Indonesian Traffic Sign Detection and Recognition Using Color and Texture Feature Extraction and SVM Classifier

Cahya Rahmad
State Polytechnic of Malang
Information Technology Department
Malang, Indonesia
cahya.rahmad@polinema.ac.id

Isna Fauzia Rahmah
State Polytechnic of Malang
Electrical Engineering Department
Malang, Indonesia
isna_fauzia@polinema.ac.id

Rosa Andrie Asmara
State Polytechnic of Malang
Information Technology Department
Malang, Indonesia
rosa.andrie@polinema.ac.id

Supriatna Adhisuwignjo
State Polytechnic of Malang
Electrical Engineering Department
Malang, Indonesia
supriatna@polinema.ac.id

Abstract—This paper presents traffic sign detection and recognition which is necessary to be developed to support several expert systems such as driver assistance and autonomous driving system. This study focused on the detection and recognition process tested on Indonesian traffic signs. There were some major issues on detecting process such as damaged signs, faded color, and natural condition. Therefore, this paper is proposed to address some of these issues and will be done in two main processes. The first one is traffic sign detection which divided into two steps. Start with segmenting image based on RGBN (Normalized RGB), then detects traffic signs by processing blobs that have been extracted by the previous process. The second process is traffic sign recognition process. In this process there are two steps to take. The first one is feature extraction, in this research we propose the combination of some feature extraction that is HOG, Gabor, LBP and use HSV color space. In next recognition stage some classifier are compared such as SVM, KNN, Random Forest, and Naïve Bayes. The propose method has been tasted on Indonesia local traffic sign. The results of the experimental work reveal that the approach of RGBN method showed precision and recall about 98,7% and 95,1% respectively in detecting traffic signs, and 100% for the precision and 86,7% for recall in recognizing process using SVM Classifier.

Keywords—traffic sign detection; RGBN; Gabor; HOG; LBP; SVM; classifier ;feature extraction; traffic sign recognition

I. Introduction

Traffic signs are one of the road equipment, in the form of symbols, letters, numbers, sentences and combinations such as warnings, prohibitions, orders or directions for road users [1]. By recognizing the traffic signs are expected the vehicle can provide information to the driver. Based on survey results in September 2010 found that 70% of people do not recognize traffic signs correctly [2]. This could have an impact on traffic safety. In addition, automotive manufacturers are also currently aggressively developing autonomous cars where detection and recognition of digital image has a very important role for driver assistance systems [3], and autonomous driving system [4].

In the previous research has been proposed the traffic sign recognition using gabor filter method as feature extraction, and KNN (K Nearest Neighbors) for the classification, with an accuracy of 79.69%. This research used template-matching method which has weakness in computation when the data that used as template is large, because the working principle of template matching is to match each testing image with an existing template [5].

Hence, this research proposed the detection and recognition of traffic signs using the RGBN [6], GABOR [7], HOG [8], LBP [9] and SVM [10] methods. This method is expected to improve the accuracy and speed of the detection process. This technology when developed will be able to recognize traffic signs that exist around the vehicle and provide information to the driver of the vehicle about the signs that are around. So it can help the driver to stay focused on the road, this can reduce the risk of accident and traffic violation [11].

II. Indonesian Traffic Sign Dataset

To quantitatively evaluate the performance of Traffic Sign Detection (TSD) and Traffic Sign Recognition (TSR) we create new dataset for Indonesian Traffic Sign.

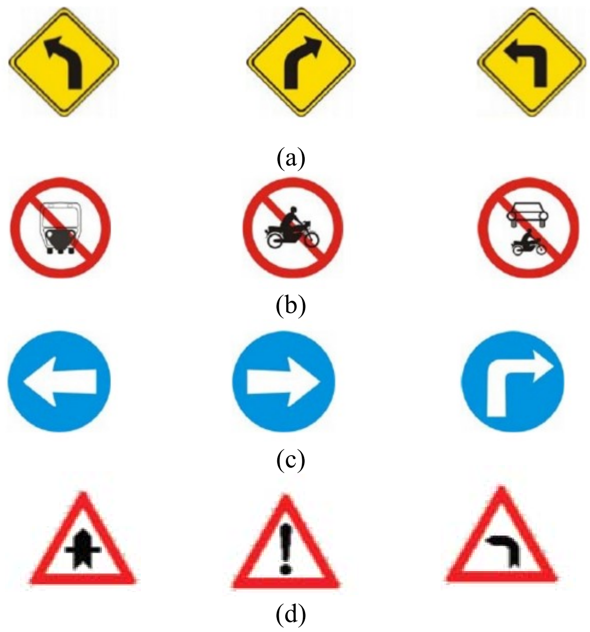

Fig. 1. (a) Indonesian Traffic Sign: Warning Sign. (b) Indonesian Traffic Sign: Prohibitive Sign. (c) Indonesian Traffic Sign: Mandatory Sign. (d) GTSDB: Warning Sign

Because the commonly used dataset of German Traffic Sign Recognition Benchmark (GTSRB) and German Traffic Sign Detection Benchmark (GTSDB) has different characteristics from Indonesian Traffic Sign dataset, especially the warning signs Fig. 1. (a) and (d). So, these two types of signs cannot be compared. Data test that used in this research is traffic signs that consist of warning signs, prohibitions, and mandatory signs. This traffic signs dataset obtained from the Department of Transportation Malang. It is taken in the afternoon and in normal condition using phone camera. Each type of this sign has different characteristics, whether different colors, shapes, or textures.

This dataset consists of 163 data training and 18 data test of TSD, this amount is based on 90% data training and 10% of data testing with details according to Table. 1 and there are traffic signs whose color is faded, covered with leaves, or partially damaged.

Table 1. The details of data training and data test of TSD

Types of Signs	Data Training	Data Test
Warning	97	11
Prohibitive	37	4
Mandatory	29	3
Total	163	18

As for the TSR consists of 150 datasets with size 60×60 pixels where from each type of signs will be selected 5 different sign and each will have 10 images of sign from random different perspectives, the details of TSR shown in Table 2

Table 2. The details of data training and data test of TSR

Types of Signs	Data Training	Data Test
Warning	50	5
Prohibitive	50	5
Mandatory	50	5
Total	150	15

III. TRAFIC SIGN DETECTION

The proposed method is achieved in three main steps. The first one is preprocessing the image using RGBN color space and ROI segmentation with the image threshold. The second step is to label the image and identify the candidate area. The last one is detecting traffic sign by analyzing each candidate region by using regionprops.

A. RGBN and Segmentation

The captured images are in the RGB color space. Therefore they are sensitive to environmental lighting conditions. To remove this effect, these images are converted to a normalized RGB color space (RGBN). The RGBN color space is formed independently of various lighting conditions or intensities. Therefore converting RGB images to RGBN will eliminate the effects of intensity variations. Three components of the RGBN color space can be obtained from the red, green and blue color components of the RGB color space, using the following formulation:

$$r = \frac{R}{R+G+B}$$

$$g = \frac{G}{R+G+B}$$

$$b = \frac{B}{R+G+B} \qquad (1)$$

r, g, b = Red Green Blue components of the RGBN image

R, G, B = Red Green Blue components of the RGB image

Table 3. Threshold values for RGBN

ThRed = 0.4	ThGreen = 0.3	ThBlue = 0.4

In this study normalized RGB imagery is used to measure the yellow, red, and blue colors of traffic signs. The threshold values for RGBN is shown in Table 3. Figure 2 is the original image before the thresholding process and Figure 3 is a binary image built after the Normalization RGB image threshold.

Fig. 2. The original image before convert to RGBN

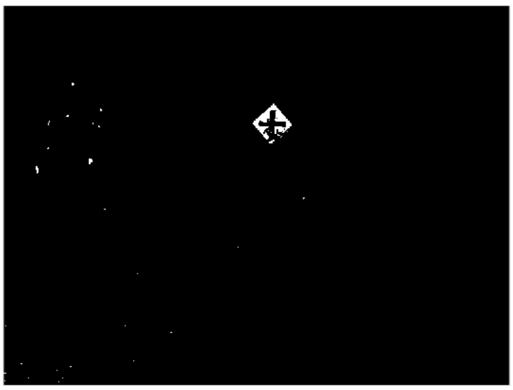

Fig. 3. Binary image after threshold the Normalized RGB image

B. *Labeling Image and Shape Classficication using regionprops()*

First it has to label the connected objects in the filtered binary image, to identify the candidate regions which may contain a traffic warning sign. Using the regionprops() method, area of the each labeled region was calculated. After calculating the area, each blob (ROI) is analyzed to find which one is the danger traffic sign. But there were lot of objects in the image, which cannot be a traffic sign. Analyzing those kinds of objects were useless and time consuming. Therefore after calculating the area of each blob, too small and too large regions were removed. Following formula (2) use to remove the unnecessary objects:

$$SpecialBlobs = \begin{Bmatrix} True, if\ blobsizes > 1000 \\ and\ blobsizes < 18000 \\ False, otherwise \end{Bmatrix} \quad (2)$$

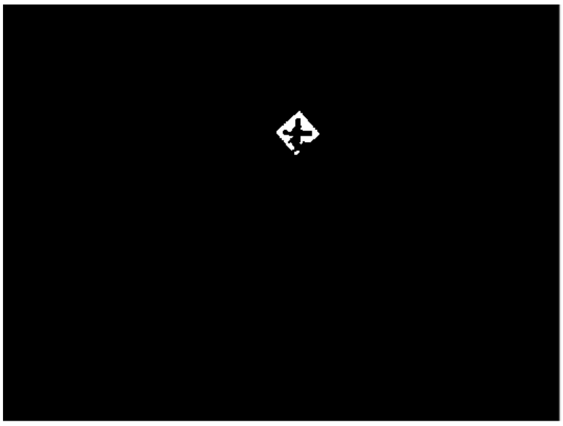

Fig. 4. Logical image, after removing the unnecessary objects

C. *Analyzing Each Candidate Region and Set Bounding Box*

When the image frame contains one or more candidate regions, which means specialBlobs1 > 0, then the program analyzed each candidate region using a "for" loop to identify whether there exist a traffic danger sign. And show the bounding box above the original image based on candidate region which is considered as a traffic sign.

Fig. 5. Detected traffic danger sign

IV. TRAFFIC SING RECOGNITION

At this stage of recognition there are two main steps.

A. *Feature Extraction*

The main idea behind the HOG method is to express the image as a group of local histograms. This histogram consists of gradient orientation and gradient magnitude. This technique calculates the gradient value in a particular area of an image. Each image has the characteristics indicated by the gradient distribution. This characteristic is obtained by dividing the image into a small area called a cell. Each cell is arranged a histogram of a gradient. The combination of this histogram is used as a descriptor that represents an object. Overall the HOG algorithm is shown in the figure 6. [12]

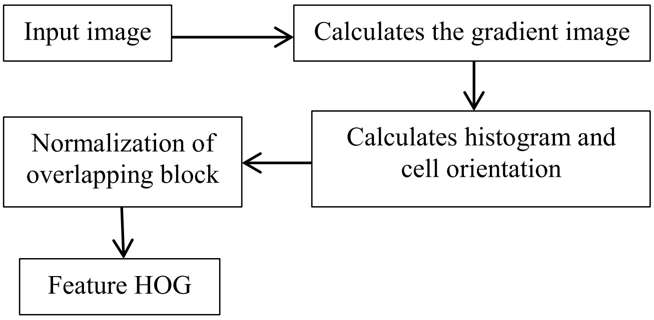

Fig. 6. Histogram of Oriented Gradient Algorithm

From Figure 6, the initial stage of HOG is calculating the gradient value of the input image. The most common method for calculating the magnitude of a gradient is to use a sobel filter in one or two directions, either horizontally or vertically. Next is to create parts called cells. Each pixel in a cell has its own histogram value based on the value generated in the gradient calculation. Cells have a size of 4×4 pixels in an image while the block has a size of 2×2 cell or 8×8 pixels.

Gabor filter is one filter that can simulate the characteristics of the human visual system in isolating the frequency and specific orientation of the image. This characteristic makes Gabor Filter appropriate for texture recognition applications in computer vision. Spatially, a Gabor function is a sinusoid that is modulated by the Gauss function. Here is a mathematical equation from gabor filter:

$$g(x,y) = \frac{1}{2\pi\sigma x\sigma y} \exp\left[-\frac{1}{2}\left(\frac{\tilde{x}^2}{\sigma x^2} + \frac{\tilde{y}^2}{\sigma y^2}\right)\right] \quad (3)$$

$$\tilde{x} = x cos\theta + y sin\theta \quad (4)$$

$$\tilde{y} = -x sin\theta + y cos\theta \quad (5)$$

θ = the Control of the orientation of the Gabor function
σ = standard deviation of Gaussian Envelope
x, y = the coordinates of Gabor Filter

LBP (Local Binary Pattern) algorithm is one of the algorithms that can be used to classify by image texture which has a very important place for computer vision. The example discussed this time is about converting the image into the image feature. LBP is not only a simple calculation, but also very effective for texture analysis. The most important of these methods is the nature of LBP is not affected by changes in light intensity.

This method works with calculate the difference intensity at the center point with 8 points of the neighbor. If the result of the difference is positive then it is given a value of 1 and if the result is negative it is given a value of 0. After that the result of that value is arranged clockwise and produces an 8-bit scale binary number. The resulting binary numbers are converted to decimal numbers. The set of decimal numbers will form a new histogram that characterizes each image. LBP computing can be denoted as follows:

$$LBP(xc + yc) = \sum_{n=0}^{7} s(i_n - i_c)\, 2^n \tag{6}$$

i_n = Gray level value of 8 points of the neighbor

i_c = Gray level value of center pixel $(xc + yc)$

B. *Classifier*

Support Vector Machine (SVM) is a technique to make predictions, both in case of classification and regression. SVM is in a class with the Artificial Neural Network (ANN) in terms of function and problem conditions that can be solved. The SVM concept can be explained simply as a search for the best hyperplane that serves as a separator of two classes in the input space. Figure 7 shows some patterns that are members of two classes: +1 and -1. Patterns belonging to class -1 are symbolized in red (square), while the pattern in class +1, symbolized by yellow (circle). Classification problems can be translated by finding the line (hyperplane) that separates into two groups. Various alternate lines of discrimination (discrimination boundaries) are shown in Figure 7. The best separator hyperplane between the two classes can be found by measuring the hyperplane's margins and look for the maximum point. Margin is the distance between the hyperplane premises with the nearest pattern of each class. The closest pattern is called a support vector. The solid line in Figure 4b shows the best hyperplane, which is located right in the middle of both classes, whereas the red and yellow dots that are in the black circle are the support vector. The effort to locate the hyperplane is at the heart of the learning process in SVM.

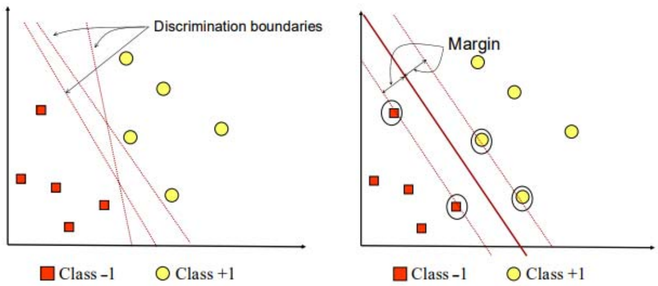

Fig. 7. The best dividing hyperplane in SVM

The available data is denoted as $\vec{x}_i \in R^d$ whereas each label is denoted $y_i \in \{-1, +1\}$ for $i = 1, 2, \ldots, l$ where l is the number of data. It is assumed that both classes -1 and +1 can be perfectly separated by the dimensioned hyperplane d, which is defined:

$$|\vec{w}.\vec{x}_i + b = 0| \tag{7}$$

Pattern $\vec{x}_i$ which belongs to class -1 (negative sample) can be formulated as a pattern that satisfies inequality

$$|\vec{w}.\vec{x}_i + b \leq 1| \tag{8}$$

Whereas pattern $\vec{x}_i$ which belongs to class +1 (positive sample)

$$|\vec{w}.\vec{x}_i + b \geq +1| \tag{9}$$

The greatest margin can be found by maximizing the distance value between the hyperplane and its nearest point, i.e. $1/||\vec{w}||$

V. Experimental Result

Traffic sign detections are considered to be true if the bounding box overlaps with at least 50% of the traffic signs. Here are the results of traffic sign detection:

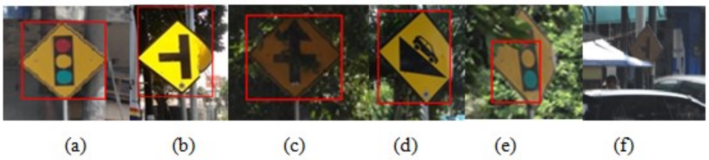

Fig. 8. Results of TSD

The traffic signs in Figure 8. (a), (b), (c) and (d) are considered successful because the bounding box area is precisely on the signs while Figure 8. (e) is considered successful because more than 50% of the bounding box area overlap with the area of signs, in this case traffic signs can be detected, but cannot be well recognized in the image recognition process (TSR), this case occurs because part of the leaf-covered sign. While figure 8 (f) is an unrecognizable dataset because of the distance that is too far away and the image of the signs that blend with the background.

The evaluation of the detection stage is performed based on precision–recall curve, where the recall and precision values are computed as follows:

$$recall = \frac{number\ of\ corectly\ detected\ sign}{number\ of\ true\ sign} \tag{10}$$

$$precision = \frac{number\ of\ corectly\ detected\ sign}{number\ of\ detected\ sign} \tag{11}$$

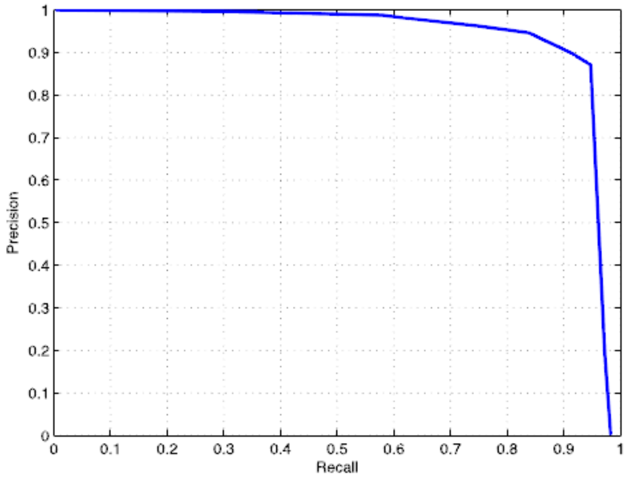

Fig. 9. Recall and Precision Curve

Table 4. Recall and precision values of TSD

Recall	Precision
95.1%	98.7%

The precision recall curve of the proposed method when applied to the data set is illustrated in Figure 8. The best trade-off between precision and recall values. It can be seen that the proposed method resulted in a 95.1% recall with a precision rate of 98.7% in the data set.

Table 5. Recall and precision values of TSR

Method	AUC	Precision	Recall
SVM	100	100	86,7
Random Forest	99,4	96,4	9,00
kNN	91,6	60,5	71,2
Naïve Bayes	93,3	96,4	86,7

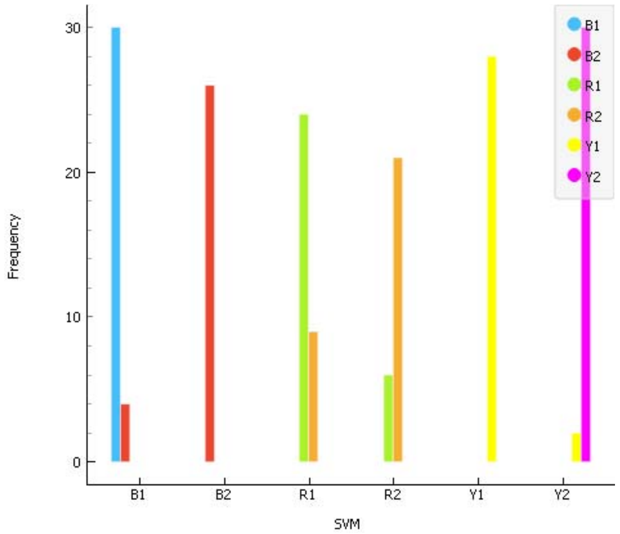

Fig. 10. The classification graph using SVM

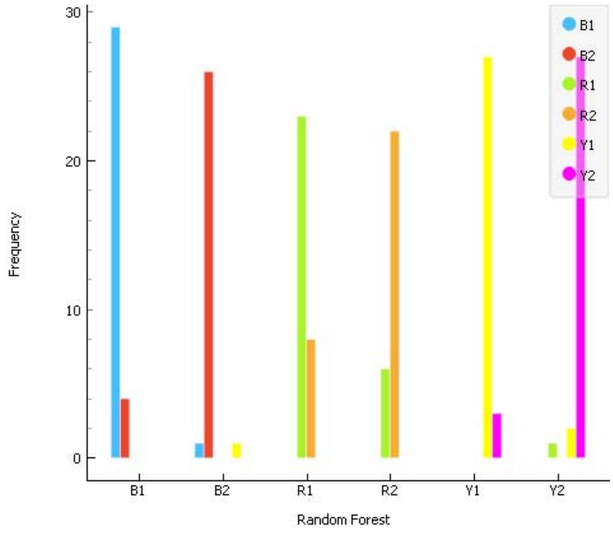

Fig. 11. The classification graph using Random Forest

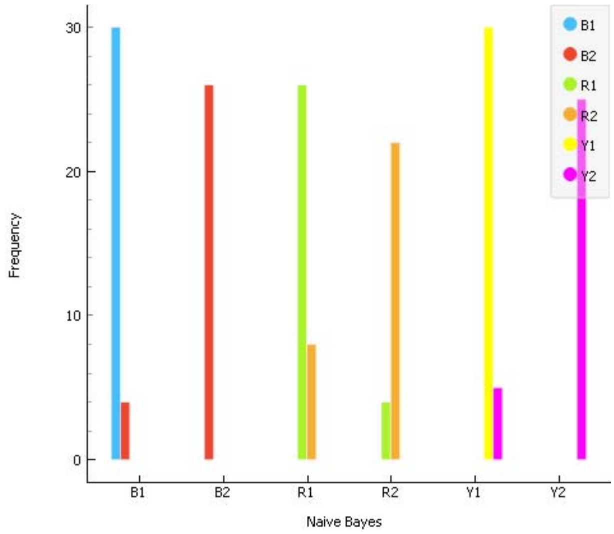

Fig. 12. The classification graph using Naïve Bayes

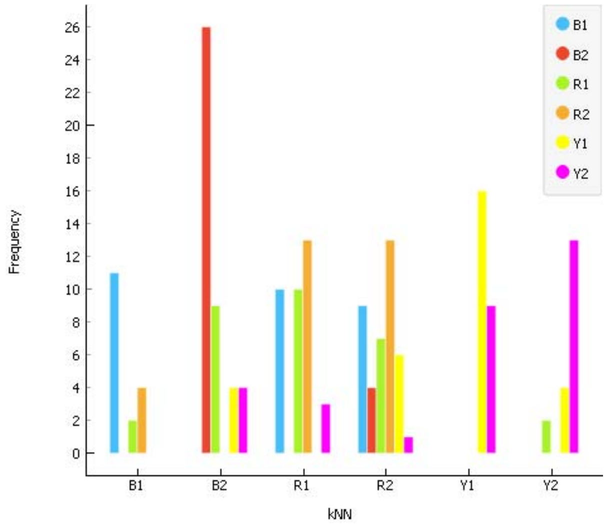

Fig. 13. The classification graph using kNN

Figure 10-13 showed graph from the result of some classification methods. The classification using SVM yields more true positive values than the other methods. Based on test result on dataset that showed in Table 5 with combined feature extraction (HOG + Gabor + LBP) it was found that the classification process using SVM has 100% AUC and has precision and recall 100% and 86,7%

VI. CONCLUSION

In this research, we have shown an efficient traffic sign detection system with strong feature extraction and representation which is a combination of HOG + Gabor + LBP. Which then tested on some classifier and got SVM has best result with precision and recall 100% and 86,7% respectively.

ACKNOWLEDGMENT

This work was supported by State Polytechnic of Malang.

REFERENCES

[1] Affandi, Faisal. 2004. Rambu Lalu Lintas Jalan Indonesia.

[2] Razi, Iqbal. (2010). Klasifikasi Kendaraan menggunakan Principal Component Analysis (PCA) dan Learning Vector Quantization (LVQ). Tugas Akhir Departemen Teknik Informatika. IT Telkom, Bandung.

[3] Timofte, R., Prisacariu, V.A., Van Gool, L.J., & Reid, I .InC. H. Chen(Ed.), "Com-bining traffic sign detection with 3d tracking towards better driver assistance", World Scientific Publishing, 2011.

[4] Levinson, J., Askeland, J., Becker, J., Dolson, J., Held, D.,Kammel, S., Kolter, J.Z., Langer, D.,Pink, O., Pratt,V., Sokolsky, M., Stanek, G., Stavens, D., Teichman ,A., Werling, M., &Thrun, S. (2011). Towards fully autonomous driving: systems and algorithms. In Proceedings of 2011 IEEE intelligent vehicles symposium (pp.163–168).

[5] D. M. Adhitama, I. I. Tritoasmoro, and Y. S. Hariyani, "Perancangan detektor nominal dan keaslian uang kertas dengan menggunakan template matching dan k-nearest neighbor (knn)," 2013, pp. 1–6.

[6] Gómez-Moreno, Hilario, et al. "Goal evaluation of segmentation algorithms for traffic sign recognition." IEEE Transactions on Intelligent Transportation Systems 11.4 (2010): 917-930.

[7] Liu,C.J., & Wechsler, H. (2002). Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition. IEEE Transactions on Image Processing, 11, 467–476.

[8] Xie, Y., Liu, L.F., Li, C.H., & Qu, Y. Y. (2009). Unifying visual saliency with HOG feature learning for traffic sign detection. In Proceedings of 2009 IEEE intelligent vehicles symposium: 1, 2 (pp.24–29).

[9] Samira, E., Sanae, B., Mounir, A.K., & Youssef,F. (2014). Traffic sign recognition based on multi-block LBP features using SVM with normalization. In Proceedings of 2014 9 th international conference on intelligent systems: theories and applications (Sita'14).

[10] Maldonado-Bascon, Saturnino, et al. "Road-sign detection and recognition based on support vector machines." IEEE transactions on

[11] Nugroho, Ignatius Agung. (2015). Pengenalan Secara Realtime Rambu Lalu Lintas Peringatan Menggunakan Ekstraksi Ciri Wavelet Haar dan Fungsi Jarak Chi Square.

[12] Dalal,N.,&Triggs,B.(2005). Histograms of oriented gradients for human detection. In Proceedings of 2005 IEEE computers ociety conference on computervision and patternrecognition:1 (pp.886–893).