A Solution for Automatic Counting and Differentiate Motorcycles and Modified Motorcycles in Remote Area

Indrabayu[1], Intan Sari Areni[2], Anugrayani Bustamin[3]
Elly Warni[4], Sofyan Tandungan[5], Rizka Irianty[6], Najiah Nurul Afifah[7]
Department of Informatics, Universitas Hasanuddin, Makassar, Indonesia[1, 3, 4, 6, 7]
Department of Electrical Engineering, Universitas Hasanuddin, Makassar, Indonesia[2]
College of Informatics Management and Computers STMIK AKBA, Makassar Indonesia[5]

Abstract—Motorcycles are the most significant contributor to the vehicle numbers in Indonesia, about 81% of all vehicles in the country. In addition, the growth of modified motorcycles has also increased in several areas, particularly remote places. Many studies have been conducted for detecting vehicles. However, most vehicle detection studies were conducted to detect cars or four-wheeled vehicles, and only a few studies were done to detect motorcycles. Further problems increase if the system is implemented in remote areas with limited electricity power resources that need low-cost budget specification computation. This study detects and calculates the number of motor vehicles and modified motorcycles passed on a highway from video data. It proposed Machine Learning instead of Deep Learning to suit the low computational video in remote areas. Computer vision-based methods used in the prediction are optical flow and Histogram Oriented Gradient (HOG) + Support Vector Machine (SVM). Five videos were used in the system testing, taken from the roadsides using a static camera with a resolution of 160x112 pixels at ±135º angle. This research showed that the accuracy of motorcycles and modified motorcycles detection and calculation systems using the HOG + SVM method is higher than the optical flow method. The average accuracy of HOG + SVM for motorcycles and modified motorcycles is 89.70% and 95.16%, respectively.

Keywords—Histogram of oriented gradient; optical flow; vehicles counting; support vector machine

I. INTRODUCTION

The government can observe the density level on the roads by utilizing cameras installed on the corners of the road. However, the utilization of the monitoring camera is still minimal. This technology does not differ from a machine that still requires other parties to operate to make it more useful [1]. Therefore, we need a system that can automatically detect moving vehicles and count vehicles through recorded video from monitoring cameras. However, several areas in southern Sulawesi where many motorcycle and modified motorcycles operate are in the remote area.

Systems that detect and count vehicles in traffic conditions can use active or passive sensors. This research will focus on systems with passive sensor technology because these sensors utilize computer vision for detection that is cheaper than active sensors [2]. Computer vision works by processing image data using combinations of image processing algorithms, artificial intelligence, pattern recognition, and computer graphics to produce information from the image [3].

Several algorithms or methods can be used to estimate the number and types of vehicles, including Optical Flow, Gaussian Mixture Models, Histogram Oriented Gradients (HOG), and Viola-Jones. In this study, the optical flow method is used in the image segmentation process in separating moving objects (vehicles or other moving objects) from stationary objects (roads or other fixed objects) by producing a motion vector that will be thresholded to distinguish objects. This method has been used in various fields, such as facial expression recognition [4], disease detection [5], virtual reality [6], object recognition [7], people counting [8], and gesture recognition [9]. Many previous researchers have used the optical flow in the vehicle recognition field. One of which is Sun et al. [10], who use it to detect and track vehicles in complex traffic conditions with shadow and occlusion. This system combines optical flow and immune particle filter, which increases tracking reliability to work well even in low visibility conditions.

Furthermore, optical flow is also used in [11] in combination with Convolutional Neural Network (CNN). The proposed method was evaluated under challenging environmental conditions. The experimental results showed 96.3% mean detection and 96.8% calculation precision.

In addition to optical flow, the HOG and Support Vector Machine (SVM) methods are also used to compare vehicle detection and counting system accuracy. These methods use the characteristics of the gradient distribution to describe the characteristic shape of an object to recognize the object. The results from HOG features are converted into feature vectors to be processed and trained in the SVM classifier. Finally, in the test stage, objects in the frame will be recognized by comparing the level of similarity of the gradient distribution trained to the gradient distribution in the test image. The combination of HOG and SVM methods has been used for various object detection. In [12], the HOG method extracted wood species and SVM to classify wood species. In [13], the researcher built an automatic mango detector system by combining SVM classifiers trained using HOG features and

image segmentation. Furthermore, HOG and SVM were also used in [14] to detect and classify airborne fungal spores.

Many studies have been conducted in vehicle detection [15],[16]. Some systems are even built to work under challenging conditions, such as dusty weather [17]. However, most vehicle detection studies were conducted on vehicles, and only a few studies were conducted to detect modified motorcycles. In addition, the growth of modified motorcycles or motorized tricycles is now increasingly out of control in some areas in Indonesia [18]. In this study, a comparative analysis of several computer vision algorithms for detecting motorcycles and modified motorcycles will be carried out. This study aims to accurately count the motorcycles and modified motorcycles and find the algorithm's influence on accuracy and processing time.

The structure of this paper is as. In Section 2, the background theory is described. The methodology is explained in Section 3. The result is discussed in Section 4. Finally, Section 5 concludes the paper and discusses future work.

II. BACKGROUND THEORY

A. Traffic Monitoring System

Traffic monitoring systems involve data collection for describing the characteristics of vehicles and their movements on the highway. Most vehicle counting and detection in traffic systems use sensor technology based on radar, microwave, tubes, and loop detectors. Sensors that reflect signals are called active sensors. The active sensor calculates the distance between the source and the target by measuring the time duration between emission and detection of the reflected signal.

On the other hand, sensors that do not reflect signals also can be used in vehicle detection on the highway. This type of sensor is called a passive sensor, an optical sensor that tends to be cheaper because it utilizes cameras and computer vision. This camera-based sensor can extract information more comprehensively, such as vehicle motion, shape, and color. The camera can track passing vehicles and their movement through complex and long-straight roads with precise positioning. A camera-based sensor can be successful if it can be operated in real-time [3].

B. Optical Flow

Optical flow is a visible movement caused by changes in brightness between two images. Optical flow occurs due to the relative movement between the observed object and the observer that can be seen in Fig. 1. The movement captured by optical flow is a movement in a two-dimensional plane [19].

The basic concept of optical flow is to see changes in the brightness of a point in two images. The brightness of a point will be compared with the surrounding area in the following image or frame. From the comparison results, the system can track an object using the brightness of the point that shows the position of the point in the following image or at a different time.
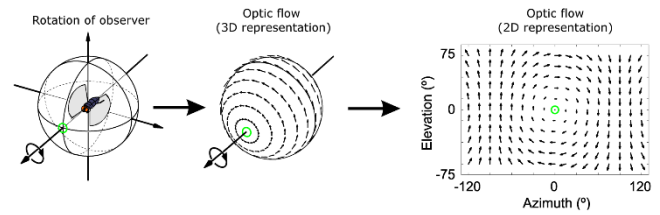

Fig. 1. Optical Flow [19].

Optical flow in computer vision is often used to mark and measure the movement of objects. By observing the intensity or brightness of two sequential images, information on the movement pattern of brightness in the image for each pixel can be obtained. If the pixel intensity value is located on *x, y* at time *t*, its value would be the same as the pixel located *(x+δx,y+ δy)* at time *t+ δt*.

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t) \qquad (1)$$

By applying the *Taylor series* to the right-hand side of eq. (1), eq. (2) can be obtained:

$$I(x + \delta x, y + \delta y, t + \delta t) = I(x, y, t) + \frac{\delta I}{\delta x}\delta x + \frac{\delta I}{\delta y}\delta y +$$
$$\frac{\delta I}{\delta t}\delta t + H.O.T \qquad (2)$$

By ignoring the higher-order terms (H.O.T.) after simplifying the notation, a simple form of optical flow is obtained, where u and v are flow vectors for each pixel, I_x and I_y are spatial gradients of brightness intensity, and I_t is the derivative of brightness intensity concerning time.

$$I_x u + I_y v + I_t = 0 \qquad (3)$$

C. Blob Analysis

Blob analysis is a technique used to express the pixel area of an image that becomes the focus of detection. This technique is used in optical flow to detect vehicles. The determination of the blob area for each object in the foreground segmentation process needs to be analyzed because the blob value for each object is different that influenced by object features such as size, type, and techniques in obtaining video data.

D. Histogram of Oriented Gradient (HOG)

Histogram of Oriented Gradient (HOG) is used in image processing for object detection purposes. This technique calculates the gradient value in a specific area of an image. Each image has characteristics indicated by a gradient distribution. These characteristics are obtained by dividing the image into small areas called cells. Each cell is composed of the histogram of gradients. Combining these histograms is used as a descriptor that represents an object.

HOG works using the window shift concept by calculating the gradient vector obtained for each window. The value and direction of the gradient vector in a particular area will display the characteristics of the gradient distribution of an image. Gradient distribution characteristics will describe the shape of an object in the image so that training can be carried out to recognize an object. Finally, objects in an image will be recognized by comparing the level of similarity of the gradient

distribution in the trained images to the gradient distribution in the target image. The results of the HOG feature are converted into feature vectors to be processed in the Support Vector Machine classifier.

E. Support Vector Machine

Supervised learning is a learning process where the training data has been labeled according to their respective class before the training begins. Furthermore, the system only checks the similarity of the features between the new incoming data and the training data, then labels them according to the most similar training data. One of the supervised learning methods is Support Vector Machines (SVM). The SVM algorithm aims to find a hyperplane that can separate classes with the maximum distance (margin/gap) between a particular class's borderlines (support vectors) and the borderlines of other classes. The basic idea of SVM itself is to find a linear decision surface (hyperplane) or barrier that separates one class from another with the most significant distance/gap/margin [6].

F. Tracking

Searching for moving objects in a sequence of frames is known as tracking. Tracking can be done by using object feature extraction and detecting moving objects/objects in the frame sequence. In computer vision, object tracking is a process that aims to follow the movement of an object. Furthermore, tracking also can be used for counting the detected vehicles.

III. METHODOLOGY

The block diagram of this research is depicted in Fig. 2. Following the input image, there is a preprocessing consisting of two stages. Subsequently, two detection algorithms will be reviewed in each model, namely Optical Flow and Histogram Oriented Gradients-Support Vector Machine (HOG-SVM). The results from the previous process will be used to calculate the number of vehicles that have been identified.

A. Input

Data acquisition was carried out on Jalan Adhyaksa Baru, Makassar, by placing a camera 4 meters above the ground using an iron pole, as shown in Fig. 3. The camera was placed on an iron pole at an angle of ±135º. This angle was used so that the passing vehicle could be seen or caught clearly by the camera. Fig. 3 illustrates the camera's position on the iron pole when collecting the data. The pole is attached to a power pole on the side of the road. Data is collected in the form of video. In this research, the data was taken from the rear corner of the vehicle where the recorded vehicles moved away from the camera.

This study designed two detection models, i.e., motorcycles and modified motorcycles. A Modified motorcycle is two-wheeled vehicles that have changed from their basic form intended to increase passenger capacity. A physical comparison of these two vehicles can be seen in Fig. 4.
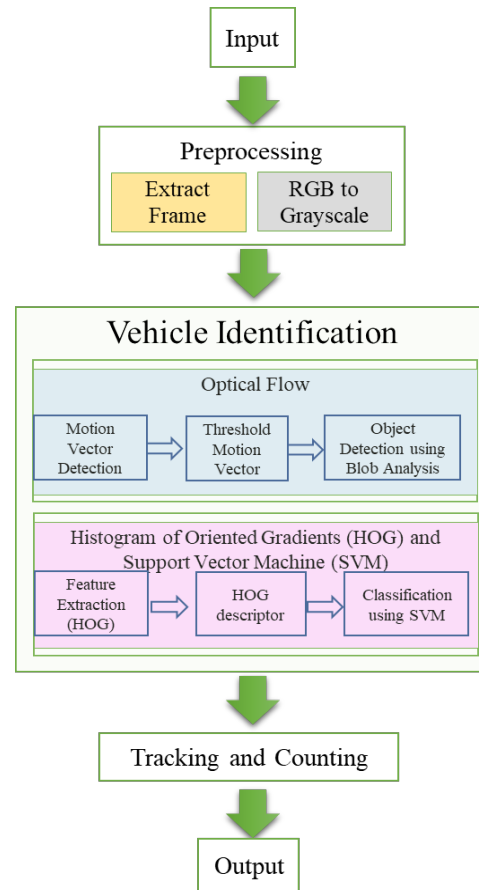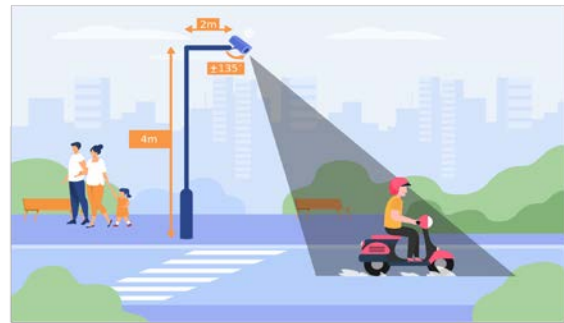

Fig. 2. Block Diagram of Vehicle Identification.

Fig. 3. Illustration of Data Acquisition Process.

(a)   (b)

Fig. 4. Sample of Dataset (a) Motorcycle and (b) Modified Motorcycle.

B. Preprocessing

After the input data from the video is obtained, the preprocessing stage will be carried out. Preprocessing is a stage for preparing the data before further processing to get maximum results. This stage will convert RGB frames to grayscale as an input for the next stage.

C. Vehicle Identification

This stage aims to identify the vehicle by implementing two different algorithms, which are optical flow and HOG + SVM classification.

1) Optical Flow method. This algorithm comprises three essential steps as follows:

a) Motion vector detection: This stage will detect vehicle movement by calculating the estimated optical flow. This process uses input in the form of grayscale frames. A yellow box around the moving object will indicate the detected motion vector.

b) Motion vector thresholding: Detected motion vectors will be subjected to a thresholding process to produce binary format frames, where foreground objects are labeled as one, and background objects are labeled as zero.

c) Objects detection using blob analysis: The thresholding process produces blobs of objects that the blob algorithm will analyze. The blob area size used in this study is 500-1000 pixels for motorcycles and 2100-2700 pixels for modified motorcycles.

Fig. 5 shows the steps of the identification process for motorcycles and modified motorcycle objects, which consist of the motion vector detection process, thresholding, and object detection with blob analysis.

2) HOG+SVM Method. This method starts by extracting features from the image that has been obtained in the previous process in preparation for the training phase. The HOG+SVM method comprises the following stages.
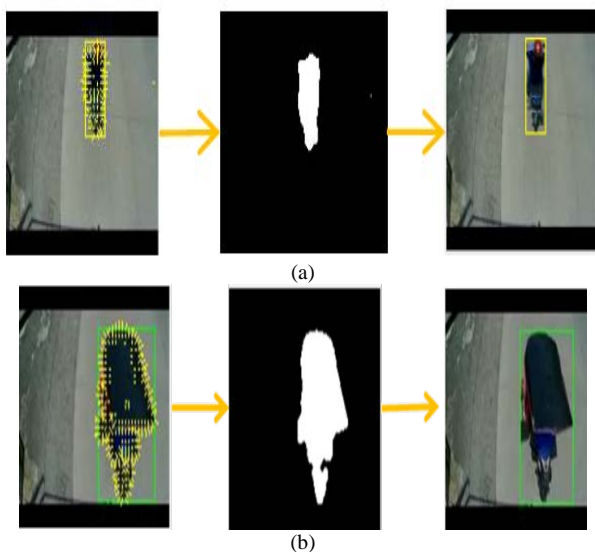
a) Training process: This stage aims to train the system for object classification (motorcycles and modified motorcycles) to generate a model used in the testing stage. This stage is carried out with several positive and negative training and negative data.

For the Motorcycle Detection Model, positive data consists of positive data, namely motorcycle images, and negative data, such as images of cars, roads, and pedestrians. At the training stage, the Modified Motorcycle Detection Model uses positive data in the form of modified motorcycle images and negative data in images of cars, roads, and other object images.

For the Motorcycle Detection Model, the training data includes 51 positive data and 300 negative data of motorcycles, and Modified Motorcycle Detection Model uses training data that consist of 100 positive data and 363 negative data of modified motorcycles.

After preparing each model's training data, the feature extraction process executes the HOG function. The output is the vector of the HOG descriptor, as shown in Fig. 6.

After the HOG extraction, classification was carried out using the SVM algorithm. This method classifies or separates motorcycles and modified motorcycles objects from other objects in the scenes. The vector obtained in the training process will be stored as a model and used for detection in the testing process.

b) Detection: A shift or sliding window is used to detect the feature descriptor in this stage. The sliding window will iterate 5 pixels per frame with a window size of 62x62 pixels for motorcycles detection and 71x71 pixels for modified motorcycles detection. HOG extraction is performed in each sliding window. The results of the HOG extraction will be classified using the SVM classifier generated at the training stage. The illustration of the detection process can be seen in Fig. 7.
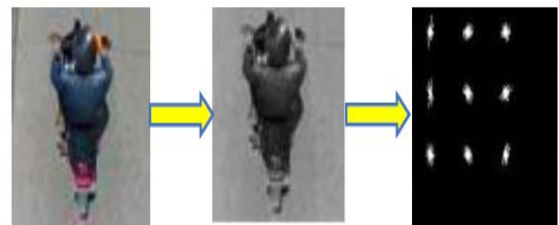

(a)

(b)

Fig. 5. Detection Process in Optical Flow (a) Motorcycle and (b) Modified Motorcycle.

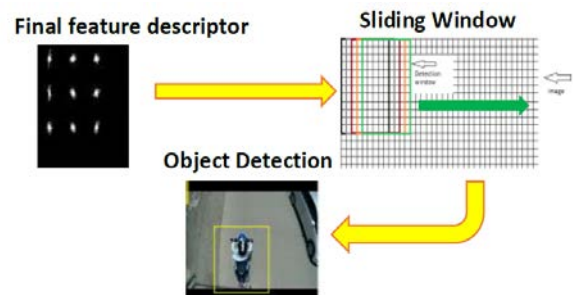Fig. 6. Illustration of the HOG Algorithm.

Fig. 7. HOG+SVM Method Object Detection Process.

D. Vehicle Tracking and Counting

The vehicle count will be based on the number of tracks generated by the tracking algorithm. This system uses the Kanade Lucas Tomasi Tracking (KLT) algorithm to count. The KLT tracking involves information from the previous frame, where the detected vehicle and the previous frame are input to this process. The output of this process is the vehicle's position in the current frame. The information of the current frame will be used as a tracking reference for the next frame. The tracking process will continue until the end of the frame. The KLT tracking is based on the feature values comparison in the frames based on a score of each bounding box.

E. Output

The output of this system is videos composed of a collection of detected frames with RGB data type. Each frame will show the number of detected vehicles in the upper left corner of the frame as shown in Fig. 8.
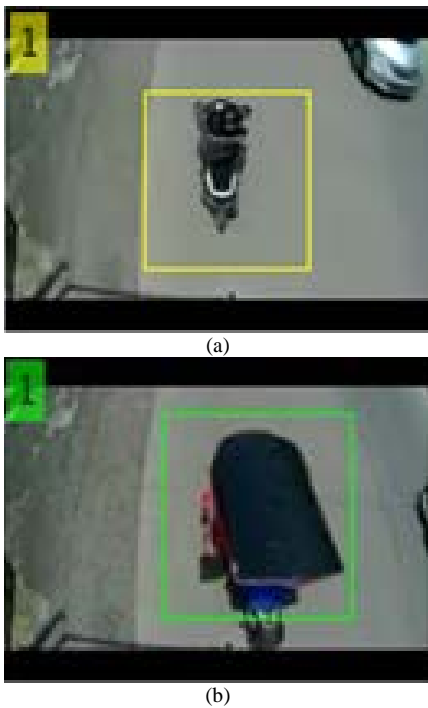

(a)

(b)

Fig. 8. Output Display System: (a). Motorcycle Detection , (b) Modified Motorcycle.

F. System Performance Analysis

The confusion matrix is used to measure the detection system performance, and the results are shown in Table I.

The system accuracy for each video is calculated using Eq. (4).

$$accuracy = \frac{tp + tn}{tp + tn + fp + fn} \qquad (4)$$

TABLE I. CONFUSION MATRIX OF THE DETECTION RESULTS

	Predicted	*Not predicted*
Positive Data	*True Positive* (*tp*)	*False Positive* (*fp*)
Negative Data	*False Negative* (*fn*)	*True Negative* (*tn*)

where:

- True Positive (tp), the motorcycles or modified motorcycles objects detected by the bbox
- True Negative (tn), not motorcycles object or modified motorcycles that is not detected by bbox.
- False Positive (fp), not motorcycles objects or modified motorcycles but detected by bbox.
- False Negative (fn), motorcycles object or modified motorcycles that is not detected by bbox.

IV. RESULTS AND DISCUSSION

The average results from both the optical flow and HOG+SVM methods are presented in Fig. 9. Five videos are used as the test data for detecting and counting the modified motorcycles and motorcycles, each of which is 1 minute long. The testing stage is carried out using the optical flow and HOG+SVM methods. The first test is accuracy analysis using the optical flow method. Table II shows the accuracy obtained using optical flow. The next test is the accuracy analysis using the HOG + SVM method. Based on (4), the measurement of system performance in this study focused on the TP value and considered the TN, FP, and FN values. The TP value for each video is relatively similar, but the FP value obtained tends to fluctuate, causing a decrease in accuracy. Table III shows the results of accurate measurement for the HOG + SVM method.
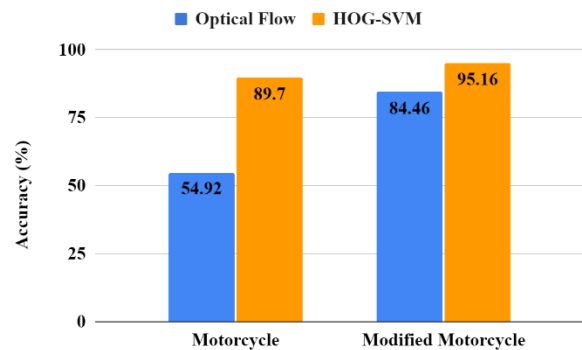

Fig. 9. Average Accuracy of the System.

TABLE II. SYSTEM ACCURACY USING OPTICAL FLOW

Model	Video	TP	TN	FP	FN	Accuracy (%)	Average Accuracy (%)
Motorcycle	1	20	3	13	0	63.89	54.92
	2	19	3	11	1	64.71	
	3	13	2	16	0	48.39	
	4	18	3	20	2	48.84	
	5	17	3	21	0	48.78	
Modified Motorcycle	1	2	33	5	0	87.50	84.46
	2	1	43	3	0	93.62	
	3	2	27	2	0	93.55	
	4	1	18	11	1	61.29	
	5	1	37	6	0	86.36	

TABLE III.   SYSTEM ACCURACY USING HOG-SVM

Model	Video	TP	TN	FP	FN	Accuracy (%)	Average Accuracy (%)
Motorcycle	1	21	11	1	3	88.89	89.70
	2	20	10	1	3	88.24	
	3	11	17	2	1	90.32	
	4	21	17	2	3	88.37	
	5	18	20	1	2	92.68	
Modified Motorcycle	1	1	38	0	1	97.50	95.16
	2	1	44	2	0	95.74	
	3	1	28	1	1	93.55	
	4	1	28	1	1	93.55	
	5	0	42	1	1	95.45	

Optical flow accuracy is 54.92% and 84.46% consecutively for motorcycles and modified motorcycles. In comparison, HOG+SVM performs 89.70% for motorcycles and 95.16% for modified motorcycles. The results show that HOG+SVM outperforms Optical flow in terms of accuracy.

To measure the performance of the system, data from the detection results for both motorcycle (in Fig. 10) and modified motor (in Fig. 11) are categorized according to the confusion matrix variable.
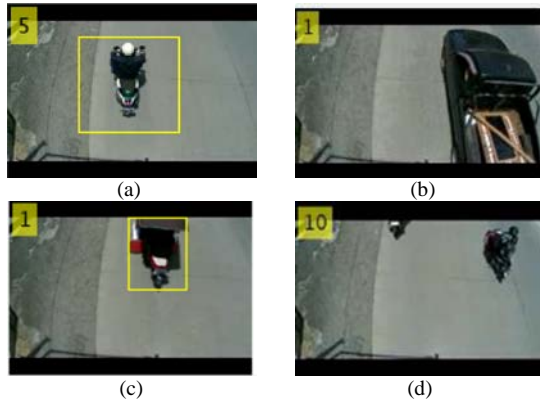

Fig. 10. Motorcycle Detection Results has Categorized in Confusion Matrix: (a) True Positive, (b) True Negative, (c) False Positive and (d) False Negative.
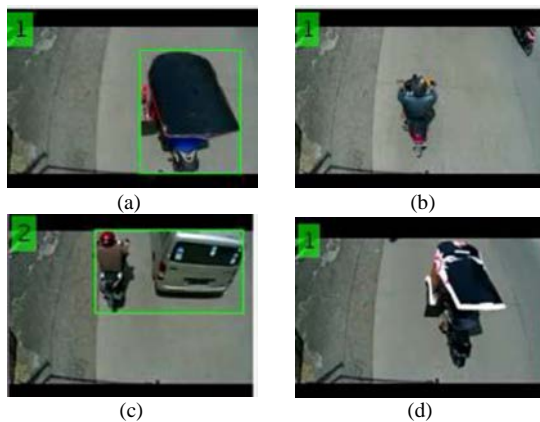

Fig. 11. Modified Motorcycle Detection Results has Categorized in Confusion Matrix: (a) True Positive, (b) True Negative, (c) False Positive and (d) False Negative.
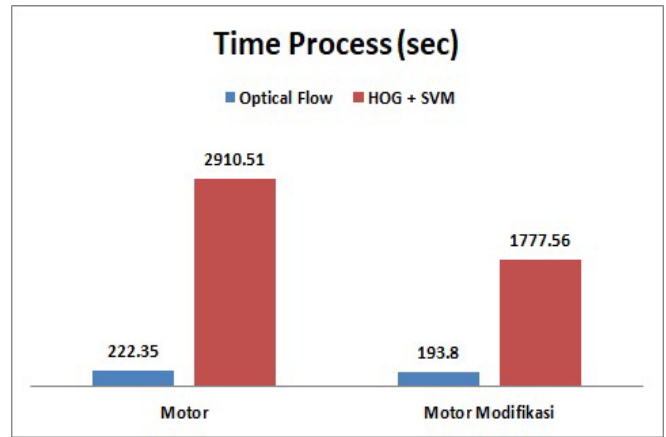

Fig. 12. Time Process of the Systems.

Fig. 12 shows the average system time, representing the processing time to detect and count the motorcycles and modified motorcycles in the video for each method. The accuracy measurement for the optical flow method is shown in the results of 54.92% accuracy for motorcycles and 84.46% for modified motorcycles. On the other hand, the HOG+SVM method produced average accuracy of 89.70% for motorcycles and 95.16% for modified motorcycles. Based on the results, the HOG+SVM method obtained higher accuracy than the optical flow method.

However, the optical flow method took less time to detect, which means the optical flow method is faster than HOG+SVM due to optical flow detected vehicles based on the blob area that had been determined in a frame, while the HOG+SVM method detected vehicles by checking the HOG feature through a window shift. HOG+SVM checked one window at a time on a frame. Thus, the HOG+SVM time process is slower than the optical flow.

V. CONCLUSION

This research aimed to design motorcycles and modified motorcycles classification and automatic counting systems through video data. The system test results showed that motorcycles and modified motorcycle detection accuracy using the HOG+SVM method was higher than the optical flow method. The average accuracy of HOG+SVM for motorcycles was 89.70%, and modified motorcycles was 95.16%, compared to the optical flow method with 54.92% accuracy for 84.46% for modified motorcycles. However, the computational time required for the optical flow method was faster than the HOG+SVM method, which were 222.35s for motorcycles and 193.80s for modified motorcycles. Meanwhile, HOG+SVM took 2910.51s and 1777.56s for motorcycles and modified motorcycles (*bentor* in Indonesia Language), respectively.

This research can be developed with variations of other vehicles in terms of categories and image specifications. Data retrieval in this study is taken in the daytime at a specific time only; hence further development for night conditions is needed before it can be used to support Intelligent Transport System technology in the real-world environment.

ACKNOWLEDGMENT

REFERENCES

[1] Jusman, "CCTV and Smart Governments," *The Fajar News,* March 14, 2016. [Online]. Available: http://makassarkota.go.id/berita-1105-cctv-dan-pemerintahan-yang-cerdas.html. [Accessed December 20, 2021] (In Indonesian).

[2] S. Jung, Y. Cho, D. Kim, and M. Chang, "Moving Object Detection from Moving Camera Image Sequences Using an Inertial Measurement Unit Sensor," *Applied Sciences*, vol. 10, no. 1, 268, 2020.

[3] V. Wiley and T. Lucas, "Computer Vision and Image Processing: A Paper Review," *International Journal of Artificial Intelligence Research*, vol. 2, no. 1, pp. 28-36, 2018.

[4] G. Patil and P. Suja, "Emotion recognition from 3D videos using optical flow method," in *2017 International Conference on Smart Technologies for Smart Nation (SmartTechCon)*, India, 2017, pp. 825-829.

[5] M. Abdel-Nasser, A. Moreno, H. A. Rashwan, D. Puig, "Analyzing the evolution of breast tumors through flow fields and strain tensors," *Pattern Recognit. Lett.,* vol. 93, pp. 162–171, 2017.

[6] G. Ren, W. Li, E. O'Neill, "Towards the design of effective freehand gestural interaction for interactive TV," *J. Intell. Fuzzy Syst.*, vol. 31, no. 5, pp. 2659–2674, 2016.

[7] S. Gujunoori and M. Oruganti, "Tracking and Size Estimation of Objects in Motion using Optical flow and K-means Clustering," in *2nd International Conference On Emerging Computation and Information Technologies (ICECIT)*, India, 2017, pp. 1-6.

[8] Tokta, Aybars and A. K. Hocaoglu, "A Fast People Counting Method Based on Optical Flow," in *2018 International Conference on Artificial Intelligence and Data Processing (IDAP)*, Turkey, 2018, pp. 1-4.

[9] M. Wrzalik and D. Krechel, "Human Action Recognition Using Optical Flow and Convolutional Neural Networks," in *16th IEEE International Conference on Machine Learning and Applications (ICMLA),* Mexico, 2017, pp. 801-805, doi: 10.1109/ICMLA.2017.00-59.

[10] Wei Sun, Min Sun, Xiaorui Zhang, and Mian Li, "Moving Vehicle Detection and Tracking Based on Optical Flow Method and Immune Particle Filter Under Complex Transportation Environments," *Complexity*, vol. 2020, 2020.

[11] A. Gomaa, M. M. Abdelwahab, M. Abo-Zahhad, T. Minematsu, and R. Taniguchi, "Robust Vehicle Detection and Counting Algorithm Employing a Convolution Neural Network and Optical Flow," *Sensors*, vol. 19, no. 20, 4588, 2019.

[12] B. Sugiarto et al., "Wood Identification Based on Histogram of Oriented Gradient (HOG) Feature and Support Vector Machine (SVM) Classifier," in *2017 2nd International conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, Indonesia, 2017, pp. 337-341, doi: 10.1109/ICITISEE.2017.8285523.

[13] M. J. C. Baculo and N. Marcos, "Automatic Mango Detection using Image Processing and HOG-SVM," in *Proceedings of the 2018 VII International Conference on Network, Communication and Computing (ICNCC 2018)*, Association for Computing Machinery, New York, 2018, pp. 211–215.

[14] M. W. Tahir, N. A. Zaidi, R. Blank, P. P. Vinayaka, M. J. Vellekoop, and W. Lang, "Detection of Fungus Through an Optical Sensor System Using The Histogram of Oriented gradients", in *IEEE Sensors 2016*, Florida, pp. 1-3, 2016.

[15] K.V. Sakhare, T. Tewari, V. Vyas, "Review of Vehicle Detection Systems in Advanced Driver Assistant Systems," *Arch. Computat. Methods Eng.,* vol. 27, pp. 591–610, 2020.

[16] A. F. Abbas, U. U. sheikh, F. T. Al-dhief, and M. N. H. Mohd, "A comprehensive review of vehicle detection using computer vision," *TELKOMNIKA*, vol. 19, no. 3, pp. 838-850, 2021.

[17] Nastaran Yaghoobi Ershadi and José Manuel Menéndez, "Vehicle Tracking and Counting System in Dusty Weather with Vibrating Camera Conditions," *Journal of Sensors*, vol. 2017, Article ID 3812301, 2017.

[18] F. M. Azzam, "Motorcycle Modification is Growing Uncontrollably, Makassar People Representative Council Asking for Government Intervention" *The PosKota*, May 3, 2021. [Online]. Available: https://sulsel.poskota.co.id/2021/05/03/bentor-menjamur-dprd-makassar-minta-pemkot-makassar-lakukan-evaluasi [Accessed December 20, 2021] (In Indonesian).

[19] S. Ranganatha and Y. P. Gowramma, "A Comprehensive Survey of Algorithms for Face Tracking in different Background Video Sequence," *International Journal of Computer Applications*, vol. 181, no. 27, pp. 43-49, 2018.