition, USA, 2010, pp. 2963–2970
[23] Khare, V., Shivakumara, P., Raveendran, P., *et al.*: 'A new sharpness based approach for character segmentation in license plate images'. Proc. Asian Conf. on Pattern Recognition, Malaysia, 2015

This is an open access article published by the IET, Chinese Association for Artificial Intelligence and Chongqing University of Technology under the Creative Commons Attribution-NoDerivs License (http://creativecommons.org/licenses/by-nd/3.0/)