Robust License Plate Detection in Complex Scene using MSER-Dominant Vertical Sobel

Gamma Kosala, Agus Harjoko, and Sri Hartati

***Abstract*—This paper presents a robust method to locate a license plate in a complex scene. In contrast to the most existing method which use the non-handcrafted feature to locate the license plate area, our method uses a modified handcrafted feature. We used Maximally Stable Extremal Region (MSER) combining with dominant vertical Sobel to construct essential biner images. Closing morphology operation is implemented to merge the contour extracted by MSER and dominant vertical edge detection. Based on the area and ratio of contour, license plate candidate area is selected. Furthermore, Support Vector Machine (SVM) is introduced to choose a license plate area by analyzing the Histogram of Oriented Gradient (HOG) of each candidate. For performance evaluation, two datasets consisting of complex scene images under different conditions are tested. The main advantage of this approach is that it is faster than non-handcrafted feature-based method while maintaining the high accuracy of plate detection.**

***Index Terms*— license plate detection, Maximally Stable Extremal Region, edge detection, vertical Sobel operator**

I. Introduction

Automatic license plate recognition (ALPR) systems are generally divided into three stages; plate detection, character segmentation, and character recognition. The first stage, plate detection, is the most critical and challenging stage in ALPR systems. Plate detection stage aims to find the license plate area in an input image. It will be more difficult on a road that has a complex background than on a park gate that has a monotonous background. Another challenge is the different illumination in each image. It depends on the location and the time when the images are being captured.

Several methods were developed to accomplish plate detection tasks in a complex background using non-handcrafted features. Non-handcrafted features are obtained by directly extracting input image using deep learning network architecture. The results of deep learning feature extraction can produce useful features for some cases such as scene recognition [1], facial expression recognition [2], medical image analysis [3], land-use classification [4] or electrical equipment analysis [5]. The disadvantage of this method is that computational time tends to be quite high because the network used many layers to build deep learning architecture.

Another approach to detecting plate can be performed by extracting handcrafted feature. Handcrafted features usually extracted based on color, texture, and shape of the object. For example, in [6], all of those features are used for identifying rice variety. While for plate detection task, the handcrafted feature usually becomes a reference to determine whether the area is plate area or not. This method has a relatively faster but has less accurate compared to non-handcrafted feature-based, especially for plate detection in the complex background.

This paper offers a license plate detection method using handcrafted feature extraction to get the candidate of the plate area. The handcrafted features we created based on MSER [7]. MSER was modified by adding the dominant vertical Sobel extraction. The dominant vertical Sobel can overcome the weaknesses of MSER when detecting plates in varying sizes [8]. The contour is created by combining MSER and Dominant Vertical Sobel extraction using closing operation. Each contour being analyzed based on area and aspect ratio to select the plate candidate.

Furthermore, each candidate is segmented using a bounding box. Each bounding box is enlarged vertically and horizontally to overcome the problem on some positive candidate which plate area is not entirely included in the bounding box. Furthermore, The bounding box refinement is applied to correct the bounding box in the candidate so that the bounding box in the positive candidate precisely locate the plate area. The refinement is done by analyzing the cumulative intensity histogram of the candidate from its MSER-Dominant Vertical Edge binary image on the vertical axis and horizontal axis. At the candidate selection stage, the HOG feature is extracted on each candidate. Then, the SVM architecture is employed to separate candidates into license plates and non-license plates.

Briefly, the contributions of this paper are:

1. Construct the handcrafted feature by combining MSER and Dominant Vertical Sobel. The vertical Sobel extraction results can help MSER in finding candidate plate areas in complex backgrounds. Vertical Sobel is able to tackle the scale invariance problem in MSER. Only the dominant vertical edge is extracted in order to avoid excessive noise.

2. Adding a mechanism for expanding the bounding box then followed by the bounding box refinement. This approach helps the detection to locate the license plate area precisely. Expansion of the bounding box is done by increasing the height and width of the candidate bounding box. While bounding box refinement is done by analyzing cumulative intensity histogram from MSER-Dominant Vertical Edge binary image

Manuscript received March 13, 2019; revised February 25, 2020. This work was supported by Directorate General of Higher Education, Ministry of Research and Education, Indonesia.

Gamma Kosala is a doctoral student in the Department of Computer Science and Electronics, Universitas Gadjah Mada (UGM), Yogyakarta, Indonesia. (e-mail: gamma.kosala@mail.ugm.ac.id)

Agus Harjoko and Sri Hartati are with the Department of Computer Science and Electronics, Universitas Gadjah Mada (UGM), Yogyakarta, Indonesia. (e-mail: {aharjoko, shartati}@ugm.ac.id})

The rest of paper is organized as follows. Section II explains some previous work on plate detection, more detailed description of the method proposed is in section III, section IV contains an explanation of experiments conducted to test and evaluate the proposed method, and last is a conclusion on section V.

II. Related Works

License plate detection based on background conditions can be divided into two classes, detection in simple background and detection in a complex background. The simple background could be classified as a monotonous background where the vehicle position tends to be invariable (e.g., parking gate and toll entrance). Whereas a complex background could be classified as various background where the vehicle position could be anywhere in a frame (e.g., intersection and city road).

Plate detection, in a simple background, tends to be easier than plate detection in complex backgrounds. Several plate detection research in simple background use handcrafted features as a reference for locating the plate. Since license plates commonly have a rectangular shape, several plate detection methods such in [9] and [10] use edge detection method to find the edges which form the rectangle of the license plate. In [9], Vertical Edge Detection Algorithm (VEDA) is introduced. It only retrieves edge information for vertical lines. The Unwanted-Line Elimination Algorithm (ULEA) is implemented before the VEDA. ULEA aims to eliminate minor edges by removing edges with an area of one pixel and having no neighbors at 0˚, 45˚, 90˚, and 135˚ angles. Other approaches in [11] and [12] find the plate based on the plate's color in a specific country. In [11], the plate area is located based on blue and yellow areas, which are standard plate colors in China. If the plate's color has the same color as the vehicle body, plate areas are determined by using vertical edge analysis.

Connected Component Analysis (CCA) also could be used to find plate location as in [13], [14], and [15]. CCA works by scanning a binary image and labels its pixels as components based on pixel connectivity. In [14], CCA is implemented to find every region extracted and classify it into plate region and non-plate region based on similarity the area, height, and width of the candidate area with them of license plate template. In [15], CCA also can be applied to plate detection in low-resolution video. While in [16], Harris corner detection also could be applied to get the plate candidate. Each detected corner is merged with its neighbor corner to form a contour. Based on its morphological shape, each contour is examined whether the contour is plat or not. The use of vertical edge detection, color detection, CCA, and corner detection are fast and good enough in locating a plate in the simple background, whereas they will produce some noise that affects the detection accuracy when implemented in a complex background.

Meanwhile, some other complicated features are employed for plate detection in a complex background. MSER method is often used to find text areas candidate in the image [17], [18], [19]. Many studies also used an MSER-based method to detect license plate since it consists of letter and number [20], [21], [22]. In [22], characters candidate area are extracted using MSER. Then each characters candidate is divided into four types and extract the suspected first node (the top-left character) based on its neighbor MSER distribution characteristic. Last, each character candidate is labeled to get plate location based on the detected suspected initial node and the corresponding label-moveable maximal MSER clique.

Several recent studies used non-handcrafted feature to get plate location [23],[24],[25],[26]. They use deep learning architecture that directly extracts the feature from input images. In [25], a 37-class convolutional neural network (CNN) is trained to detect characters in an input image. The detected character which is not included in plate area then eliminated by another CNN classifier.

Non-handcrafted feature tends to reach high accuracy in locating plate but have expensive computational time. In order to get more efficient computational time, we create a new feature based on the handcrafted feature. This new feature is based on MSER. Vertical edge is added to overcome the lack of MSER for detecting an object in varying sizes.

III. The Proposed Method

This section explains our approach to find the license plate in a complex background. It also gives more detail technique on the proposed contribution. Our method is divided into four main stages; starting from the preprocessing stage, the candidate plate extraction stage, the bounding box refinement stage, and last is the plate selection stage. Overall stages of the proposed plate detection method can be seen in Fig. 1.

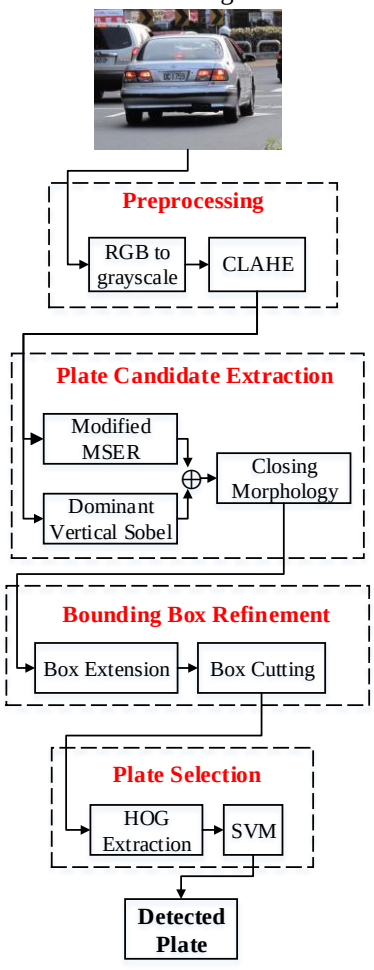
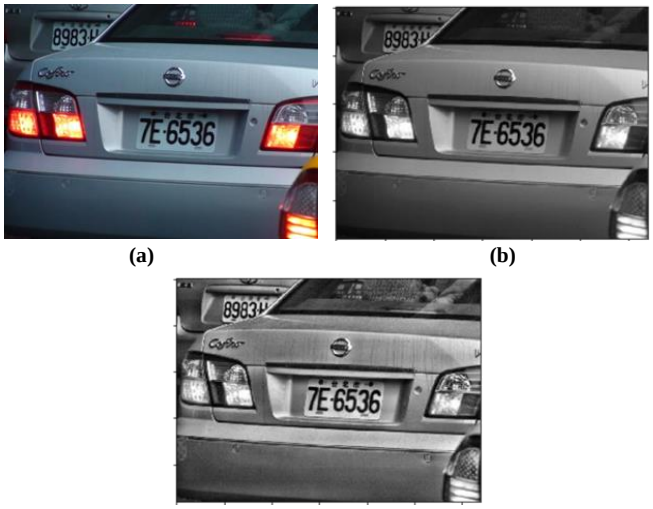


Fig. 1. The overall stages of our proposed method.

A. Preprocessing

First, the input image will be converted into the grayscale. Furthermore, the Contrast-Limited Adaptive Histogram Equalization (CLAHE) [27] is implemented in the greyscale image to improve contrast in the image. The difference of input image, grayscale image, and image after CLAHE implementation can be seen in Fig. 2



(a) (b) (c)

Fig. 2. Preprocessing stages, (a) input image (b) grayscale image (c) after CLAHE implemented

B. Candidate Extraction

The candidate extraction stage is divided into two subsections. First is MSER-Dominant Vertical Sobel feature extraction then followed by closing morphology operation. The first step aims to get blobs, which can be used as a candidate plate area, while the second stage aims to unite the blobs so that it becomes a single unit contour.

1) MSER-Dominant Vertical Sobel Extraction

The MSER extraction is applied by examining each region in the image after thresholding process. It used the threshold value from 0 to 255 in turn. The regions that change little or even stay the same are selected as MSER. The mathematical definition of linear MSER [28] is in (1),

$$q(i) = \frac{|Q_i - Q_{i-\Delta}|}{|Q_{i-\Delta}|} \tag{1}$$

Where Q_i is the connected region of threshold i, Δ denotes the small change of the grey value and $q(i)$ is the changing rate of region Q_i with threshold i. When the $q(i)$ is the local minimum, the Q_i is the MSER.

We modified MSER by limiting the aspect ratio of MSER blob as in (2). The width Q_{iw} is divided by the height Q_{ih} in each blob from MSER extraction to get their aspect ratio. Since we want the blob to represent the character in the plate, which has no significant difference in length and height, so we define the aspect ratio threshold τ. The blobs which have aspect ratio more than τ are removed, where we define $\tau = 5$.

$$Q_{i\_modified} = \begin{cases} 255, & if \quad \frac{Q_{iw}}{Q_{ih}} < \tau \\ 0, & otherwise \end{cases} \tag{2}$$

MSER can detect objects from various points of view. It is beneficial when detecting the plates in the complex background such as in intersection where the vehicle has various positions. However, some character still cannot be extracted into the blob due to MSER lack in overcoming the problem of object size scaling. Fig. 3 shows some MSER extraction results fail to extract all the character on the input image.

(a) (b)

Fig. 3. MSER extraction, (a) input image (b) contour after MSER extraction. Some characters cannot extracted properly.

Overcoming the problem of object scaling in MSER, we tried to combine MSER and vertical Sobel. Vertical Sobel is extracted by applying convolution (⊗) between CLAHE image I and Sobel Vertical operator G_y as in (3)

$$S = I \otimes G_y = I \otimes \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \tag{3}$$

Vertical Sobel is also often used for detecting plate in the simple background. The vertical Sobel is hardly implemented to detect the plate in a complex background because it will produce too much noise. Avoiding that problem, we limit the threshold when we convert the result of vertical Sobel extraction into a binary image. As in (4), λ is the threshold parameter where λ is 80% of 255. It will extract only the top 20% edge intensity x in grayscale image $S(x)$ to convert into a binary image. Only the outermost point of the edge will be extracted and combined with the MSER blob.

$$S(x) = \begin{cases} 255, & if \; x > \lambda \\ 0, & otherwise \end{cases} \tag{4}$$

Fig. 4 shows the difference between the dominant vertical Sobel extraction results and the typical vertical Sobel extraction results. The dominant Sobel vertical extraction has less blob than the typical vertical Sobel extraction but still remain the important blob in plate area.

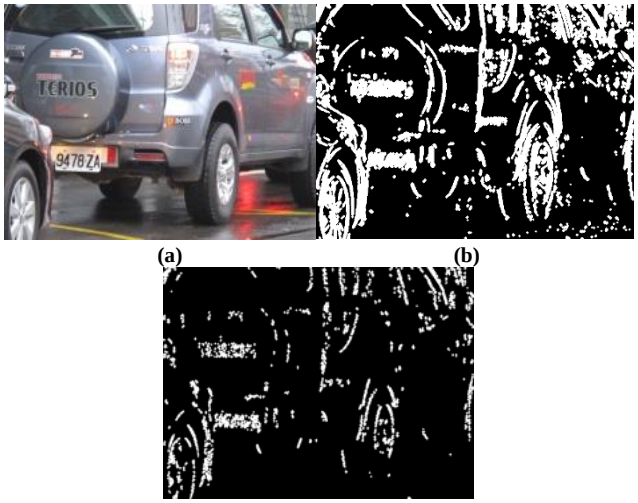


Fig. 4. Vertical edge extraction, (a) input image (b) binary image after vertical edge extraction (c) binary image after only top 20% vertical edge extraction

After extraction of each feature, merging MSER and dominant vertical Sobel is done using OR logic operations as in (5).

$$MSER\_DVE = Q_{i_{modified}} \vee S(x) \quad (5)$$

Some results from MSER-Dominant Vertical Sobel merging process are shown in Fig.5. Dominant vertical Sobel could fill the blank space in the plate area where MSER has not covered yet.

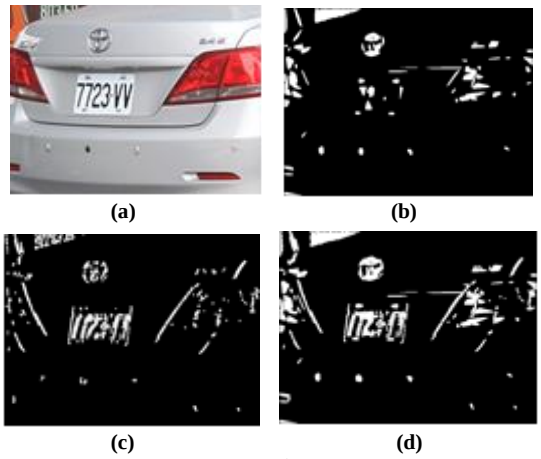


Fig. 5. MSER-Dominant Vertical Edge features extraction, (a) input image (b) MSER extraction (c) Dominant Vertical Edge extraction (d) MSER-Dominant Vertical Edge

2) Closing morphology

The closing morphology operation is implemented to close the space between characters so that the license plate becomes a single contour. Accordingly, the closing operation of image X by structuring element P, denoted $X \bullet P$, is defined as

$$X \bullet P = (X \oplus P) \ominus P \quad (6)$$

Where $\oplus$ and $\ominus$ are dilation and erosion operations in (6), respectively. The closing operation begins with the dilation process for thickening pixels then followed by erosion, which serves to eliminate noise in the form of small separated pixels.

Contours from the closing operation are detected by implementing connected component extraction method. We choose a contour as a plate candidate based on the area and the aspect ratio. After being selected, each candidate is isolated using a bounding box. Fig. 6 shows the process to find the candidate after closing morphology is implemented.

Fig. 6. Plate candidate extraction

C. Bounding Box Refinement

In some positive candidates, bounding boxes do not cover the entire plate area. Resolving that problem, we expand the bounding box on the vertical axis and horizontal axis. The bounding box is expanded up and down by 10% of the candidate's height, whereas it is expanded left and right by 35% of the candidate's width. The example of bounding box expansion can be seen in Fig. 7

Fig. 7. Bounding box extension in the vertical and horizontal axis

However, the bounding box extension causes some candidate's boxes to exceed the plate area. To fix that problem, the box refinement is applied based on the horizontal and vertical projection of binary images after closing operation. First, we build histogram based on the distribution of cumulative intensity in the vertical axis and horizontal axis. Then we define the threshold value. The threshold is set to 85% of the average intensity in each histogram.

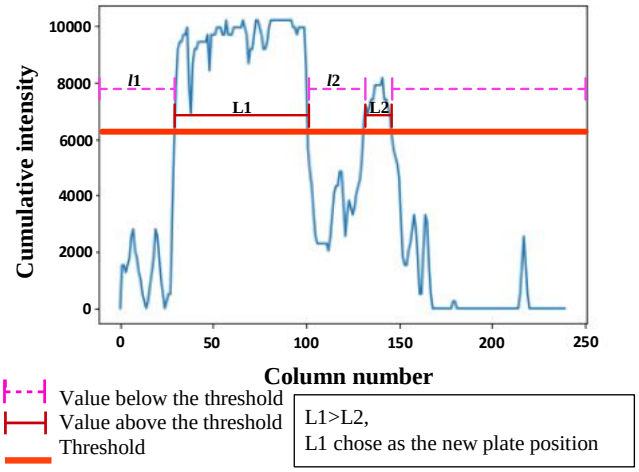


Fig. 8. Plate candidate area refinement based on histogram analysis

When the cumulative intensity below the threshold of more than 5 pixels in a row, it will be not considered as the plate area. The plate area location is assumed as the longest range above the threshold. Based on that range, a new bounding box is created. Fig.8 is an illustration of the bounding box refinement based on the histogram of vertical projection, while Fig. 9 shows the example of how the bounding box refinement process works on the candidate.

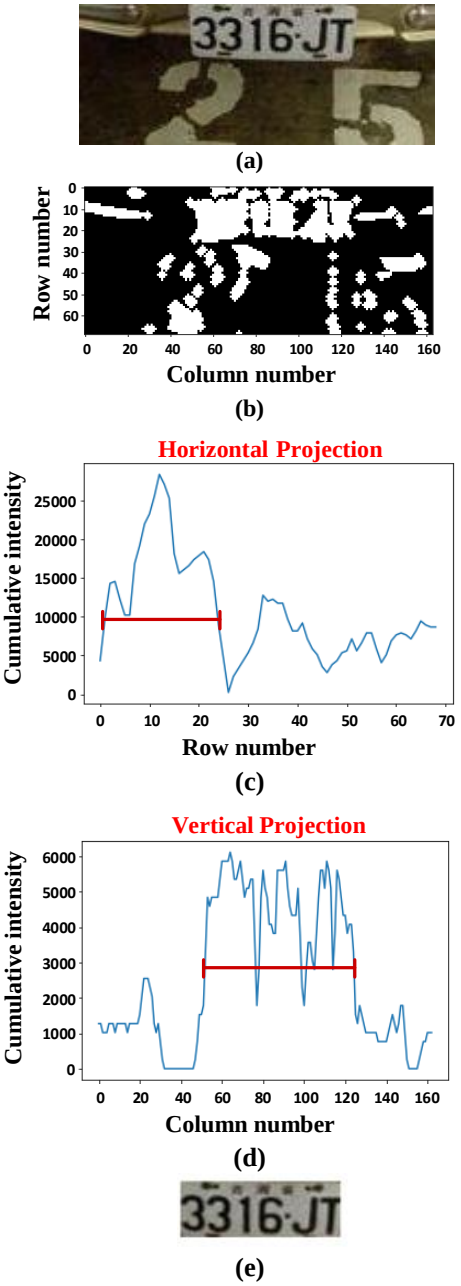


Fig. 9. Bounding box refinement step by step, (a) plate candidate (b) MSER_DVE binary image (c) horizontal projection analysis (d) vertical projection analysis (e) candidate after box refinement

D. Plate Selection

After each candidate passes the bounding box refinement stage, candidates are selected based on the aspect ratio. Since the license plate has a longer width than the height and has the aspect ratio around 3, we eliminate the candidate which has the aspect ratio below 2.5 and above 5.5. Furthermore, the selection of candidates is made by extracting HOG features in each candidate then classify it using SVM architecture.

In order to get the same dimension of HOG feature, All candidate is resized into 30 x 90 pixels. We divide the candidate into cells and compute the gradient in every cell. Gradient value is stored in one of 9 bins available based on the direction of the gradient. Normalization using the block is applied to make gradient resistance for changes in illumination and contrast. Block consist of some cells and typically overlap each other.

Equation (7) shows how to compute the gradient meanwhile (8) shows how to find the direction of the gradient.

$$R = \sqrt{{G_x}^2 + {G_y}^2} \tag{7}$$

$$\theta = \tan^{-1}\left(\frac{G_y}{G_x}\right) \tag{8}$$

Where G_x is the gradient in the vertical axis and G_y is the gradient in the horizontal axis. HOG feature in each candidate become an input into the SVM. SVM employed Radial Basis Function (RBF) as its kernel. SVM will classify the candidates into plate area and non-plate area.

IV. Experiments

This section will explain the performance evaluation method and those result on the proposed method. The experiments are implemented using python on i5 CPU and NVIDIA GTX 1070 GPU with 8GB memory.

Performance evaluation includes two aspects, the accuracy and the computational time. In the calculation of plate detection accuracy, a plate detection is said to be successful if the bounding box covers the entire plate area and the Intersection over Union (IoU) value does not exceed 50%. IoU is an evaluation protocol used for text detection on natural scenes [29] and also used for the evaluation of some plate detection [25],[26]. Meanwhile, for the computational time, it is measured from the time to read the input image until the license plate area is detected by applying the proposed method. The computational time in each dataset is calculated by averaging the computational time in its subsets.

A. Dataset

To measure the performance of the proposed method, two public datasets are used. Both datasets consist of vehicle images on highways that have varied and complex backgrounds.

The first dataset is the Application-Oriented License Plate (AOLP) [20]. It is a collection of vehicle images on highways in Taiwan. This dataset consists of 2049 images which are divided into 3 subsets with different position of the camera. Access Control (AC) subset refers to the cases that a vehicle passes a fixed passage at a reduced speed or with a full stop. Parking gate or toll entrance are examples of this case. Traffic Law Enforcement (LE) subset refers to the cases that a vehicle travels at normal or higher speed but violates traffic laws, such as a traffic signal or speed limit, and is captured by a roadside camera. Road Patrol (RP) subset refers to the cases that the camera is installed or handheld on a patrolling vehicle which takes images of the vehicles with arbitrary viewpoints and distances.

The second dataset is PKUData [30]. It consists of vehicle images in China traffic. This dataset has 5 subsets (G1-G5) with the total images are 3977. Each subset has varied condition and background characteristics. The G1 subset includes a collection of car images on toll roads during the shady daytime. The G2 subset includes the image of cars and trucks on toll roads during sunny daytime. The G3 subset includes the image of cars and trucks on the highway during nighttime. The G4 subset includes the image of cars and trucks on urban roads during sunny daytime. The G5 subset includes the image of cars and trucks in urban areas during the daytime till nighttime. In subset G1-G4, there is only one vehicle in each image. Meanwhile, in G5, some images consisted of more than one vehicle.

B. Performance Evaluation on AOLP

The candidate extraction stage produces a collection of positive candidates and negative candidates in each subset. Positive candidates are the candidates that contain plate area, while negative candidates are the candidates that do not contain plate area. We made a performance comparison between the proposed method and its constituent method in extracting candidates. Table I shows the number of candidates detected in each subset, both positive and negative candidate for each method. For the performance evaluation in the candidate extraction stage, recall value was calculated and plotted in Fig.10.

TABLE I
THE CANDIDATE EXTRACTION EVALUATION IN AOLP DATASET. PERFORMANCE COMPARISON BETWEEN MSER, VERTICAL SOBEL AND MSER-DOMINANT VERTICAL SOBEL

Method	Subset	#license plate	#positive candidate	#negative candidate	Recall (%)
Vertical Sobel	AOLP AC	681	661	3341	89.72
	AOLP LE	783	664	9360	84.80
	AOLP RP	611	402	6825	65.79
MSER	AOLP AC	681	672	3314	98.68
	AOLP LE	783	760	10417	97.06
	AOLP RP	611	593	3728	97.05
Proposed Method	AOLP AC	681	679	1275	99.71
	AOLP LE	783	776	3586	99.11
	AOLP RP	611	607	1374	99.35

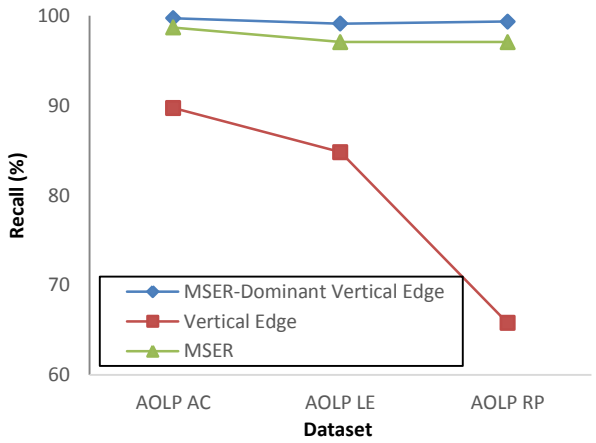


Fig. 10. Recall in the candidate extraction stage on AOLP dataset for each method.

Compared to Vertical Sobel and MSER, the proposed method achieved the highest recall, with an average of 99.39%. While the average recall for Vertical Sobel is 80.11% and the average recall for MSER is 97.60%. It shows that adding dominant vertical edges on MSER could increase the recall by 1.79 %, as shown in Fig.10. Moreover, based on the negative candidates produced, our method produces at least negative candidates with an average of 3 negative candidates per frame, while for Vertical Sobel 10 negative candidates per frame and MSER 9 candidates per frame.

However, this method still failed in extracting some license plates as the candidate because of bad illumination or stain in those plates. In LE, the number of negative candidates is quite high compared to other subsets in AOLP. It is influenced by the background in LE, which is more complex and the distance between the camera and the car is so varied. More candidates in LE also make the computational time slower than the other subset, as seen in Table II.

TABLE II
COMPUTATIONAL TIME ON AOLP DATASET

Subset	SPEED (ms)
AOLP AC	38
AOLP LE	73
AOLP RP	36
AVERAGE	**49**

Candidates classification stage divide the candidates into plate area and non-plate area. SVM is employed to classify the candidates based on HOG features. Missing plates in candidate extraction stage will be added as a false negative in the classification stage. Result comparison with other methods is presented in Table III. It shows that the accuracy of the proposed method is a little bit below to the accuracy obtained on [25] and [26], but we still got a better result than the previous works in [20]. In [20], they used MSER and edge clustering.

Meanwhile, in [25] and [26], they used deep learning architecture, so it does not only increase the accuracy but also increase the computational time. Among other methods, the proposed method reached the best computational time since our method based on fast handcrafted features. This method not only can get near to the accuracy of the state-of-the-art method but also decrease the computational time around 87%.

TABLE III
EXPERIMENTAL RESULT ON AOLP DATASET

METHOD	DETECTION PERFORMANCE (%)			RECOGNITION SPEED (ms)
	AC	LE	RP	
HSU ET AL [20]	96	95	94	260
LI ET AL [25]	98.38	97.62	95.58	1000 - 2000
LI ET AL [26]	**99.56**	**99.34**	**98.85**	400
PROPOSED METHOD	98.01	95.96	96.85	**49**

C. Performance Evaluation on PKU

In PKU, the proposed method will also be compared with Vertical Sobel and MSER methods in candidate extraction

stage. The detection procedures are identical with the procedure used in AOLP, and the only difference is the size of the images. PKU images are larger than AOLP images. In average, AOLP images are 320 × 438, while PKU images are 924 × 1290.

TABLE IV
THE CANDIDATE EXTRACTION EVALUATION IN PKU DATASET. PERFORMANCE COMPARISON BETWEEN MSER, VERTICAL SOBEL, AND MSER-DOMINANT VERTICAL SOBEL

Method	Subset	#license plate	#positive candidate	#negative candidate	Recall (%)
Vertical Sobel	PKU G1	810	674	13809	83.21
	PKU G2	701	570	9645	81.31
	PKU G3	743	516	10540	69.45
	PKU G4	572	472	7076	82.52
	PKU G5	1464	1240	22305	84.70
MSER	PKU G1	810	805	3964	99.38
	PKU G2	701	697	3514	99.43
	PKU G3	743	738	4028	99.33
	PKU G4	572	527	6892	92.13
	PKU G5	1464	1369	22090	93.51
Proposed Method	PKU G1	810	807	2398	99.62
	PKU G2	701	700	1964	99.85
	PKU G3	743	741	2288	99.73
	PKU G4	572	551	3737	96.32
	PKU G5	1464	1407	12518	96.10

Since the images in PKU are larger than the images in AOLP, the extracted candidates in PKU tend to be more in number than the extracted candidates in AOLP. Candidate extraction result in Table IV shows that G4 and G5 have more missed plate than other subsets. In both subsets, several input images have very bright or very dark lighting conditions. Our MSER-Dominant vertical Sobel feature is not good enough in detecting candidate on terrible illumination images since it finds the candidate based on the different intensity in an object with its neighbors.

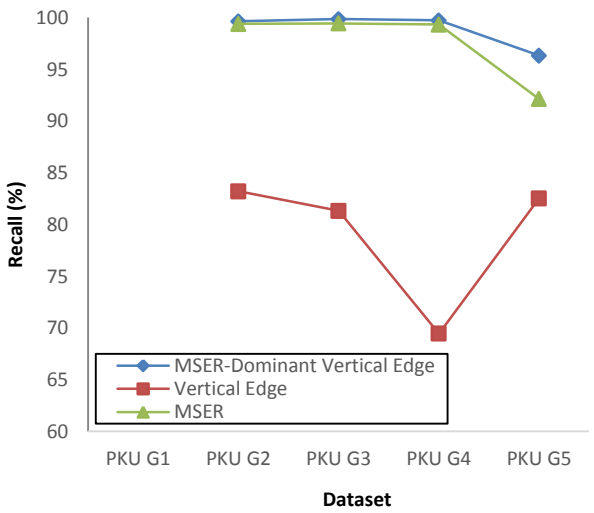


Fig. 11. Recall in the candidate extraction stage on PKU dataset for each method.

However, the proposed method still get the highest recall in candidate extraction stage compared with two other methods, MSER and Vertical Sobel, in PKU Dataset. The average recall reach by the proposed method is 98.32%. Meanwhile, the average recall in Vertical Sobel and MSER are 80.24% and 96.76%, respectively. As shown in Fig.11, in PKU Dataset, the proposed method also could increase the recall of MSER by 1.57%. The negative candidate produced by the proposed method in PKU Dataset is also less than the negative candidate produced by Vertical Sobel or by MSER. In average, the proposed method produces 6 negative candidates per frame. Meanwhile, Vertical Sobel and MSER produce 16 and 10 negative candidates per frame, respectively.

Larger input images also affect the computational time in PKU to be longer than the computational time in AOLP. The computational time in PKU shown in Table V. G4 - G5 also take longer time than G1 - G3 due to larger input images. They images dimension are 1200 × 1600 while G1 - G3 images dimension are 728 × 1082.

TABLE V
COMPUTATIONAL TIME ON PKU DATASET

Subset	SPEED (ms)
PKU G1	102
PKU G2	98
PKU G3	95
PKU G4	220
PKU G5	221
AVERAGE	**147**

After all the candidates were classified in candidate classification stage, the detection accuracy and the computational time in proposed method had been compared with other methods in [31], [21], [30], and [26]. The comparison both the accuracy and the computational time in each subset between the proposed method and other methods are shown in Table VI. The proposed method reaches the second highest accuracy, with the average accuracy is 99.73%. It is 1.38% below the highest accuracy achieved by [26]. Meanwhile, in computational time, we acquire 47% faster than [26]. It affirms that our method could maintain high accuracy while decreasing the computational time in the license plate detection task. Some plate detection results for AOLP dataset and PKU dataset are shown in Fig. 12.

V. CONCLUSION

In this paper, we have presented an MSER-Dominant Vertical Sobel feature to detect license plate in a complex background. We add vertical Sobel on MSER to overcome the lack of MSER to detect scaling object. We also modified the vertical Sobel by limiting only the top 20% to be extracted. The fusion of both features can work well to detect license plate in a complex background. Evaluation performance on AOLP dataset and PKU dataset show that our method can improve the recall of MSER method by around 2%. It also achieves high accuracy and decreases the computational time until 47% from the state-of-the-art method. While other methods use the deep learning architecture that needs many layers, our method use modified handcrafted feature to get fast computational time. Since the computational time is fast, our method could be used for the real-time task.

TABLE VI
EXPERIMENTAL RESULT ON PKU DATASET

METHOD	DETECTION PERFORMANCE (%)						SPEED (ms)
	G1	G2	G3	G4	G5	AVERAGE	
ZHOU ET AL [31]	95.43	97.85	94.21	81.23	82.37	90.22	475
LI ET AL [21]	98.89	98.42	95.83	81.17	83.31	91.52	672
YUAN ET AL [30]	98.76	98.42	97.72	96.23	97.32	97.69	42
LI ET AL [26]	99.88	99.86	99.60	100	99.31	99.73	279
PROPOSED METHOD	98.81	98.52	99.00	97.80	97.61	98.35	148

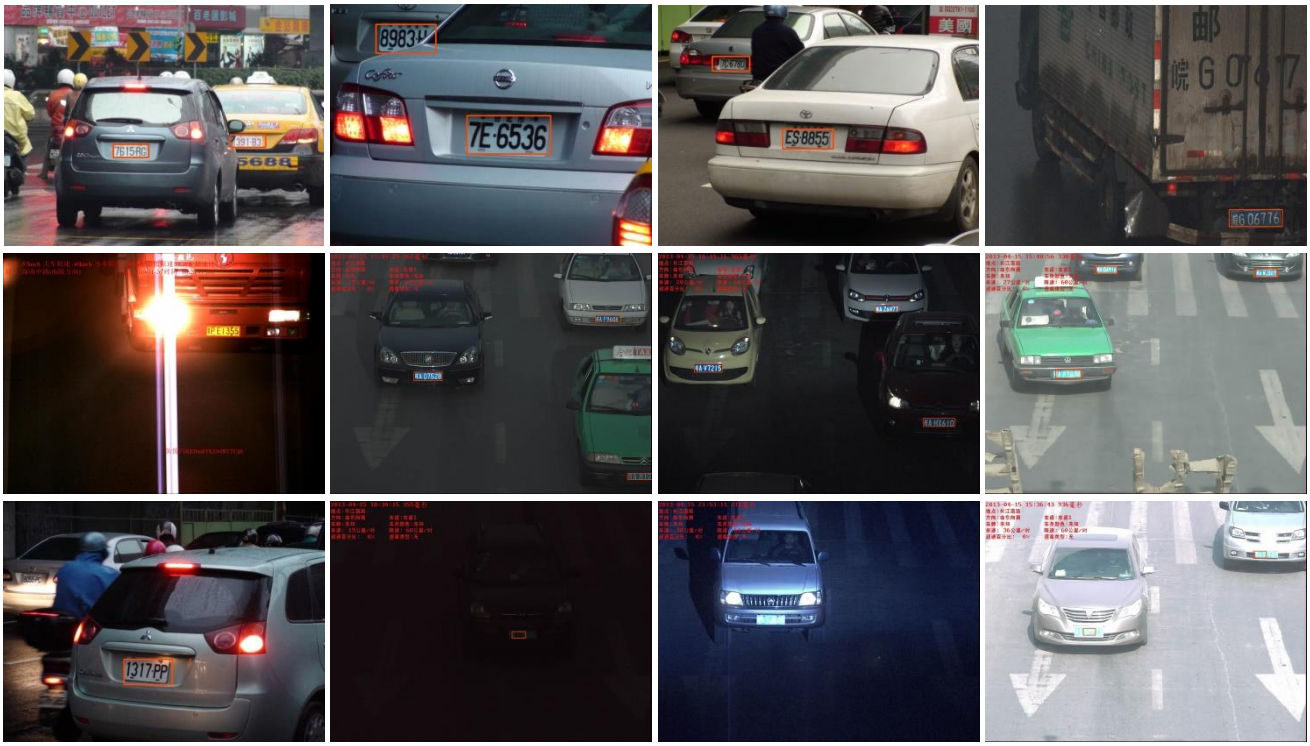


Fig. 12. Several results of plate detection in AOLP dataset and PKU dataset, success plate detection are shown in the first row and second row. Meanwhile, the last row shows fail plate detection in order to lousy illumination (very dark and very bright)

ACKNOWLEDGMENT

The authors would like to extend their gratitude to Directorate General of Higher Education, Ministry of Research and Education, Indonesia for funding this research.

REFERENCES

[1] Y. Yuan, L. Mou, and X. Lu, "Scene Recognition by Manifold Regularized Deep Learning Architecture," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 26, no. 10, pp. 2222–2233, 2015.

[2] T. Zhang, W. Zheng, Z. Cui, Y. Zong, J. Yan, and K. Yan, "A Deep Neural Network-Driven Feature Learning Method for Multi-view Facial Expression Recognition," *IEEE Trans. Multimed.*, vol. 18, no. 12, pp. 2528–2536, Dec. 2016.

[3] J. Ker, L. Wang, J. Rao, and T. Lim, "Deep Learning Applications in Medical Image Analysis," *IEEE Access*, vol. 6, pp. 9375–9379, 2017.

[4] H. Kong, N. Shigei, K. Mandai, S. Sugimoto, R. Takaesu, and Y. Ishizuka, "Land-Use Classification Using Convolutional Neural Network with Bagging and," in *Proceedings of the International MultiConference of Engineers and Computer Scientists 2019*, 2019.

[5] S. Rokrakthong, T. Suesut, and N. Tumrukwatthana, "Applying CNN to Infrared Thermography for Preventive Maintenance of Electrical Equipment," in *Proceedings of the International MultiConference of Engineers and Computer Scientists 2019*, 2019.

[6] L. Sumaryanti, A. Musdholifah, and S. Hartati, "Digital Image Based Identification of Rice Variety Using Image Processing and Neural Network," *Telkomnika Indones. J. Electr. Eng.*, vol. 16, no. 1, pp. 182–188, 2015.

[7] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions," in *Image and Vision Computing*, 2004, vol. 22, no. 10 SPEC. ISS., pp. 761–767.

[8] A. Śluzek, "Śluzek - 2016 - Improving performances of MSER features in matching and retrieval tasks.pdf," in *European Conference on Computer Vision*, 2016, pp. 759–770.

[9] A. M. Al-Ghaili, S. Mashohor, A. R. Ramli, and A. Ismail, "Vertical-edge-based car-license-plate detection method," *IEEE Trans. Veh. Technol.*, vol. 62, no. 1, pp. 26–38, 2013.

[10] A. D. X, "License Plate Localization by Sobel Vertical Edge Detection Method," *Int. J. Emerg. Technol. Eng. Res.*, vol. 4, no. 6, pp. 48–53, 2016.

[11] L. Xu, "A new method for license plate detection based on color and edge information of Lab space," *Proc. - 2011 Int. Conf. Multimed. Signal Process. C. 2011*, vol. 1, pp. 99–102, 2011.

[12] A. C. Roy, M. K. Hossen, and D. Nag, "License plate detection and character recognition system for commercial vehicles based on morphological approach and template matching," *2016 3rd Int. Conf. Electr. Eng. Inf. Commun. Technol.*, pp. 1–6, 2016.

[13] I. C. Tsai, J. C. Wu, J. W. Hsieh, and Y. S. Chen, "Recognition of vehicle license plates from a video sequence," *IAENG*

International Journal of Computer Science, vol. 36, no. 1, pp. 26–33, 2009.
[14] I. Ruslianto and A. Harjoko, "Pengenalan Karakter Plat Nomor Mobil Secara Real Time," *IJEIS*, vol. 1, no. 2, pp. 35–44, 2011.
[15] H. H. P. Wu, H. H. Chen, R. J. Wu, and D. F. Shen, "License plate extraction in low resolution video," *Proc. - Int. Conf. Pattern Recognit.*, vol. 1, pp. 824–827, 2006.
[16] T. Panchal, H. Patel, and A. Panchal, "License Plate Detection Using Harris Corner and Character Segmentation by Integrated Approach from an Image," *Procedia Comput. Sci.*, vol. 79, pp. 419–425, 2016.
[17] X.-C. Yin, X. Yin, K. Huang, and H.-W. Hao, "Robust Text Detection in Natural Scene Images," *{IEEE} Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 5, pp. 970–983, 2014.
[18] M. Nian *et al.*, "Method for unconstrained text detection in natural scene image," *IET Comput. Vis.*, vol. 11, no. 7, pp. 596–604, 2017.
[19] C. Yan, H. Xie, S. Liu, J. Yin, Y. Zhang, and Q. Dai, "Effective Uyghur Language Text Detection in Complex Background Images for Traffic Prompt Identification," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 1, pp. 220–229, 2018.
[20] G.-S. Hsu, J.-C. Chen, and Y.-Z. Chung, "Application-Oriented License Plate Recognition," *IEEE Trans. Veh. Technol.*, vol. 62, no. 2, pp. 552–561, 2013.
[21] B. Li, B. Tian, Y. Li, and D. Wen, "Component-based license plate detection using conditional random field model," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 4, pp. 1690–1699, 2013.
[22] Q. Gu, J. Yang, L. Kong, and G. Cui, "Multi-scaled license plate detection based on the label-moveable maximal MSER clique," *Opt. Rev.*, vol. 22, no. 4, pp. 669–678, 2015.
[23] J. Cheng, D. Zang, Z. Chai, D. Zhang, and J. Zhang, "Vehicle license plate recognition using visual attention model and deep learning," *J. Electron. Imaging*, vol. 24, no. 3, p. 033001, 2015.
[24] S. G. Kim, H. G. Jeon, and H. I. Koo, "Deep-learning-based license plate detection method using vehicle region extraction," *Electron. Lett.*, vol. 53, no. 15, pp. 1034–1036, 2017.
[25] H. Li, P. Wang, M. You, and C. Shen, "Reading car license plates using deep neural networks," *Image Vis. Comput.*, vol. 72, pp. 14–23, 2018.
[26] H. Li, P. Wang, and C. Shen, "Toward End-to-End Car License Plate Detection and Recognition With Deep Neural Networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 2, pp. 1–11, 2018.
[27] G. H. Park, H. H. Cho, and M. R. Choi, "A contrast enhancement method using dynamic range separate histogram equalization," *IEEE Trans. Consum. Electron.*, vol. 54, no. 4, pp. 1981–1987, 2008.
[28] D. Nistér and H. Stewénius, "Linear time maximally stable extremal regions," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 5303 LNCS, no. PART 2, pp. 183–196, 2008.
[29] D. Karatzas *et al.*, "ICDAR 2015 competition on Robust Reading," *Proc. Int. Conf. Doc. Anal. Recognition, ICDAR*, vol. 2015-Novem, pp. 1156–1160, 2015.
[30] Y. Yuan, W. Zou, Y. Zhao, X. Wang, X. Hu, and N. Komodakis, "A Robust and Efficient Approach to License Plate Detection," *IEEE Trans. Image Process.*, vol. 26, no. 3, pp. 1102–1114, 2017.
[31] W. Zhou, H. Li, Y. Lu, and Q. Tian, "Principal visual word discovery for automatic license plate detection," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 4269–4279, 2012.

Gamma Kosala received the bachelor degree in Electronics and Instrumentation from Universitas Gadjah Mada (UGM), Yogyakarta, Indonesia, in 2015. He is currently working toward the Doctoral degree in the Department of Computer Science and Electronics, Universitas Gadjah Mada, Yogyakarta, Indonesia. His research interests are image processing and artificial intelligence.

Agus Harjoko received the bachelor degree in degree in Electronics and Instrumentation from Universitas Gadjah Mada (UGM), Yogyakarta, Indonesia, in 1986. He received his M.Sc. and Ph.D. degree in Computer Science University from University of New Brunswick, Canada, in 1990 and 1996 respectively. He is currently an associate professor and head of the Department of Computer Science and Instrumentation, UGM, Yogyakarta, Indonesia. His research interest are computer vision, pattern recognition, instrumentation, and sensor network.

Sri Hartati received the bachelor degree in degree in Electronics and Instrumentation from Universitas Gadjah Mada (UGM), Yogyakarta, Indonesia, in 1986. She received his M.Sc. and Ph.D. degree in Computer Science University from University of New Brunswick, Canada, in 1990 and 1996 respectively. She is currently a professor and member of Intelligent System Lab in Department of Computer Science and Instrumentation, UGM, Yogyakarta, Indonesia. Her research interest are artificial & computational intelligence and decision support system.