Paper_rudini.pdf

by

An Approach for Vehicle's Classification Using BRISK Feature Extraction

[1st]M. Rudini Kurniawan Amiruddin
Postgraduate Student of Electrical Engineering
Hasanuddin University
Makassar, Indonesia
amiruddinmrk19d@student.unhas.ac.id

[2nd]Indrabayu*
Department of Informatics
Hasanuddin University
Makassar, Indonesia
Corresponding Author: indrabayu@unhas.ac.id

[3rd]Intan Sari Areni
Department of Electrical Engineering
Hasanuddin University
Makassar, Indonesia
intan@unhas.ac.id

[4th]Anugrayani Bustamin
Department of Informatics
Hasanuddin University
Makassar, Indonesia
anugrayani@unhas.ac.id

Abstract **- This study aimed to classify vehicles according to their categories, consisting of motorcycles, light vehicles, and heavy vehicles. For this purpose, there were three main techniques discussed: vehicle detection using Background Subtraction, feature extraction using Binary Robust Invariant Scalable Keypoint (BRISK), and vehicle classification using the K-Nearest Neighbors (KNN) algorithm for most cases. The dataset consisted of 432 images for the training stage and one video data for the testing stage. The system performance was evaluated by reviewing the BRISK threshold value ranging from 10 to 80 with a *k*-value on KNN of 6. Results showed that the highest F1 scores were 96%, 86%, and 67% for motorcycles, light vehicles, and heavy vehicles, consecutively.**

Keywords—Intelligent Transportation System, Vehicle Classification, Background Subtraction, BRISK, KNN

I. INTRODUCTION

The growth of Indonesia's transportation sector is very high, but it is not accompanied by adequate infrastructure development. Based on Statistics Indonesia's data in 2015-2019, the number of vehicles has increased quite significantly per year by 6.13 percent [1]. Thus, often the accumulation of the number of vehicles causes traffic accidents. Therefore, Traffic Monitoring System (TMS) which leads to the development of the Intelligent Transportation System (ITS) is needed. TMS is used to collect traffic data, such as the number of vehicles, vehicle categories, and vehicle speeds, etc, which can be utilized to analyze future transportation needs and improve transportation safety [2]. The use of Computer Vision-based cameras is a solution to pro-existing sensor methods based on various ITS problems [3].

Fundamentally, Computer Vision works by utilizing processed video and image data to produce the required information. Several tasks in the ITS field use the Computer Vision system, i.e., vehicle detection, vehicle classification, vehicle counting, and vehicle tracking. Each vehicle has different characteristics so that an effective and efficient approach is required to identify vehicles based on shapes.

Research related to ITS utilizes several methods, one of which is Deep Neural Network for vehicle detection and classification [4]. This system acquired an accuracy of 90%; even so, this method required quite a lot of data, roughly 90.000 image data. Another technique to identify vehicles is using the Gaussian Mixture Model (GMM) method with a Virtual Detection Zone (VDZ). However, vehicles in this case, were classified based on their size and width. Hence, the approach taken for distinguishing vehicles of the same size was not effective [5]. The same technique is also used for vehicle classification which is done through the Blob Analysis method based on the foreground-background segmentation process. It should be noted that this study only classified vehicles in the form of motorbikes and cars [6].

Furthermore, the feature-based method is used to classify vehicle types through the Bag of SIFT method. The vehicle features were extracted using the Scale Invariant Feature Transform (SIFT) which was then grouped using the Bag of Visual Word (BOVW) concept. This work used vehicle front view data and was classified as a Support Vector Machine (SVM) which produced an accuracy of 89% [7]. Meanwhile, another approach for vehicle model recognition used the Binary Robust Invariant Scalable Keypoint (BRISK) method. Nevertheless, this approach only classified vehicles by comparing images with the same point of view against different backgrounds. The results obtained showed a system accuracy of 96.25% [8].

Several studies on vehicle detection and classification only used vehicle dimensions. Therefore, to identify vehicles by type, model, or category, it is necessary to take advantage of the unique characteristics of each vehicle to produce a high level of classification accuracy. This research aimed to build an ITS application for detecting and classify vehicles based on their categories with different approaches. The combination of detection methods and feature extraction is the main discussion. The function of the Background Subtraction and BRISK methods is to classify by utilizing the K-Nearest Neighbor (KNN) algorithm, as it is a fairly simple algorithm but with good performance [7, 8].

Accordingly, part II explains the proposed method in this paper, Part III describes the results and discussion of system testing, and part IV closes with the conclusions.

II. PROPOSED METHOD

This research uses video data format (.MP4) with a frame rate of 30 fps (frame per second) and a resolution of 1920 x 1080 pixels. Video data were collected from the pedestrian bridge on Jl. Perintis Kemerdekaan in Makassar City, South Sulawesi via Vivotec IP Camera. This collected video data was converted into images while vehicle objects were cropped in each image. The coordinate points on the front of the vehicle were manually marked using the Labeling tool to simplify the cropping process. Each of these coordinates was stored in an XML (Extensible Markup Language) file. The image will then be cropped based on the coordinates in the file. Afterward, each cropped vehicle image was grouped

978-1-6654-3398-3/21/PRINT ©2021 IEEE

according to its respective categories. The data processing process is as shown in Fig. 1.

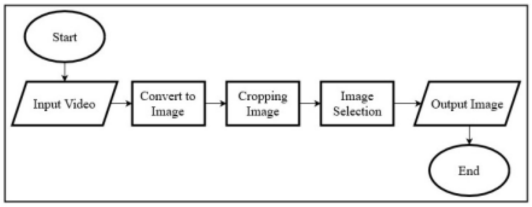


Fig. 1. Data Processing Flowchart

The data processing resulted in 432 images with three vehicle categories, namely motorcycle, light vehicle, and heavy vehicle, as shown in Fig. 2 (a), (b), and (c). The images were used as input data at the training stage. Meanwhile, 22 vehicle images from three categories obtain from video data were used in testing stage. Furthermore, this input data would be processed through several stages, as shown in Fig. 3.

Fig. 2. Data process of three vehicle categories: (a) Motorcycle, (b) Light Vehicle, and (c) Heavy Vehicle.

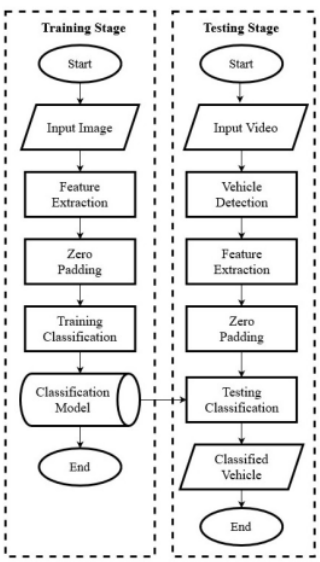


Fig. 3. Flowchart Training Stage and Testing Stage

The training and testing stages have a closely similar process. However, at the testing stage, the vehicle detection process is executed before the vehicle classification stage due to the input data originated in the form of a video. These stages are described as follows.

A. Vehicle Detection

In this section, the process of detecting vehicles on each frame by using the steps shown in Fig. 4. will be described.

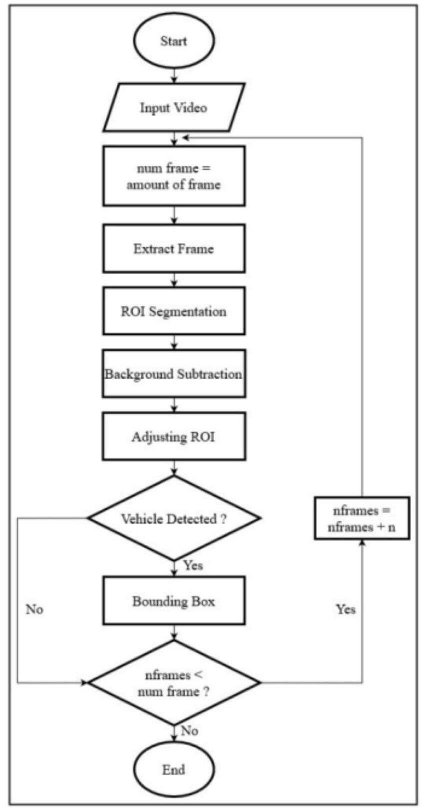


Fig. 4. Vehicle Detection Flowchart

1) Extract Frame

The first step is to process the video data into frames, followed by processing one by one until the last frame in the video.

2) ROI Segmentation

ROI Segmentation aims to limit the area to be processed to optimize vehicle detection at the focal detection point of each frame [11]. Therefore, it requires ROI Segmentation is required to determine the pixel coordinates that cover the non-focus detection area (Non-ROI). The specified pixel coordinate points are (706, 0), (1215, 0), (1917, 1079), (2, 1079). Fig. 5 shows the output segmentation of the ROI and Non-ROI area boundaries.

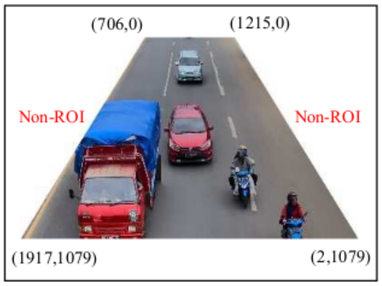


Fig. 5. Region of Interest Segmentation

3) *Background Subtraction*

Background Subtraction is a method used to separate the background from objects or commonly known as the foreground. This method detects motion effectively and efficiently by paying attention to moving objects in each frame [10, 11]. The approach taken is the Frame Difference technique by comparing the current frame ($frame_n$) with the previous frame ($frame_{n-1}$). First, the frame image is converted to grayscale. Then, the absolute difference in $frame_n$ is calculated to obtain the value for each pixel of the moving object as defined in equation (1).

$$f_n = |frame_n - frame_{n-1}| \tag{1}$$

Once the pixel value of the moving object is acquired, the binarization process is carried out by referring to the threshold value (*Th*) and the following equation (2).

$$b(x,y) = \begin{cases} 1, f_n(x,y) \geq T_h \\ 0, f_n(x,y) < T_h \end{cases} \tag{2}$$

The pixel value of one (1) is marked in white as the foreground area, while the pixel value of zero (0) is marked in black as the background area shown in Fig. 6 (a). After the foreground is segmented, the filter process is carried out to fill the contour holes in the foreground with the Morphology Operation process. The opening and closing operations are two techniques to produce a better foreground contour [14]. The output of the Morphology Operation filter is shown in Fig. 6 (b).

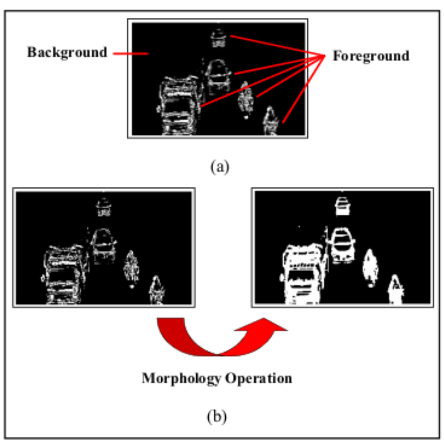


Fig. 6. (a) Foreground and Background, (b) Morphology Operation

4) *Adjusting ROI*

ROI adjustments are made so that process of detecting vehicles when crossing the ROI area runs more efficiently [15]. The blue line in Fig. 7 marks the ROI in the road area. Each vehicle that crosses the blue line will be detected according to the foreground contour area. This process is determined by the height and width of the three categories of vehicle objects: min-max height (240, 600) and min-max width (100, 600). The foreground that corresponds to the contour area will be detected and given a Bounding Box (green box), as shown in Fig. 7. Each vehicle that has been detected will be processed in the feature extraction section.

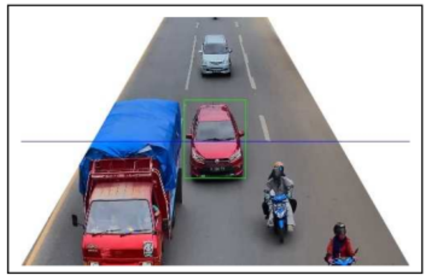


Fig. 7. The vehicle detected that given a Bounding Box

B. *Feature Extraction*

Classification of vehicle categories is performed based on the features of each vehicle image to be identified. Vehicle features are obtained from the feature extraction process using the BRISK method. This method is a more efficient alternative for extracting features in the form of a keypoint on an image compared to the SIFT and SURF methods [16].

In the BRISK algorithm, keypoint features can be detected by using the From Accelerated Segment Test (FAST) method. This method calculates the maximum value on the scale-space in determining the keypoint for each image based on the pattern in the area around interconnected keypoint by utilizing the Gaussian function to smooth the gradient (*g*) at each pixel point (*x, y*). The format pattern will generate two kinds of point pairs (*p*): the short-distance pair (*S*) and the long-distance pair (*L*). These point pairs are calculated based on equations (3-4).

$$S = \{(p_i, p_j) | \|p_i - p_j\| \delta_{max}\} \tag{3}$$

$$L = \{(p_i, p_j) | \|p_i - p_j\| \delta_{min}\} \tag{4}$$

Each keypoint in the image is described as a vector with a length of 64 bits. Hence, the features in each image will be a matrix with dimensions of *m* * 64, with *m* is the number of detected keypoints. Several parameters such as threshold value, octaves, and pattern scale affect the number of keypoints generated. This process requires a varying threshold value ranging from 10 to 80 to get optimal results. The lower the threshold value is used, the greater the number of keypoint features generated from the image, which in turn will affect processing time.

This process is done at the training stage with vehicle image input and the testing stage with vehicle images through the vehicle detection process on video input. The vehicle image is converted to grayscale before the feature extraction on the image by detecting the keypoint is executed. Each image will generate multiple keypoints with different descriptors. The average number of keypoint features generated was for 'motorcycle': 727, 'light vehicle': 2700, 'heavy vehicle': 5083. The illustration of the Feature Extraction process is shown in Fig.8.

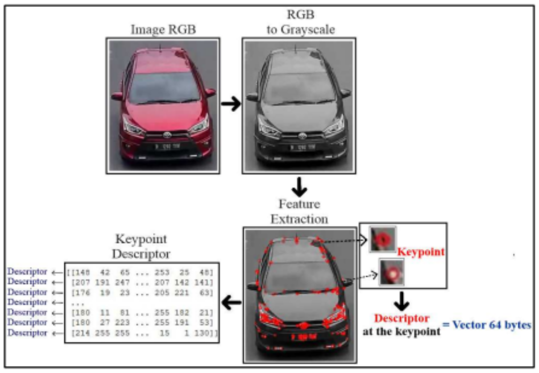


Fig. 8. Illustration of Feature Extraction

C. Zero Padding

Before entering the classification process, the keypoint descriptor matrix on each image is converted into a one-dimensional array (a vector) beforehand. Due to the difference in the number of descriptors in each image, the Zero Padding technique is used by referring to the maximum vector length. The number "0" is added at the end of the vector until the length of each vector is equal to the maximum length of the vector. This process is performed on all features (features in training and testing data).

D. Classification

The vehicle classification by their categories is the final step in this research. This process is carried out using KNN, a simple classification algorithm, by finding the closest distance between the testing data feature and the k value (neighbor) closest in the training data feature. This feature is a keypoint descriptor processed at the Zero Padding stage of each image vehicle.

The distance function is one of the parameters used to calculate the closest distance through the default distance function, namely Euclidean Distance (ED), according to equation (5).

$$ED = \sqrt{\sum_{i=1}^{k}(e_i - r_i)^2} \quad (5)$$

Where e_i is testing data, r_i is training data, and k is the Number of Neighbors.

In general, there are two stages in the classification process which include training and testing, which are shown in Fig. 2. The input to these processes is a keypoint descriptor that represents each feature of each vehicle. Next, the training stage is carried out to look for features patterns to distinguish each vehicle category label. The output of the training process is stored as a training model which will later be used for the testing stage using the video.

E. Performance parameter

In this study, the evaluation metric used is the F1 Score. This refers to the condition of imbalanced data, so that system performance can be measured better and informatively than classification accuracy, etc [17]. The metric is calculated based on the average harmonic mean of precision and recall, as shown in equations (6-8).

$$Precision = \frac{TP}{TP+FP} * 100\% \quad (6)$$

$$Recall = \frac{TP}{TP+FN} * 100\% \quad (7)$$

$$F1\ score = 2x\ \frac{Precision\ x\ Recall}{Precision+Recall} * 100\% \quad (8)$$

in which:

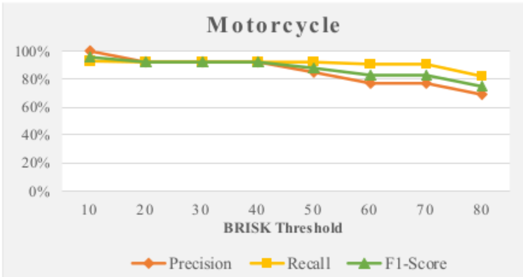
- True Positive (TP): positive data is correctly classified as the positive class.
- False Positive (FP): positive data is classified incorrectly and is in the negative class.
- True Negative (TN): negative data is classified correctly as negative class.
- False Negative (FN): negative data that is classified as false and is in a positive class.

Precision is obtained by comparing the TP results with the amount of data predicted to be positive, as shown in formula (5). Meanwhile, recall is obtained by comparing the TP results with the amount of data that is actually positive, as shown in formula (6). Finally, the values obtained from precision and recall will be averaged to obtain the F1 score, as shown in formula (7).

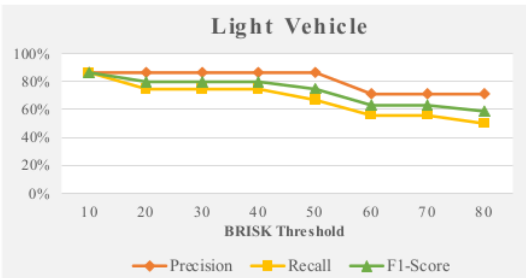
III. RESULTS AND DISCUSSION

This paper proposes an ITS application to classify vehicles according to their categories by combining the Background Subtraction method for vehicle detection and the BRISK for Feature Extraction method. This process is continued with the classification stage using the KNN. This process is continued with the classification stage using the KNN. The training process used 432 vehicle images, while the testing stage used one video, which contained 22 vehicle image data (13 motorcycles, 7 light vehicles, 2 heavy vehicles). The threshold for the BRISK method and the k value on the KNN are the test parameters in this study. The values of these parameters affect the number of features for each object. KNN will then classify new objects based on the majority of the values from the k-nearest neighbor category. The value of k thus needs to be determined first. In this study, the number of k = 6 was used to test the BRISK threshold value of 10 to 80. The three data classes in this study were motorcycle, light vehicle, and heavy vehicle. The determination of k value is done by considering the number of classes and dimensions of the data being tested.

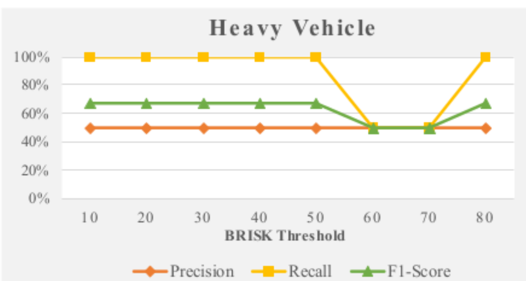
System performance is measured using the F1 score value obtained from precision and recall values with the evaluation metrics following imbalanced data conditions. The performance of the system is thus reviewed by looking at the F1 score of each class. Fig. 9 shows a graph of the evaluation of precision, recall, and F1score based on changes in the BRISK threshold values given for the three classes.



(a)



(b)



(c)

Fig. 9. The comparison of the evaluation metric results based on the threshold value for each class, (a) motorcycle, (b) light vehicle, and (c) heavy vehicle

The feature length in each image is influenced by the threshold value determined in BRISK from the observations. The optimal threshold value for three data classes in this study is 10. The highest F1 score value of 96% is obtained in the class of motorcycle which can be classified well. However, this did not occur in the other two classes, namely heavy vehicles and light vehicles, which respectively had F1 scores of 67% and 86%. Misclassification in these two classes occurs because there is one image of a heavy vehicle classified as a light vehicle and one light vehicle image classified as a motorcycle. Besides, the imbalance of data for the three classes also significantly affects system performance where only two data tests for heavy vehicles. The successes and failures of the system in classifying vehicles are shown in Fig. 10-13.

Fig. 10. A light vehicle that successfully classified

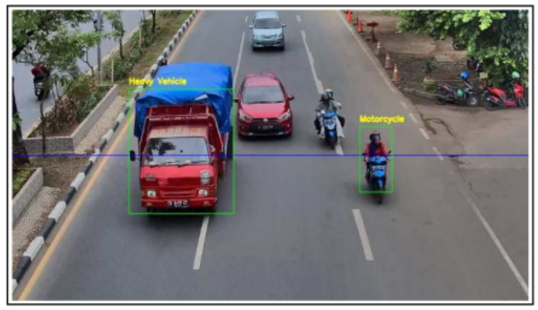


Fig. 11. A heavy vehicle and motorcycle that successfully classified

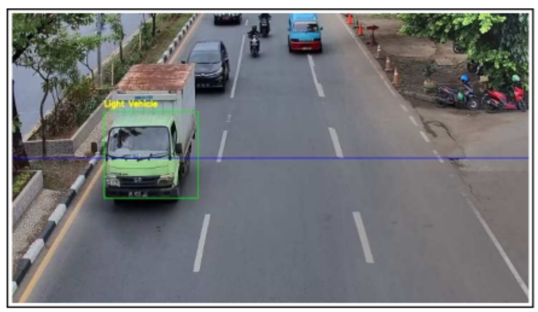


Fig. 12. A heavy vehicle classified as a light vehicle

Fig. 13. A light vehicle classified as a motorcycle

IV. CONCLUSION

This paper has proposed a detection and classification system for vehicle categories. Background Subtraction and BRISK methods are combined to detect and extract features in the vehicle images. This stage is then continued with classification using the KNN method. The threshold value of BRISK of 10 is used for the three classes, and the $k = 6$ value for KNN. The test results of the vehicle detection system for the classification of vehicle categories resulted in F1 score from the highest to the lowest, namely 96% for the motorcycle class, 86% for the light vehicle class, and 67% for the heavy vehicle class. The development system can identify vehicles better, especially in the motorcycle category. However, the data imbalance becomes the shortage of this study, due to real conditions that occur on the road.

In the future, this research can be improved by performing feature selection so that the resulting feature dimensions are simpler. Thus, a more efficient computational process will be obtained without compromising the classification results. In addition, more data should be added to avoid cases of imbalanced data.

ACKNOWLEDGMENT

This research is supported by LPPM – Hasanuddin University (UNHAS) via PDUPT Ristekdikti Grant 2021 and *Artificial Intelligence Multimedia Processing* (AIMP) Research Group.

REFERENCES

[1] "Statistic Land Transportation," Indonesia: BPS-Statistics Indonesia, 2019, pp. 1–73. [In Bahasa Indonesia]

[2] M. Won, S. Sahu, and K. J. Park, "DeepWiTraffic: Low cost WiFi-based traffic monitoring system using deep learning," *Proc. - 2019 IEEE 16th Int. Conf. Mob. Ad Hoc Smart Syst. MASS 2019*, pp. 476–484, 2019, doi: 10.1109/MASS.2019.00062.

[3] M. A. Butt *et al.*, "Convolutional Neural Network Based Vehicle Classification in Adverse Illuminous Conditions for Intelligent Transportation Systems," *Complexity*, vol. 2021, 2021, doi: 10.1155/2021/6644861.

[4] C. C. Tsai, C. K. Tseng, H. C. Tang, and J. I. Guo, "Vehicle Detection and Classification based on Deep Neural Network for Intelligent Transportation Applications," in *2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, APSIPA ASC 2018 - Proceedings*, 2019, no. November, pp. 1605–1608, doi: 10.23919/APSIPA.2018.8659542.

[5] N. Seenouvong, U. Watchareeruetai, C. Nuthong, K. Khongsomboon, and N. Ohnishi, "Vehicle detection and classification system based on virtual detection zone," 2016, doi: 10.1109/JCSSE.2016.7748886.

[6] Indrabayu, Basri, A. Achmad, I. Nurtanio, and F. Mayasari, "Blob Modification in Counting Vehicles using Gaussian Mixture Models under Heavy traffic," in *ARPN Journal of Engineering and Applied Sciences*, 2015, vol. 10, no. 16, pp. 7157–7163.

[7] M. A. Manzoor and Y. Morgan, "Vehicle Make and Model classification system using bag of SIFT features," 2017, doi: 10.1109/CCWC.2017.7868475.

[8] M. Huzaifa and I. S. Suwardi, "Car Model Recognition from Frontal Image using BRISK," in *International Conference on Electrical Engineering and Computer Science (ICECOS)*, 2017, pp. 104–108, doi: 10.1109/ICECOS.2017.8167114.

[9] F. A. I. Achyunda Putra, F. Utaminingrum, and W. F. Mahmudy, "HOG Feature Extraction and KNN Classification for Detecting Vehicle in The Highway," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 14, no. 3, p. 231, 2020, doi: 10.22146/ijccs.54050.

[10] Y. Yu, J. Wang, J. Lu, Y. Xie, and Z. Nie, "Vehicle logo recognition based on overlapping enhanced patterns of oriented edge magnitudes," in *Computers and Electrical Engineering*, 2018, vol. 71, no. June, pp. 273–283, doi: 10.1016/j.compeleceng.2018.07.045.

[11] S. Han and N. Vasconcelos, "Object-based regions of interest for image compression," in *Data Compression Conference Proceedings*, no. May, pp. 132–141.

[12] R. Roopa Chandrika, N. S. Gowri Ganesh, A. Mummoorthy, and K. M. Karthick Raghunath, "Vehicle Detection and Classification using Image processing," 2019, doi: 10.1109/ICESE46178.2019.9194678.

[13] F. C. Cheng, B. H. Chen, and S. C. Huang, "A hybrid background subtraction method with background and foreground candidates detection," *ACM Trans. Intell. Syst. Technol.*, vol. 7, no. 1, 2015, doi: 10.1145/2746409.

[14] B. Pawar, V. T. Humbe, and L. Kundnani, "Morphology based moving vehicle detection," in *Proceedings of the 2017 International Conference On Big Data Analytics and Computational Intelligence, ICBDACI 2017*, 2017, pp. 217–223, doi: 10.1109/ICBDACI.2017.8070837.

[15] Basri, Indrabayu, and A. Achmad, "Gaussian Mixture Models optimization for counting the numbers of vehicle by adjusting the Region of Interest under heavy traffic condition," in *International Seminar on Intelligent Technology and Its Applications, Proceeding (ISITIA)*, 2015, pp. 245–249, doi: 10.1109/ISITIA.2015.7219986.

[16] R. Y. S. Leutenegger Stefan, Margarita Chli, "'BRISK: Binary Robust Invariant Scalable Keypoints.' Computer Vision (ICCV)," in *International Conference on Computer Vision*, 2011, pp. 2548–2555, doi: 10.1109/ICCV.2011.6126542.

[17] D. Zhang, J. Wang, and X. Zhao, "Estimating the uncertainty of average F1 scores," *ICTIR 2015 - Proc. 2015 ACM SIGIR Int. Conf. Theory Inf. Retr.*, pp. 317–320, 2015, doi: 10.1145/2808194.2809488.

Paper_rudini.pdf

ORIGINALITY REPORT

7%	3%	7%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Apoorva Ojha, Satya Prakash Sahu, Deepak Kumar Dewangan. "Vehicle Detection through Instance Segmentation using Mask R-CNN for Intelligent Vehicle System", 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), 2021 Publication	1%
2	Indrabayu, Rizki Yusliana Bakti, Intan Sari Areni, A. Ais Prayogi. "Vehicle detection and tracking using Gaussian Mixture Model and Kalman Filter", 2016 International Conference on Computational Intelligence and Cybernetics, 2016 Publication	1%
3	repository.unhas.ac.id Internet Source	1%
4	Shigeru Kashihara, Muh. Arief Wicaksono, Doudou Fall, Muhammad Niswar. "Supportive Information to Find Victims from Aerial Video in Search and Rescue Operation", 2019 IEEE	1%

International Conference on Internet of Things and Intelligence System (IoTaIS), 2019
Publication

5	Submitted to American University of Beirut Student Paper	1%
6	Sitti Wetenriajeng Sidehabi, Ansar Suyuti, Intan Sari Areni, Ingrid Nurtanio. "Classification on passion fruit's ripeness using K-means clustering and artificial neural network", 2018 International Conference on Information and Communications Technology (ICOIACT), 2018 Publication	<1%
7	export.arxiv.org Internet Source	<1%
8	Indrabayu Indrabayu, Rahmat Hardian Putra, Ingrid Nurtanio, Intan Sari Areni, Anugrayani Bustamin. "Blob adaptation through frames analysis for dynamic fire detection", Bulletin of Electrical Engineering and Informatics, 2020 Publication	<1%
9	Annisa Nurul Puteri, Dewiani, Zulkifli Tahir. "Comparison of Potential Telemarketing Customers Predictions with a Data Mining Approach using the MLPNN and RBFNN Methods", 2019 International Conference on Information and Communications Technology (ICOIACT), 2019	<1%

Publication

10	Nilakorn Seenouvong, Ukrit Watchareeruetai, Chaiwat Nuthong, Khamphong Khongsomboon, Noboru Ohnishi. "Vehicle detection and classification system based on virtual detection zone", 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE), 2016 Publication	<1 %
11	Awad, Mariette, and Rahul Khanna. "Support Vector Machines for Classification", Efficient Learning Machines, 2015. Publication	<1 %
12	Malisa Huzaifa, Iping Supriana Suwardi. "Car model recognition from frontal image using BRISK", 2017 International Conference on Electrical Engineering and Computer Science (ICECOS), 2017 Publication	<1 %
13	"Intelligent Systems and Applications", Springer Science and Business Media LLC, 2019 Publication	<1 %
14	Ming Qin, Yao Lu, Huijun Di, Wei Huang. "A Background Basis Selection-Based Foreground Detection Method", IEEE Transactions on Multimedia, 2016	<1 %

Publication

15	Moch Arief Soeleman, Irfanul Haq, Muslih, W Karis, P R. Anggi. "The Use of Gaussian Mixture Model for Counting Human Object on Video", Journal of Physics: Conference Series, 2021 Publication	<1 %
16	sci.ldubgd.edu.ua Internet Source	<1 %
17	Myounggyu Won. "Intelligent Traffic Monitoring Systems for Vehicle Classification: A Survey", IEEE Access, 2020 Publication	<1 %

Exclude quotes	On	Exclude matches	< 5 words
Exclude bibliography	On		