Performance Evaluation of Feature Extraction Algorithms for Vehicle Shape Classification

Sunitha Patel[1], Srinath S[2]

[1]Department of Computer Science and Engineering, ATME College of Engineering, Mysore, Karnataka, India (mssunithapatel@gmail.com) ORCID 0000-0001-6676-8678; [2]Department of Computer Science and Engineering, SJCE College of Engineering, Mysore, Karnataka, India (srinath@sjce.ac.in) ORCID 0000-0003-4771-2614

Abstract

Vehicle classification is a classic application of automotive image processing that is necessary for a variety of modern vehicle safety and comfort features. While machine learning-based solutions are effective in these fields and are currently employed extensively in various automotive applications. However, the most challenging aspect of automotive image processing with machine learning methods is gathering adequate quality and quantity datasets to develop such applications. Additionally, imbalanced datasets are common in multiclass automotive image processing, as is the case with the current topic of vehicle platoon management. The effectiveness of available handcrafted feature extractors and classifiers employed for vehicle class categorization varies greatly due to the effect of dataset imbalance. This study aims to examine how the performance of four prominent feature extractors alters when used with the imbalance dataset for vehicle shape classification. Also, the use of image augmentation techniques to increase the dataset size for three vehicle classes: car, bus, and truck, has been presented. Further, using the Support Vector Machine (SVM) classifier, experimental analysis was performed using feature extractors such as Histogram of Oriented Gradient (HOG), Scaled Invariant Feature Transform (SIFT), Speeded-Up Robust Feature (SURF), and Haar. Vehicle shape classification, which is an important characteristic in vehicle platoon management, has been evaluated using Receiver Operating Characteristic (ROC) for both the unbalanced dataset and the augmented dataset. The experimental results demonstrate that using the HOG feature extractor performs better when compared to SIFT, SURF, and HAAR feature extractors on the imbalance dataset. After using an image augmentation technique to add images, output performance improved significantly, with HOG output of 95%, SIFT output of 91%, SURF output of 91%, and HAAR output of 96%.

Author Keywords. Feature Extractor. SVM. HOG. SIFT. SURF. HAAR. Vehicle Platoon.

Type: Research Article

1. Introduction

For its many vehicle safety and comfort features, the automotive industry heavily relies on vision-based vehicle detection and classification systems. Vehicle shapes can be classified using a variety of methods, including feature-based, model-based, and machine learning-based approaches. In the work of Chen, Pan, and Zou (2016), a vision-based method was utilized to perform lane-line recognition in conjunction with vehicle detection using a feature-based strategy that uses an edge feature and vertical projection to estimate vehicles in the image. Li and Wang (2015) developed a new hybrid image template for describing vehicle features through the use of texture and color. Zakria et al. (2018) worked on a model-based approach and proposed a new probabilistic framework for vehicle detection.

Chen et al. (2018) proposed a vehicle detection model for complex traffic video surveillance. Penghua and Yanping (2019) proposed an improved feature extractor using an SVM classifier. Machine learning-based vehicle shape classification has gained popularity in recent years due to its simplicity and increased accuracy. SVM is one of the popular machine learning algorithms used for image processing-based classification problems.

Local Binary Pattern (LBP) (Wen et al. 2015), Gist (Sikirić, Brkić, and Šegvić 2013), HOG (Lee et al. 2015), SIFT (Chen and Meng 2013), SURF (Song et al. 2019), HAAR (Wei et al. 2019) are some of the prominent handcrafted feature extraction techniques. According to the literature review, feature extractors like HOG, SIFT, SURF, and Haar have been used in a variety of vehicle applications and have produced notable results for various scale changes. The current work focuses on using popular handcrafted feature extractors to do machine learning-based vehicle shape categorization. Feature extractors that are invariant to image augmentation approaches have been chosen for the current study. In addition, the current effort aims to improve the dataset for bus and truck images by incorporating multiple image augmentation techniques such as rotation, crop, blur, and tilt.

The rest of the article is organized as follows: Section 2 provides an overview of the image processing problem statement for vehicle platoon management. Section 3 describes the algorithm's implementation approach, Section 4 offers simulation results for the given problem using various feature extraction techniques, Section 5 summarizes and discusses the results; and Section 6 wraps up the study with concluding remarks.

2. Problem Statement

Vehicle platoon management entails the coordination of many vehicles driving in the same longitudinal direction, as shown in Figure 1. Vehicles in platooning systems typically feature at least one front camera for Advanced Driver Assistance Systems (ADAS) and communicate with one another using advanced telematics systems. In platoon management, the lead vehicle will direct the other followers' vehicles to reap the benefits of platooning. Reduced fuel or energy consumption, reduced carbon dioxide (CO_2) emissions, reduced traffic congestion, and reduced accidents are all important advantages of vehicle platoon management. Various researchers have demonstrated the potential advantages of vehicle platooning in their work. Liu et al. (2019) presented a new platoon technique that maximizes fuel economy during lane changes. Ballarin and Zeilinger (2017) showed how platoon management benefits society and customers by lowering fuel consumption, CO_2 emissions, traffic congestion, and road accidents in their study. Peng, Bonsangue, and Xu (2019) suggest a novel timed automata model for safe vehicle platooning.

Each following vehicle will transmit the image of its front vehicle to the lead vehicle through a front vehicle camera. As illustrated in Figure 1, after a new merger, the platoon consists of four vehicles: V1 is the lead vehicle, while V2, V3, and V4 are follower vehicles. Follower vehicles V2, V3, and V4 shall capture front images of their respective front vehicles and transmit them to lead vehicle V1 via vehicle-to-vehicle communication. It is assumed in the current work that all vehicles travel at a constant distance d. The lead vehicle V1 is responsible for identifying the various types of vehicles in platooning and arranging them in a certain order based on vehicle area. To optimize fuel efficiency, however, it is important to group vehicles according to their shape before arranging them in a prescribed manner.
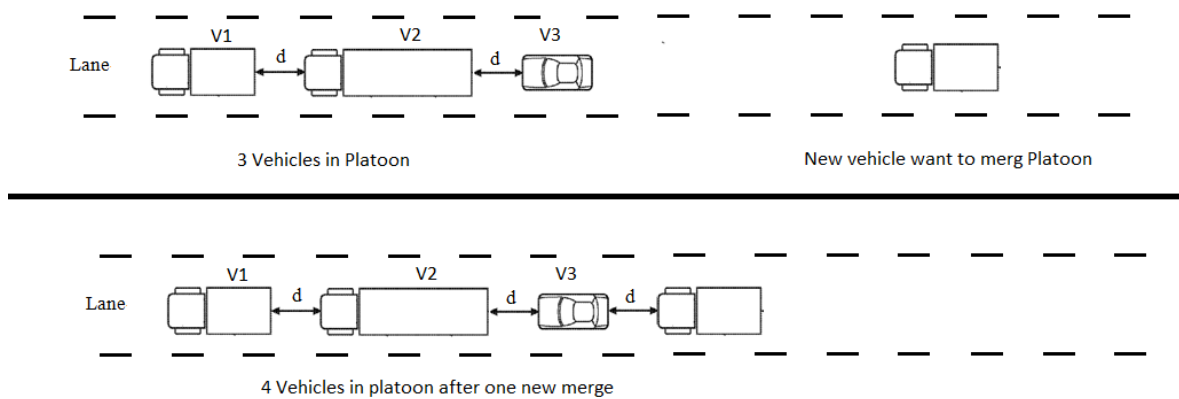
Figure 1: Vehicle platoon management example

Numerous studies have been conducted to date utilizing various feature extraction algorithms as a pre-processing step for vehicle form detection in automotive image processing using single-class and multiclass machine learning classifiers. Nayak, Gopalswamy, and Rathinam (2019) conducted a study employing various features and different classifiers on a variety of vehicles, demonstrating that HOG with Adaboost delivers a 40% more accurate and precise classification result. In Amraoui et al. (2019), the accuracy of SURF has been compared to other extractors; it is 100 percent when compared to other bio-inspired approaches; however, the authors proposed enhancing the average duration of reaction before offering the bio-inspired approach. Kapoor et al. (2007) showed that utilizing SVM and SIFT, the Gaussian Process with Pyramid Match Kernal may be used to recognize a wide variety of target vehicles. Wei et al. (2019) demonstrated and discussed the importance of performing research in multi-vehicle detection settings, as well as the advantages of combining HOG and Haar feature extraction with an SVM classifier. Bougharriou, Hamdaoui, and Mtibaa (2018) demonstrated that linear SVM with HOG feature benefits advanced vehicle driver assistance systems by improving the performance of their proposed algorithms. Lin et al. (2012) researched blind spot identification using a front and sidecar model with SURF as the feature extractor and probabilistic as the classifier. In Sun, Bebis, and Miller (2006), the HOG and Gabour feature extractors were compared to the SVM classifier and demonstrated their superiority.

The feature extraction approach in image processing frequently results in the loss of the original image features, according to the literature review. Furthermore, each feature extractor has its own set of pros and downsides; therefore, a comprehensive investigation of the specific application should be undertaken before selecting a feature extractor. The goal of this study is to learn more about the performance of four popular feature extractors for vehicle shape detection, which will be employed in vehicle platoon management. Furthermore, the study demonstrates how image augmentation may be used to improve the quality of image datasets, hence boosting the performance of machine learning systems.

3. Algorithm Implementation

The experimental work has been carried out using Python 3.7.9, with IDE PyCharm and OpenCV Library. PyCharm is used as it is the dedicated Integrated Development Environment (IDE) for Python, which provides essential tools for the development of machine learning and data science-based applications; OpenCV (Open Source Computer Vision Library) is an open-source software library for computer vision, machine learning, and image processing.

The implementation flow of the proposed algorithm is depicted in Figure 2, with the first stage being the creation of the required image dataset, followed by filtering and training with multiple feature extractors. Following the extraction of features depending on the region of

interest, the vehicles are classified using an SVM classifier before being arranged in a specified manner for platoon management.
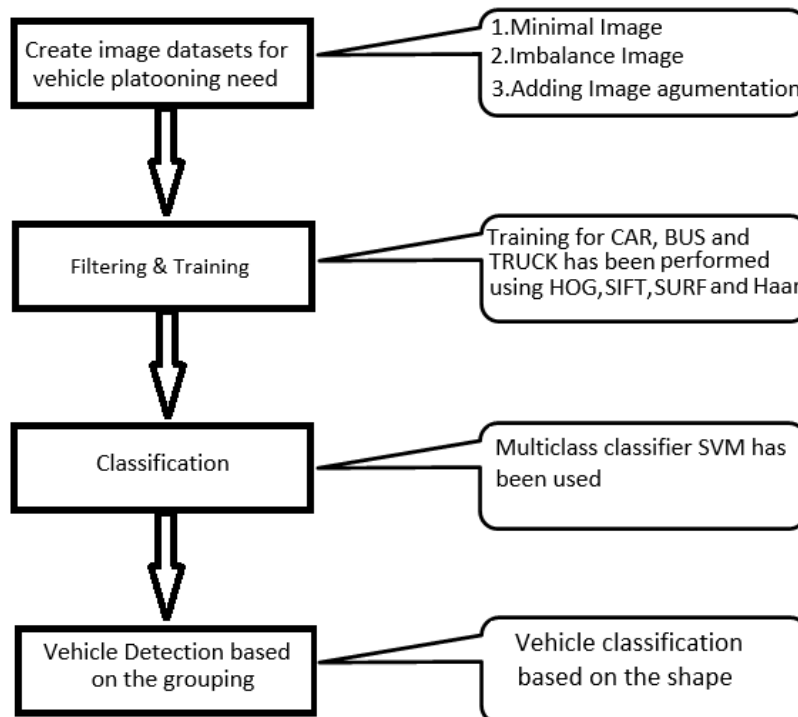

Figure 2: Algorithm flow

3.1. Vehicle dataset creation

A brief survey was undertaken to ascertain the availability of datasets for the stated topic, as identifying or developing data sets for the specified topic is a vital step in achieving the best results for any classification challenges involving machine learning algorithms. Zakria et al. (2018) have developed the datasets of rear-view car images and a few vehicle datasets have been made open source. Additionally, the datasets mentioned in these papers have also been made public (Chang et al. 2018; Lou et al. 2019; Naphade et al. 2021). According to the results of the survey, open repository datasets may not contain all of the images required to conduct an experimental study on a certain issue. As a result, to expand the dataset size for the current problem study, data augmentation techniques were used.

Currently, the effort is being made to recognize certain vehicles based on their size and organize them suitably so that the platoon's lead vehicle can influence the rearrangement of other vehicles to improve fuel or energy economy, which benefits consumers and society. For the proposed problem statement, the vehicles investigated are cars, trucks, and buses, and the associated image collection was used for the preliminary study. The images have been gathered and created, as mentioned in Figure 3. The images were taken from various open-source repositories and 5000 images of car types were obtained from these open datasets (Krause et al. 2013). However, there were about 3500 images available for Truck (Tabassum et al. 2020), while nearly 2900 images were available for Bus (Tabassum et al. 2020). Initially, an analysis was carried out using the available images, with minimal image quantities of 2900 images for each car, bus, and truck. Additionally, studies were carried out utilizing an unbalanced image dataset acquired from an open-source repository (Arróspide, Salgado, and Nieto 2012) which included 5000, 2900, and 3500 images of each car, bus, and truck, respectively. The dataset size was increased for the final analysis, with an equal number of

images for each group, resulting in 5000 images for each car, truck, and bus vehicle type. Table 1 represents the number of images considered for analysis in three separate scenarios.

Scenarios	Number of Bus images	Number of Car images	Number of Truck images
Minimal Image Dataset	2900	2900	2900
Imbalance Image Dataset	2900	5000	3500
Added Augmented Image Dataset for Bus and Truck	5000	5000	5000

Table 1: Datasets Scenarios

3.1.1. Image augmentation

Researchers can use image augmentation techniques to create more diverse images in less time without sacrificing quality, such as obtaining a variety of images for different vehicles in various weather scenarios, illuminations, orientation, etc. is time-consuming in the majority of cases. Image augmentation is a technique in which the quality and number of available images can be boosted by the use of augmentation techniques such as cropping, flipping, resizing, adjusting the brightness, adding noise, etc. The image augmentation technique is a boon for researchers, mainly because machine learning algorithms call for diversified images for a better learning process. Shorten and Khoshgoftaar (2019) conducted a survey on the different image augmentation techniques that can be utilized for problems being solved using deep learning. Lemley, Bazrafkan, and Corcoran (2017) suggested a technique for smart data augmentation in which the proposed network learns to generate enhanced data during the learning process. Because there were fewer images accessible for buses and trucks than for cars in the current study, the number of images for buses and trucks was enhanced through the use of image augmentation techniques. An effort has been made to capture photos of approximately 400 quantities each for different types of buses and trucks on the road. By utilizing different image augmentation techniques, as shown in Figure 3, quantities of captured images have increased. Even though the image augmentation techniques greatly aid in increasing the data size, one should be careful when applying the required image augmentation technique for the given problem because sometimes wrong augmentation techniques can degrade the quality of the dataset, which in turn degrades learning capability.
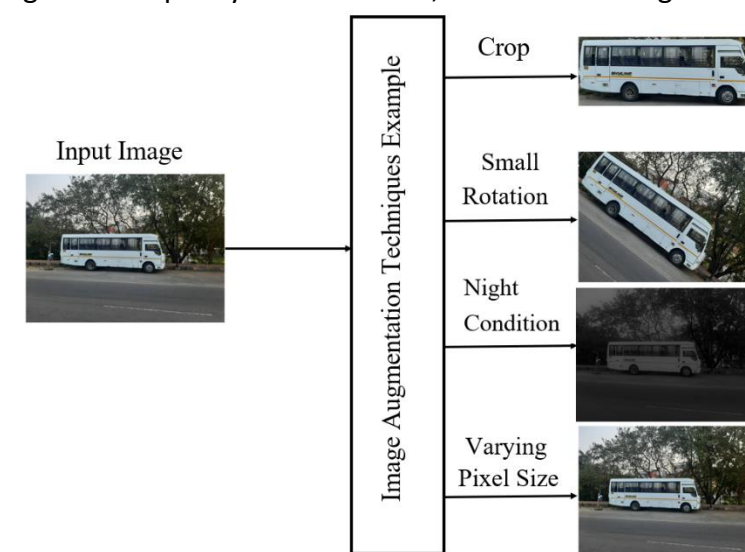

Figure 3: Example of Image Augmentation Techniques for bus image

3.1.2. Training, validation, and testing data set

In the proposed work, from the total image dataset, 80% of the images have been used to train the learning algorithm and the remaining 20% have been used for testing the performance of the classifier. For tuning the SVM classifier, approximately 20% of the available dataset in the training set has been utilized. Dataset modeling, training, and testing have been demonstrated in Figure 4.
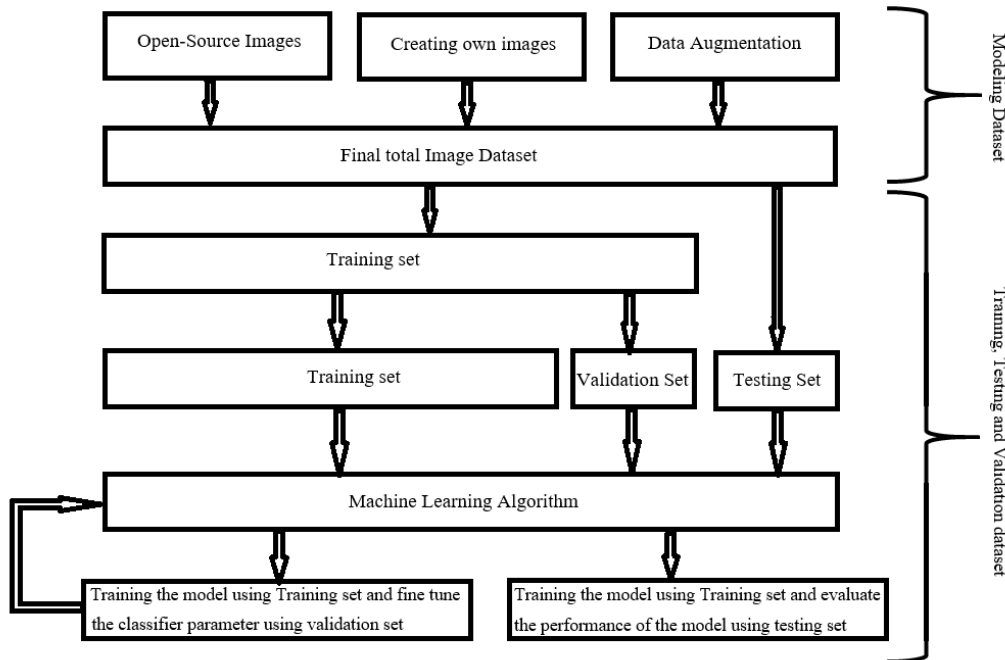

Figure 4: Dataset Creation Method

4. Classification Methodology

4.1. Feature extraction algorithms

Many feature extractor techniques are available to execute the aforementioned task; however, in the current work, HOG, SIFT, SURF, and Haar have been examined for experimental analysis, and the same algorithms have been explained as follows.

4.1.1. Histogram of Oriented Gradient (HOG)

Dalal and Triggs (2005) first derived HOG algorithm for human detection in 2005. Since then, several studies have been carried out in the automotive field for pedestrian detection, vehicle detection, etc. HOG operates using the sliding window technique with width and height that slides across an image. Usually, sliding window size is eight times more compared to HOG cell size; that is, sliding window pixels will be 32x32, 48x48, and 64x64 for the respective HOG pixel size of 4x4, 6x6, and 8x8. HOG extraction computes the magnitude and orientation of the gradient. The orientation histogram shall be derived from magnitude and orientation. If p(x,y) is a pixel data at (x,y) position of an image, then the HOG feature is to calculate the difference value, dx and dy for the x and y direction, respectively, as mentioned in Equations (1) and (2). The gradient magnitude and gradient orientation can be calculated as mentioned in Equations (3) and (4).

$$dx = p(x+1, y) - p(x-1, y) \tag{1}$$

$$dy = p(x, y+1) - p(x, y-1) \tag{2}$$

$$m(x, y) = \sqrt{(dx^2 + dy^2)} \tag{3}$$

$$\theta(x,y) = tan^{-1}\frac{dy}{dx} \tag{4}$$

4.1.2. Scaled Invariant Feature Transform (SIFT)

SIFT technique shall extract scale-invariant features from the image and this was published in 1999 by David Lowe (Lowe 1999). For a given set of reference images and for the given new image, an object is recognized individually, comparing each feature. Here Gaussian function is applied in the scale space to locate maxima and minima in the horizontal and vertical directions, as mentioned in Equation (5).

$$g(x) = \frac{1}{\sqrt{(2\pi\sigma)}} e^{-x^2/2\sigma^2} \tag{5}$$

Where input filtered signal is x and σ is Gaussian bases width and it is always constant feature space dimension number of 128. It means in SIFT extractor, previous image characteristics shall be the result of a numerical matrix multiplied by 128 integers. At each pixel (Ai,j,), the image gradient Magnitude (Li,j) and orientation (Ri,j) are computed using pixel differences as mentioned in Equations (6) and (7), respectively.

$$L_{i,j} = \sqrt{(A_{i,j} - A_{i+1,J})^2} + \sqrt{(A_{i,j} - A_{i,J+1})^2} \tag{6}$$

$$R_{i,j} = a\,tan\,2\,(A_{i,j} - A_{i+1,J}, A_{i,J+1} - A_{i,j}) \tag{7}$$

4.1.3. Speed-Up Robust Feature (SURF)

The SURF detector descriptor has been proposed by Bay et al. (2008) to improve speed and accuracy compared to scale-invariant feature transform (SIFT). The blob detector method has been used with the Hessian matrix to detect intersecting points, as mentioned in Equation (8). Once intersect point is made, then local neighborhood description is done using Hessian matrix; given a pixel, the Hessian of the pixel is as shown in Equation (9).

$$S(x,y) = \sum_{i=0}^{x} \sum_{j=0}^{y} I(i,j) \tag{8}$$

$$H(S(x,y)) = \begin{bmatrix} \dfrac{\delta^2 s}{\delta x^2} & \dfrac{\delta^2 s}{\delta x \delta y} \\ \dfrac{\delta^2 s}{\delta x \delta y} & \dfrac{\delta^2 s}{\delta y^2} \end{bmatrix} \tag{9}$$

4.1.4. Haar

Haar is a rectangular feature and it can be defined as the difference between the sum of image pixels inside the rectangle within the original image at any position. Later this simple rectangular feature was modified for 2, 3, and 4 rectangle features. Viola and Jones (2004) take credit for defining 3-rectangle and 4-rectangle features. The retrieval process of the four corner coordinates is done by calculating the image features in the integral image. From Equation (10), each point is scanned in the image, and then the integral image is calculated to calculate Harr features more efficiently.

$$AI(x,y) = \sum_{x'\leq x, y'\leq y}^{x} A(x',y') \tag{10}$$

An (x', y') is the integral image for the input image of AI(x,y).

4.2. Vehicle classification

Many machine learning classifiers are available for image processing applications, and multiclass SVM has been chosen in the current work because of its widespread use in automotive image processing. SVM classifier is a supervised learning algorithm that separates two classes with the maximum margin possible. The support vector in the classifier constructs the hyperplane that acts as a class descriptor between two datasets optimally. As shown in the final SVM equation (Equation 11), where Xi is the input image and X is the dataset image used for training with hyperplane parameters w and w0. The architecture for SVM-based vehicle classification is shown in Figure 5.

$$f(x) = \sum_{i \in S} \lambda i \; yi \; K \; (Xi \, . \, X \;) + w0 \qquad \qquad \text{(11)}$$
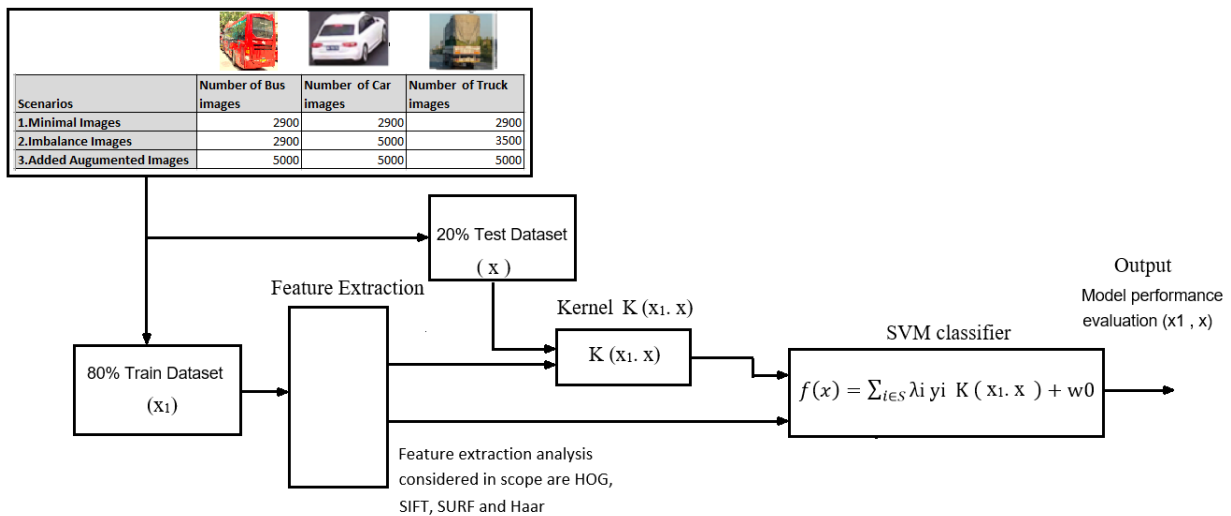

Figure 5: SVM Architecture for Vehicle Classification

4.3. Vehicle detection and grouping

The classification model assigns vehicles to the car, bus, or truck class types, and the lead vehicle in platoon management groups the vehicles accordingly, depending on the classification model's results. In the current study, classification is limited to only three vehicles: truck, bus, and car; therefore, grouping will be constrained as well. Based on the grouping outcome, the lead vehicle in the platoon management shall decide on the optimal placement of vehicles to achieve better fuel or energy efficiency. The current research effort does not focus on the optimal vehicle placement aspect, but it is being considered for future work awaiting the completion of the current study's learning algorithm.

5. Results and Discussion

The accuracy of the HOG, SIFT, SURF, and Haar feature extraction methods have been evaluated. Accuracy can be calculated as mentioned (Equation 12).

$$Accuracy = \frac{(True\ Positive\ +\ True\ Negative)}{(True\ Positive\ +\ True\ Negative +\ False\ Positive\ +\ False\ Negative)} \qquad \text{(12)}$$

Where true positive and true negative represents the correctly predicted classes, false positive and false negative are those that are incorrectly predicted as positive or negative. As mentioned in Table 1, for scenario-1: Minimal Image Dataset, each class of bus, car, and truck has 2900 images in the dataset, and images are gathered as mentioned earlier. For this, the experimental result is shown in Table 2. With the HOG feature extraction algorithm, the correctly classified images are 2250,2770 and 2240 images for bus, car, and truck class,

respectively. Misclassified images are mentioned in the red color cells. Based on the misclassified images, the false positive is 610,150 and 680 images for bus, car, and truck classes, respectively. The False negative is 650,130 and 660 images for bus, car, and truck classes, respectively. By using Equation (12), total accuracy has been computed and the model has achieved 83% and the same can be visualized in the experimental result from Figure 6 for the HOG feature extractor.

		Bus	2250	50	600	650		Correctly Classified
Actual Class		Car	50	2770	80	130		Misclassified
		Truck	560	100	2240	660		False Positive
			610	150	680	83%		False Negative
			Bus	**Car**	**Truck**			Average accuracy
			Predicted Class					

Table 2: Confusion matrix for scenario-1(Minimal Image Dataset) with HOG

As mentioned in Table 1, for scenario-2: Imbalance Image Dataset, each class of bus, car, and truck have 2900, 5000, and 3500 images in the dataset. Experimental results are shown in Table 3 for the HOG feature; here, correctly classified images are 1995,4800 and 2800 images for bus, car, and truck class, respectively. Misclassified images are mentioned in the red color cells. Based on the misclassified images, the false positive rate is 560,300 and 845 images for bus, car, and truck class respectively. A false negative is 905,200 and 600 images for bus, car, and truck class respectively. By using Equation (12), the total accuracy achieved is 85% and the same can be visualized in the experimental result from Figure 7 for the HOG feature extractor.

		Bus	1995	150	755	905		Correctly Classified
Actual Class		Car	110	4800	90	200		Misclassified
		Truck	450	150	2800	600		False Positive
			560	300	845	85%		False Negative
			Bus	**Car**	**Truck**			Average accuracy
			Predicted Class					

Table 3: Confusion matrix for scenario-2(Imbalance Image Dataset) with HOG

As mentioned in Table 1 for scenario-3: Added Augmented Image Dataset, each class of bus, car, and truck have 2900, 5000, 3500 images in the dataset. Additional images for bus and truck class have been increased by taking photos for bus and trucks of nearly 400 each and then using image augmentation technique, the quantity of the image dataset size has been increased. After increasing the image dataset size for bus and truck, the total images in each class is 5000. The results are shown in Table 4. Where for the HOG feature extractor, correctly classified images are 4680,4900 and 4710 images for bus, car, and truck class, respectively. Misclassified images are mentioned in the red color cells. Based on the misclassified images, the false positive is 310, 60, and 340 images for bus, car, and truck class, respectively. For the bus, car, and truck classes, the false negative is 320, 100 and 290 images, respectively. By using Equation (12), the total accuracy achieved is 85% and the same can be visualized in the experimental result from Figure 8 for the HOG feature extractor.

		Bus	Car	Truck		
Actual Class	Bus	4680	30	290	320	
	Car	50	4900	50	100	
	Truck	260	30	4710	290	
		310	60	340	95%	
		Bus	**Car**	**Truck**		
			Predicted Class			

Correctly Classified	
Misclassified	
False Positive	
False Negative	
Average accuracy	

Table 4: Confusion matrix for scenario-3 (Added Augmented Image Dataset) with HOG

A similar approach has been used to calculate average accuracy for all other feature extractors, SIFT, SURF, Haar, and consolidated values of all the mentioned feature extraction methods, as shown in Table 5. Where Figure 6, Figure 7, and Figure 8 show the experimental result and compare the true positive versus false-positive rates of HOG, SIFT, SURF, and Haar through the ROC curve for scenario-1, scenario-2, and scenario-3, respectively. In particular, for scenarios-2 and 3, the SIFT and SURF feature extractions performed approximately identically. This can be seen in Figure 7 and Figure 8, where the ROC curve of SURF overlaps with the ROC curve of SIFT, allowing just SURF's performance curve to be visualized.
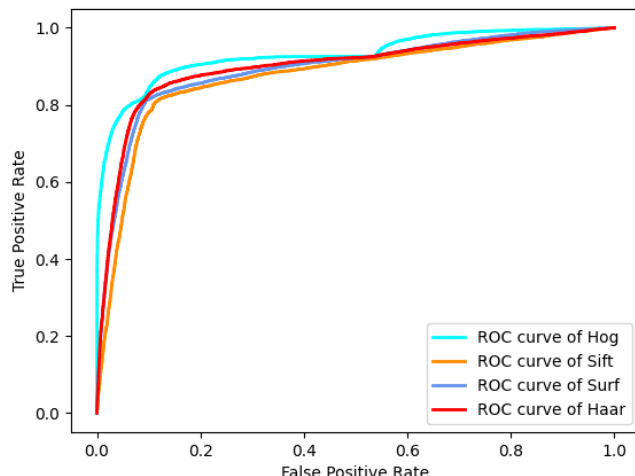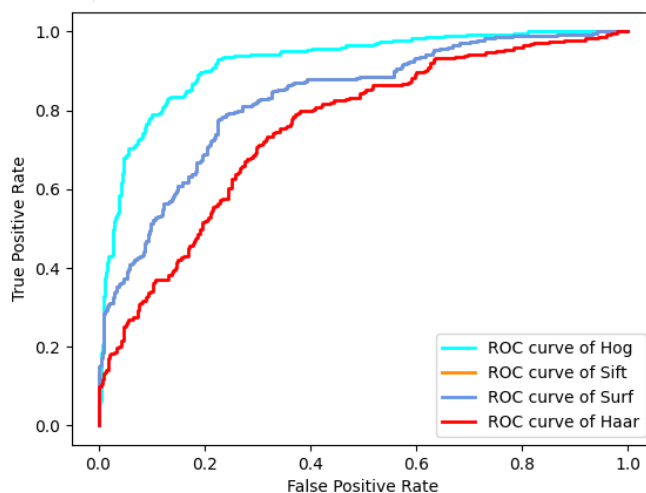

Figure 6: ROC for scenario-1(Minimal Image Dataset)

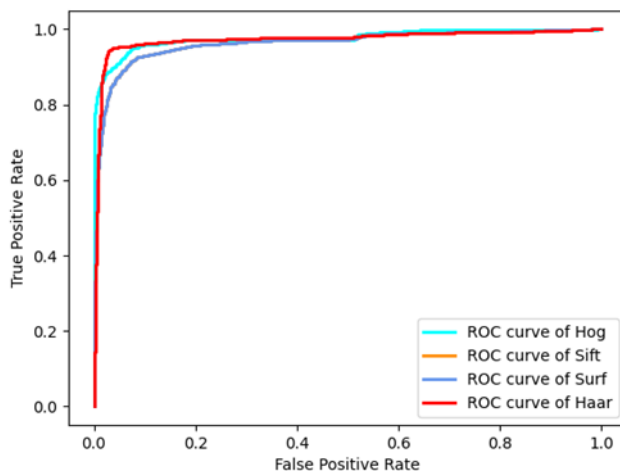**Figure 7**: ROC curve for scenario-2 (Imbalance image dataset)

Figure 8: ROC curve for scenario-3 (Added Augmented Image Dataset)

Classifier	Feature Extractor	Class	Accuracy in % for Scenario-1	Accuracy in % for Scenario-2	Accuracy in % for Scenario-3
SVM	HOG	Bus	78	78	94
		Car	94	94	99
		Truck	76	77	93
		Average	83	85	95
	SIFT	Bus	75	73	89
		Car	91	88	96
		Truck	73	73	88
		Average	80	78	91
	SURF	Bus	74	73	89
		Car	91	88	96
		Truck	72	73	88
		Average	79	78	91
	Haar	Bus	77	64	95
		Car	94	83	99
		Truck	75	65	94
		Average	82	71	96

Table 5: Performance evaluation of scenario 1(Minimal Image Dataset)

From Table 5, it can be summarized that for image dataset scenario-3, performance accuracies have increased more than 10% for all feature extractors when compared to scenario-1 and scenario-2. When the imbalanced image dataset was considered for experimentation, performance accuracy for SIFT, SURF, and Haar was degraded, but there was a slight increase only with the HOG feature extractor. For scenario-3, HOG and Haar feature extractors perform better, while HOG feature extractors have also performed better in scenarios-1 and 2. This experimental result illustrates that by increasing the dataset size via image augmentation techniques, the machine learning algorithm's overall accuracy and performance can be enhanced. In conclusion, the HOG feature extractor with SVM outperformed the other two methods in all three circumstances.

6. Conclusion

Because gathering appropriate image datasets for machine learning algorithm-based solutions in automotive image processing takes a lot of time and effort, this work shows how image augmentation can be used to create the necessary dataset and perform vehicle classification in automotive image processing. For various dataset sizes, analysis was conducted using prominent feature extractors such as HOG, SIFT, SURF, and Harr. In comparison to previous algorithms, HOG with SVM has shown improved performance for minimal images, unbalanced images, and augmented image datasets. As a result, the image augmentation technique can be used to increase the machine learning algorithm's accuracy. However, researchers have to aptly analyze and select appropriate image augmentation techniques for their research problem to avoid adverse effects. Based on the current work outcome, further work is being carried out to increase the classification of different vehicle classes in the vehicle platoon management system. The future study will expand the vehicle class and group them according to their diverse shapes, thereby improving the performance of vehicle platoon management utilizing HOG and SVM.

References

Amraoui, M., R. Latif, A. Elouardi, and A. Tajer. 2019. "Features extractors evaluation based V-SLAM applications". In *Proceedings of 2019 IEEE World Conference on Complex Systems (WCCS'19)*, Article number 8930787. https://doi.org/10.1109/ICoCS.2019.8930787.

Arróspide, J., L. Salgado, and M. Nieto. 2012. "Video analysis-based vehicle detection and tracking using an MCMC sampling framework". *Eurasip Journal on Advances in Signal Processing* 2012, no. 1. https://doi.org/10.1186/1687-6180-2012-2.

Ballarin, C., and M. Zeilinger. 2017. "The truck of the future: Autonomous and connected driving at Daimler trucks". Paper presented at the 9th AVL International Commercial Powertrain Conference 2017, Graz, Austria, May 2017. https://doi.org/10.4271/2017-01-1931.

Bay, H., A. Ess, T. Tuytelaars, and L. Van Gool. 2008. "Speeded-Up Robust Features (SURF)". *Computer Vision and Image Understanding* 110, no. 3: 346-59. https://doi.org/10.1016/j.cviu.2007.09.014.

Bougharriou, S., F. Hamdaoui, and A. Mtibaa. 2018. "Linear SVM classifier based HOG car detection". In *18th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering, STA 2017-Proceedings*, 241-45. https://doi.org/10.1109/STA.2017.8314922.

Cai, Y., Z. Liu, H. Wang, X. Chen, and L. Chen. 2018. "Vehicle detection by fusing part model learning and semantic scene information for complex urban surveillance". *Sensors* 18, no. 10: Article number 3505. https://doi.org/10.3390/s18103505.

Chang, J., L. Wang, G. Meng, S. Xiang, and C. Pan. 2018. "Vision-based occlusion handling and vehicle classification for traffic surveillance systems". *IEEE Intelligent Transportation Systems Magazine* 10, no. 2: 80-92. https://doi.org/10.1109/MITS.2018.2806619.

Chen, C., C. Huang, S. Sun, and J. Li. 2018. "Vehicle detection algorithm based on multiple-model fusion". *Journal of Computer-Aided Design and Computer Graphics* 30, no. 11: 2134-40. https://doi.org/10.3724/SP.J.1089.2018.17096.

Chen, X., and Q. Meng. 2013. "Vehicle detection from UAVs by using SIFT with implicit shape model". In *Proceedings - 2013 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2013*, 3139-44. https://doi.org/10.1109/SMC.2013.535.

Chen G.-h, X. Pan, and Z. Hou. 2016. "Preceding vehicle detection algorithm based on lane recognition and multi-characteristics". *Sci. Technol. Eng.* 16: 245-50. http://en.cnki.com.cn/Article_en/CJFDTOTAL-KXJS201615044.htm.

Dalal, N., and B. Triggs. 2005. "Histograms of oriented gradients for human detection". In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 886-93. https://doi.org/10.1109/CVPR.2005.177.

Kapoor, A., K. Grauman, R. Urtasun, and T. Darrell. 2007. "Active learning with Gaussian processes for object categorization". In *Proceedings of the IEEE International Conference on Computer Vision*, 1-8. IEEE. https://doi.org/10.1109/ICCV.2007.4408844.

Krause, J., M. Stark, J. Deng, and L. Fei-Fei. 2013. "3D object representations for fine-grained categorization". In *4th IEEE Workshop on 3D Representation and Recognition, at ICCV 2013* 554-61. https://doi.org/10.1109/ICCVW.2013.77. https://ai.stanford.edu/~jkrause/cars/car_dataset.html.

Lee, S. H., M. Bang, K. H. Jung, and K. Yi. 2015. "An efficient selection of HOG feature for SVM classification of vehicle". In *Proceedings of the International Symposium on Consumer Electronics, ISCE*, Article number 7177766. https://doi.org/10.1109/ISCE.2015.7177766.

Lemley, J., S. Bazrafkan, and P. Corcoran. 2017. "Smart augmentation learning an optimal data augmentation strategy". *IEEE Access* 5: 5858-69. https://doi.org/10.1109/ACCESS.2017.2696121.

Li, Y., and F. Y. Wang. 2015. "Vehicle detection based on And-Or Graph and Hybrid Image Templates for complex urban traffic conditions". *Transportation Research Part C: Emerging Technologies* 51: 19-28. https://doi.org/10.1016/j.trc.2014.10.009.

No in-text citation

Lin, B. F., Y. M. Lin, L. C. Fu, P. Y. Hsiao, L. A. Chuang, S. S. Huang, and M. F. Lo. 2012. "Integrating appearance and edge features for sedan vehicle detection in the blind-spot area". *IEEE Transactions on Intelligent Transportation Systems* 13, no. 2: 737-47. https://doi.org/10.1109/TITS.2011.2182649.

Liu, D., J. Wu, Y. Ye, J. Shi, and Y. Wang. 2019. "Research on multi-vehicle coordinated lane change of connected and automated vehicles on the highway". *SAE International Journal of Advances and Current Practices in Mobility* 1, no. 4: 1386-97. https://doi.org/10.4271/2019-01-0678.

Lou, Y., Y. Bai, J. Liu, S. Wang, and L. Duan. 2019. "VERI-wild: A large dataset and a new method for vehicle re-identification in the wild". In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 3230-38. https://doi.org/10.1109/CVPR.2019.00335.

Lowe, D. G. 1999. "Object recognition from local scale-invariant features". In *Proceedings of the Seventh IEEE International Conference on Computer Vision*, 1150-57. https://doi.org/10.1109/ICCV.1999.790410.

Naphade, M., S. Wang, D. C. Anastasiu, Z. Tang, M.-C. Chang, X. Yang, Y. Yao, et al. 2021. "The 5th AI City Challenge". *ArXiv*. [Summary of the 5th AI City Challenge Workshop in conjunction with CVPR 2021]. https://doi.org/10.48550/arXiv.2104.12233.

Nayak, A., S. Gopalswamy, and S. Rathinam. 2019. "Vision-based techniques for identifying emergency vehicles". *SAE Technical Papers* 2019-April. https://doi.org/10.4271/2019-01-0889.

Peng, C., M. M. Bonsangue, and Z. Xu. 2019. "Model checking longitudinal control in vehicle platoon systems". *IEEE Access* 7: 112015-25. https://doi.org/10.1109/ACCESS.2019.2935423.

Penghua, G., and H. Yanping. 2019. "Vehicle type classification based on improved HOG_SVM". In *Proceedings of the 3rd International Conference on Mechatronics Engineering and Information Technology (ICMEIT 2019)*, 640-47. Atlantis Press. https://doi.org/10.2991/icmeit-19.2019.102.

Shorten, C., and T. M. Khoshgoftaar. 2019. "A survey on Image Data Augmentation for Deep Learning". *Journal of Big Data* 6, no. 1: Article number 60. https://doi.org/10.1186/s40537-019-0197-0.

Sikirić, I., K. Brkić, and S. Šegvić. 2013. "Classifying traffic scenes using the GIST image descriptor". ArXiv, submitted October 1, 2013 [Part of the Proceedings of the Croatian Computer Vision Workshop, CCVW 2013]. https://doi.org/10.48550/arXiv.1310.0316.

Song, H., H. Liang, H. Li, Z. Dai, and X. Yun. 2019. "Vision-based vehicle detection and counting system using deep learning in highway scenes". *European Transport Research Review* 11, no. 1: Article number 51. https://doi.org/10.1186/s12544-019-0390-4.

Sun, Z., G. Bebis, and R. Miller. 2006. "Monocular precrash vehicle detection: Features and classifiers". *IEEE Transactions on Image Processing* 15, no. 7: 2019-34. https://doi.org/10.1109/TIP.2006.877062.

Tabassum, S., Md. S. Ullah, N. H. Al-nur, and S. Shatabda. 2020. Poribohon-BD. Mendeley Data, V2. https://doi.org/10.17632/pwyyg8zmk5.2.

Viola, P., and M. J. Jones. 2004. "Robust real-time face detection". *International Journal of Computer Vision* 57, no. 2: 137-54. https://doi.org/10.1023/B:VISI.0000013087.49260.fb.

Wei, Y., Q. Tian, J. Guo, W. Huang, and J. Cao. 2019. "Multi-vehicle detection algorithm through combining Harr and HOG features". *Mathematics and Computers in Simulation* 155: 130-45. https://doi.org/10.1016/j.matcom.2017.12.011.

Wen, X., L. Shao, Y. Xue, and W. Fang. 2015. "A rapid learning algorithm for vehicle classification". *Information Sciences* 295: 395-406. https://doi.org/10.1016/j.ins.2014.10.040.

Zakria, J. Cai, J. Deng, M. S. Khokhar, and M. Umar Aftab. 2018. "Vehicle classification based on deep convolutional neural networks model for traffic surveillance systems". In *2018 15th International Computer Conference on Wavelet Active Media Technology and Information Processing*, 224-27. IEEE. https://doi.org/10.1109/ICCWAMTIP.2018.8632593.