DOI: https://doi.org/10.33103/uot.ijccce.22.4.1

Machine Learning Techniques for Vehicle Detection

Maryam Raad Shihab[1], Rana Farid Ghani[2], Athraa Jasim Mohammed[3]

[1,2,3] Computer Sciences Department, University of Technology, Baghdad, Iraq

[1]110926@student.uotecnology.edu.iq, [2]110016@uotechnology.edu.iq, [3]10872@uotechnology.edu.iq

Abstract— *The traffic surveillance system is a type of intelligent system of traffic control. Traffic control provides solutions to most problems faced by people. It helps to monitor, detect traffic congestion and traffic accidents. As science evolved, it became possible to control traffic using video surveillance. Video surveillance is the most economical option that does not involve high costs or changes in infrastructure. Vehicle detection is one of the main parts of the traffic surveillance system. In this paper, vehicles will be detected using two different artificial intelligence methods (the YOLO method and the HAAR cascade classifier method). The first one is smarter than the second method, and both of them contain machine learning. The first processing step will read the video. Then vehicle detection algorithms are applied using two different ways. The comparison between them depends on the results to find the most effective and applicable vehicle detection method. After implementing the two methods, results were obtained using YOLO, that the accuracy is 91.31% and the error rate is 8.69%, in time 10 sec. As for using the XML (HAAR cascade classifier method) method, the accuracy is 86.9%, the quality is 86.9%, completeness is 90.9%, and the error rate is 13%, in time 17 sec. Thus, we conclude that the YOLO method has better results than the second method.*

Index Terms— *Artificial Intelligence, Machine Learning, Vehicle Detection, Yolo Method.*

I. INTRODUCTION

Since traffic has become worse and traffic accidents have increased, researchers have resorted to more effective monitoring of traffic. This technology detects vehicles, calculates the number of cars, predicts traffic, and gives the driver about it. With the development of science, it became possible to use artificial intelligence in vehicle detection techniques, proving its efficiency by previous scientists. The first method to detect vehicles was in 1978 using the computer vision method [1]. After that, many different studies have implemented to measure the performance and improve its effectiveness, and they discovered two different approaches to detect vehicles. The first approach uses the information of moving vehicles, and the second approach uses cars' inherent features to detect them in videos [2, 3].

The two methods contain advantages and disadvantages. When the method of detecting vehicles is adopted through vehicles' movement, this causes the stopped cars not to be detected. The second method detects things other than vehicles, such as humans, animals, and moving organisms. When the difference Lighting, i.e. in the dark, the two ways find it challenging to detect vehicles, artificial intelligence solves these problems.

Received 27/November/2021; Accepted 03/January/2022

 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.22.4.1

Several researchers have suggested ways to solve these challenges. Some of them Suggested a background subtraction algorithm to improve the segmentation of a moving object in a video [4, 5]. Then proposed an improved format for the presentation of background subtraction to mitigate the adverse effects of incremental changes. After that, the level set method uses to identify blobs. Finally, the Kalman filter and support vector machine uses to improve the accuracy of the classification of vehicles detection [6].

In this paper, the Yolo method and the Haar cascade classifier (XML) method will be used, because they are considered two of the most advanced and smarter methods than the previously mentioned methods. These two methods can detect parked cars, also discover vehicles only, not like the old methods that were discovering things other than vehicles, such as humans, animals and moving organisms.

II. OBJECT DETECTION METHODS

This paper implements two different methods in artificial intelligence: the cascade classifier method and the Yolo method. The cascade classifier is a machine learning used to classify the vehicles then detect them. The Yolo method is a deep learning method with a CNN (convolutional neural networks) and machine learning to classify and detect vehicles.

A. Vehicle Detection Using Haar Classifier Method

i. Haar-Like Features

The Haar-like feature depends on dividing the image into a rectangular and this rectangular divided into multiple parts. Mostly it will be a rectangle in black and white. As *Fig. 1* [7].

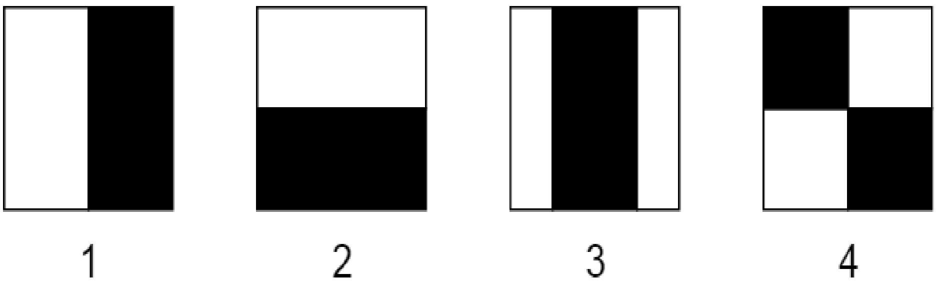


FIG. 1. BASIC HAAR-LIKE RECTANGLE FEATURE [7].

ii. AdaBoost Algorithm

The AdaBoost algorithm used to select features and improve performance frequently. AdaBoost builds an assertive synthesizer that combines many weak classifiers. The Viola-Jones AdaBoost algorithm uses to blend a series of AdaBoost matrices as a filter series. Each candidate is a discrete AdaBoost classifier consisting of a few weak classifiers. If each of these filters in the image acceptance area shows the vehicle failure, this area is immediately classified as not a vehicle. When a filter accepts a region of the image as a vehicle, it enters the next filtered region in the series. If this image area successfully crosses all of the series filters, it is classified as composite. In this algorithm, each cycle is defined to enhance an advantage among all other potential features. In the end, the final classification will be linear sets of the weak primary sort [8].

Received 27/November/2021; Accepted 03/January/2022
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.22.4.1

iii. Integral Image

The integrated image is a quick way to calculate the Haar-like feature, where the integrated image is the sum of all pixel values at the top and left of the position (x, y) [9].

iv. Cascaded Classifier

The cascade classifier is used to reject error windows and improve processing speed quickly. Every node of trees has a non-car branch, which means the image will not be a car. Through this technology, the negative rate is at least wrong. Use the Viola-Jones algorithm to detect vehicles. At first, a sequential file needs to be trained separately by OpenCV (open source computer vision) and must provide an XML file from it. OpenCV is used in classification to detect the vehicle. Images of the car in this software dataset are given to make training files and XML files. After that, the car can detect using the Haar-Like approach [10]. *Fig. 2* shows how the detection in cascade classifiers work.

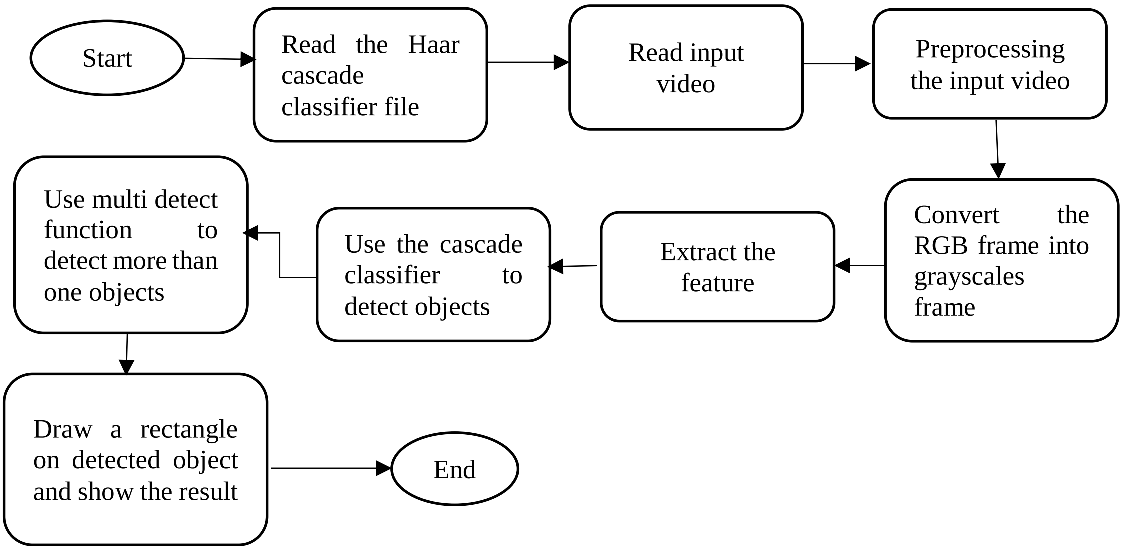


FIG. 2. FLOWCHART FOR HAAR CLASSIFIER METHOD.

v. Haar classifier Algorithm:

Haar feature-based cascade classifier is a machine learning that classifies the object depending on the cascade function to train a lot of positive and negative images then use these images to detect the object.

Vehicle detection algorithm using haar classifier

Input: Read input video

Output: Car detection

Processing:

Step1: Use python with OpenCV library to open the video.

Step2: Read the input video frames.

Step3: Read the haar cascade classifier.

Step4: Converting for every frame to grayscale because the cascade classifier works on a grayscale image.

Step5: Remove the noise from the extracted feature:

- **X= x + weight * height**

Received 27/November/2021; Accepted 03/January/2022

DOI: https://doi.org/10.33103/uot.ijccce.22.4.1

- **Y= y + weight * width**

Step6: Use a multi-detect function to detect vehicles of different sizes in every frame.
Step7: Identify the width and height of the rectangle.
Step8: Draw rectangle throughout the detected vehicles.
Step9: Show the result of vehicle detection.
End

B. Vehicle Detection Using Yolo Method

The meaning of the name YOLO is 'You Only Look Once. Yolo is an algorithm that detects and recognizes various objects in a picture (in real-time). Also, the Yolo is deep learning and a part of artificial intelligence. Object detection in YOLO is done as a regression problem and provides the class probabilities of the detected images. YOLO algorithm employs convolutional neural networks (CNN) to detect objects in real-time. As the name suggests, the algorithm requires only a single forward propagation through a neural network to detect objects. This means that the prediction in the whole picture is done in the process of running one algorithm. The CNN is used to predict various class probabilities and bounding boxes simultaneously [11].

The CNN consists of 106 layers composed of several types: convolutional layers are 75 layers, shortcut layer is 23 layers, route layer is 4, upsample layer is two, and finally, Yolo layers represent to detect the objects are three layers. The yolov3 weights are downloaded as a file, and the coco name of objects are 80 objects stored in a file as object names. *Fig. 3* represents the flowchart of the Yolo method.

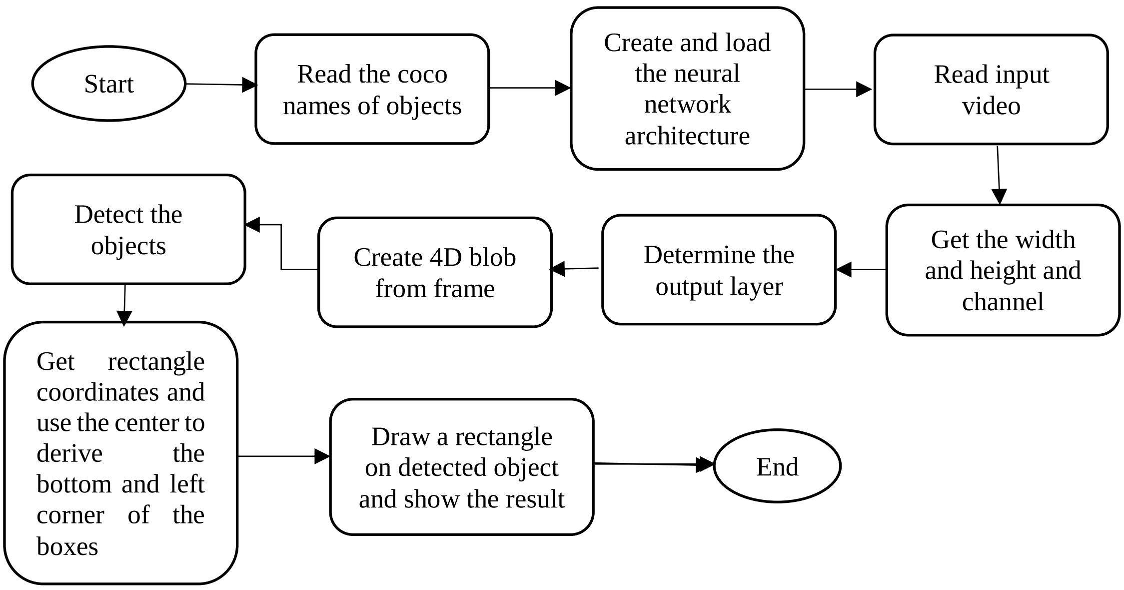


FIG. 3. FLOWCHART FOR YOLO METHOD.

Vehicle detection algorithm using Yolo
Input: Read input video
Output: Car detection

Received 27/November/2021; Accepted 03/January/2022

Iraqi Journal of Computers, Communications, Control & Systems Engineering (IJCCCE), Vol. 22, No. 4, December 2022

DOI: https://doi.org/10.33103/uot.ijccce.22.4.1

Processing:
Step1: Use python with OpenCV libraries.
Step2: Read the object names from the coco name file.
Step3: Create and load the CNN with weight using a dark-net network
Step4: Divided the image into different grids by getting the width, height, and channel where each grid has a dimension of S X S.
Step5: Determine the output layer starting with location [0].
Step6: Create a 4D blob form frame because the Yolo doesn't work on RGB images so the blob converts the RGB image to a grayscale image.
Step7: Detect the object by selecting areas of interest in the image.
Step8: Classify the selected areas using the created network that contained the CNN and weights.
Step9: Get rectangle coordinates and use the center to derive the bottom and left corner of the boxes.
Step10: A bounding box is drawn on each object detected by:

- **Width (w)**
- **Height (h)**
- **Class (for example, person, car, bus, etc.)**
- **Bounding box center (x, y)**

Step11: show the result of vehicle detection.
End

C. TRAINING PROPOSED METHODS

i. Training Haar Classifier Method

Haar cascade classifier has more than 710 nodes; every node has a threshold, left value, and right value in positive and negative numbers. Moreover, this node called features also has rectangles (between -1 and 22), and the size of this cascade is 24 x 24; this means that the width and height are 24 x 24, the version of XML is 0.1. The following tables I, II, and III are trains the Haar classifier; note these tables are samples of the total training where the total cascade has over 710 nodes.

TABLE I. TRANING HAAR CLASSIFIER TEST1

Parameters	Node one	Node two	Node three
Threshold	-0.181526198983192	-1.03817600756883 62e-003	-7.99356587231159 21e-003
Left value	0.5787503719329834	0.4794977903366089	0.629184186458587
Right value	-0.681496918201446	-0.52260810136795	-0.28534609079360
Rectangles	(1, 8, 22, 16, -1)	(20, 22, 4, 2, -1)	(3, 0, 18, 2, -1)
Size	24 x 24	24 x 24	24 x 24

This training gives the best result.

Received 27/November/2021; Accepted 03/January/2022
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.22.4.1

Traning two:

TABLE II. TRANING HAAR CLASSIFIER TEST2

Parameters	Node one	Node two	Node three
Threshold	2.875470090657472 6e-003	1.622635056264698 5e-003	-2.36251707974588 50e-005
Left value	-0.332929611206054	0.0397286005318165	0.1558447033166885
Right value	0.0741278976202011	-0.54077547788620	-0.16690459847450
Rectangles	(10, 7, 4, 5, 2)	(12, 6, 1, 2, 2)	(7, 21, 1, 1, 2)
Size	22x 22	22x 22	22x 22

Traning three:

TABLE III. TRANING HAAR CLASSIFIER TEST3

Parameters	Value of parameters in training one	Value of parameters in training two	Value of parameters in training three
Threshold	0.0262028109282255	-0.014905779622495	-0.029286060482263
Left value	-0.095133870840072	-0.110595896840095	0.2650567889213562
Right value	0.2311744987964630	0.0111535498872399	-0.08584027737379
Rectangles	(11, 5, 6, 1, 2)	(13, 1, 1, 9, 2)	(4, 1, 8, 5, 2)
Size	20 x 20	20 x 20	20 x 20

In an implementation, the training one gives the best result. Haar-like features divided the image into rectangles. That's why there is a rectangle parameter in training; also, there are two values left and right because the integral image has two values (x, y). The positive and negative values are the features of the left and right values in every node.

ii. Training Yolo Method

Training the Yolo three times shown in tables IV, V, and VI then will find the best result.

Training one:

TABLE IV. TRANING YOLOV3 TEST 1

Parameter	Value
Batch	32
Subdivisions	10
Width, the height of the input images	416 x 416
Channels	5
Momentum	0.9
Decay	0.0005
Saturation and exposure	1.5
Angle	0
Hue	0.1

Received 27/November/2021; Accepted 03/January/2022

DOI: https://doi.org/10.33103/uot.ijccce.22.4.1

Learning rate	0.012
Burn in	1000
Max batches	500200
Policy	Steps
Steps	400000,450000
Scales	.1,.1

This training did not give a result because it couldn't find the output layer.

Training two:

TABLE V. TRANING YOLOV3 TEST 2

Parameter	Value
Batch	45
Subdivisions	14
Width, height of the input images	416 x 608
Channels	4
Momentum	1.0
Decay	0.0005
Saturation and exposure	1.05
Angle	1
Hue	0.001
Learning rate	0.001
Burn in	1000
Max batches	500200
Policy	Steps
Steps	400000,450000
Scales	.1,.1

This training gives a bad result because the size must be equal (the width and height), batch and subdivisions are less than the actual batch, and subdivisions, hue, angle, and learning rate also didn't give the best result in these values.

Training three:

TABLE VI. TRANING YOLOV3 TEST 3

Parameter	Value
Batch	64
Subdivisions	16
Width and height of the input images	608 x 608
Channels	3
Momentum	0.9
Decay	0.0005
Saturation and exposure	1.5
Angle	0
Hue	0.1
Learning rate	0.001
Burn in	1000
Max batches	500200
Policy	Steps
Steps	400000,450000
Scales	.1,.1

This training gives the best result and high accuracy.

Received 27/November/2021; Accepted 03/January/2022

ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.22.4.1

III. IMPLEMENTATION AND RESULT

The first step is to input the video into Python and set a snapshot from the video. The second step is processing the image to detect the appeared cars. Then show the result of vehicle detection using the Haar classifier and Yolo, compare these two methods and suggest the best method with reasons.

i. Result of Haar Classifier Method

The total number of detected vehicles using this method is 20 vehicles. The actual vehicle is 23, the processing time is 17 seconds, and the performance measurement [12] will test the proposed method and verify its efficiency.

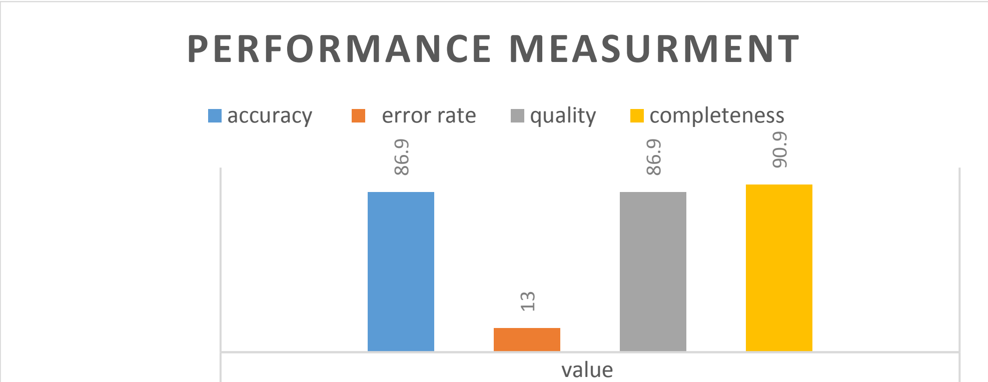
Accuracy = TP / (TP + FP) (1)

Completeness = TP / (TP + FN) (2)

Quality = TP / (TP + FP) (3)

Error rate = 1 – Accuracy (4)

Where: TP = True positive (the number of true detected vehicles; FP = False positive (the number of false vehicles); FN = False negative (the number of vehicles which did not detect). The purpose of the performance measurement is to measure the performance of the proposed system which is measured by accuracy, completeness, quality also measures the error rate. *Fig. 4* shows the result of performance measurement, whereas *Fig. 5* is a sample of vehicle detection using the Haar classifier method.



	value
accuracy	86.9
error rate	13
quality	86.9
completeness	90.9

FIG. 4. PERFORMANCE MEASUREMENT.

Received 27/November/2021; Accepted 03/January/2022
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.22.4.1

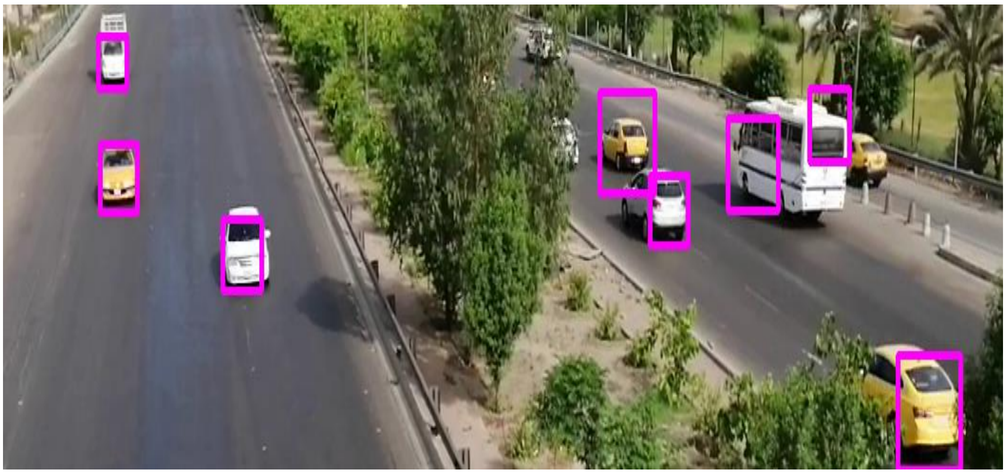


FIG. 5. SAMPLE OF VEHICLE DETECTION USING HAAR CLASSIFIER METHOD.

The Harr classification method can only detect vehicles, which means it cannot detect a motorcycle, human, bike, or anything other than vehicles because the features in the Haar classifier are for cars only. The result of the Haar classifier method shows its disability to detect three vehicles because; either the cars are far away from the camera or Vehicles do not appear well in front of the camera, so the cars cannot be detected.

ii. Result of Yolo Method

The dataset in the start is labelled for 80 classes: [person, bicycle, car, motorbike, plane, bus, train, truck, etc.]. It uses a parameter to train the yolov3 network in the Table VII

TABLE VII. INITIAL VALUES OF YOLO NETWORK PARAMETER

Parameter	Value
Batch	64
Subdivisions	16
Width and height of the input images	608 x 608
Channels	3
Momentum	0.9
Decay	0.0005
Saturation and exposure	1.5
Angle	0
hue	0.1
Learning rate	0.001
Burn in	1000
Max batches	500200
Policy	Steps
Steps	400000,450000
Scales	.1,.1

The Yolo has 75 CNN layers (convolutional layers) + 31 other layers (shortcut, route, upsample, Yolo) = 106 layers in total, the min confidence (Minimum confidence threshold) is 0.14, increasing this will improve false positives but will also reduce detection rate.

Table VIII below shows the detection for different kinds of vehicles with their probability, accuracy, error rate, and time.

Received 27/November/2021; Accepted 03/January/2022

DOI: https://doi.org/10.33103/uot.ijccce.22.4.1

TABLE VIII. RESULT OF VEHICLE DETECTION

Confidence	Probability for car	Probability for trunk	Probability for bus
0.5	0.8	0.8	0.7
0.4	1.0	0.4	0.3
0.3	0.9	0.2	0.9
0.9	0.6	0.3	1.0
0.8	0.9	0.8	0.3
0.8	1.0	1.0	0.7
0.3	0.7	0.7	1.0

The confidence score is the result of the score as the probability for an object, which it's per bounding box is one of the outputs of the neural network; note that it is not recomputed, but it is used in making the final output based on which boxes have the highest confidence. Probability for car, trunk, and a bus is this may be a car, trunk, or bus?

The time of the detection process is 10 secs, the total number of vehicles detected are 25 (detect 23 cars and two persons), the total number of actual vehicle is 23 vehicles. *Fig. 6* shows the result of vehicle detection accuracy and error rate. *Fig. 7* shows a sample of vehicle detection using yolov3.

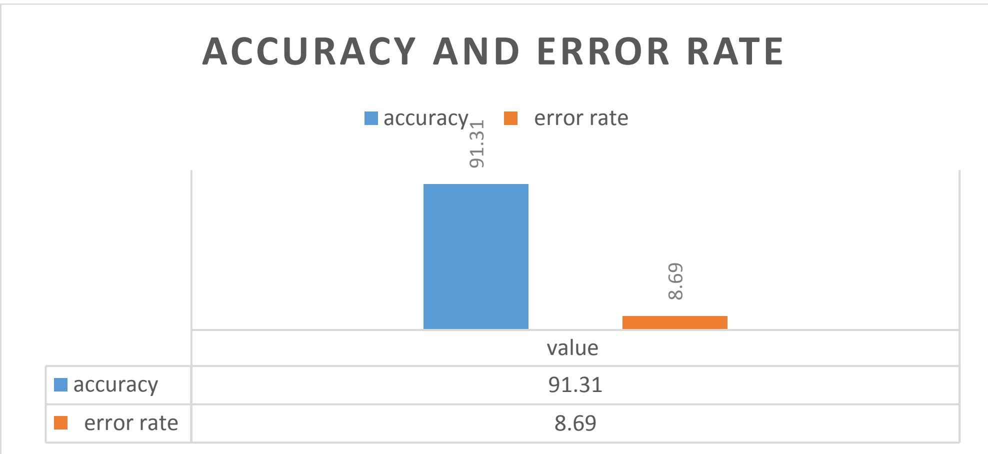


FIG. 6. ACCURACY OF VEHICLE DETECTION AND ERROR RATE.

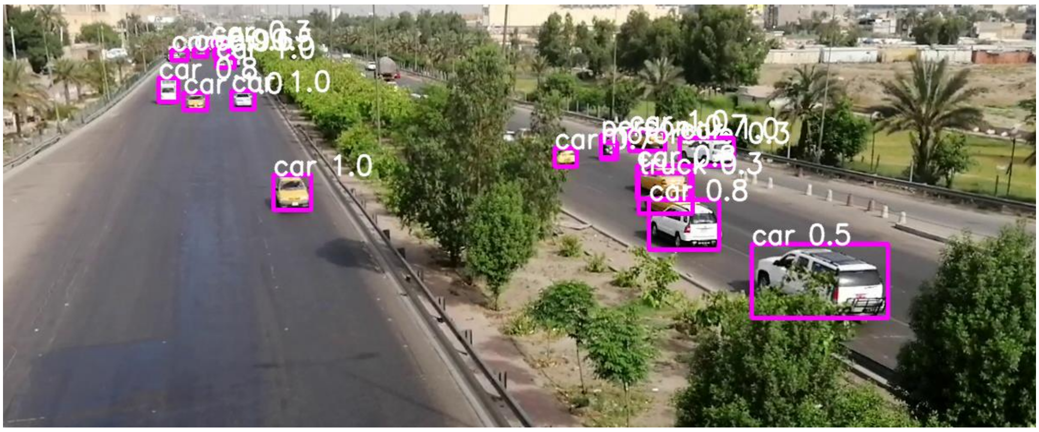


FIG. 7. VEHICLE DETECTION USING YOLO.

Received 27/November/2021; Accepted 03/January/2022

DOI: https://doi.org/10.33103/uot.ijccce.22.4.1

The Yolo method detects 80 objects, not only vehicles, but it can also determine which object must be detected. As the result of the Yolo method, it detects all the vehicles. Additionally, it detects persons and motorbikes because we didn't select to detect only vehicles.

IV. COMPARING BETWEEN YOLO, HAAR CLASSIFIER, AND OTHER OLD METHODS

The other old methods are: the first approach used the information of moving vehicle and the second approach used the inherent features of vehicles to detect them in videos.

Haar Classifier	Yolo	Other Old Methods
Slower than yolo	Faster than haar classifier	Fast
Can detect more than one time if there is no condition to detect one time	Detect only one time	Detect multi objects
Easier to implement	Harder to implement	Easy to implement
More accurate	More accurate than haar and other old methods	Less accuracy
It has only the machine learning	More intelligent because it has the CNN and machine learning	Don’t have any intelligence it depends on computer vision to detect objects
Simple	More complicated	Simple than haar classifier
The error rate is 0.03	The error rate is 0.13	The error rate more than 0.5
It detects only vehicle	It detects more than 80 object	Can detect vehicles and non-vehicles
Have 106 layers to detect objects	Have more than 710 nodes to detect vehicle only	Don’t have any layers or node
Detect moving and stopped vehicles	Detect moving and stooped vehicles	When the method of detecting vehicles is adopted through the movement of vehicles, this causes the stopped vehicles to not be detected. In the second method, things other than vehicles such as humans, animals, and other moving organisms are detected.

Received 27/November/2021; Accepted 03/January/2022

DOI: https://doi.org/10.33103/uot.ijccce.22.4.1

The best method is the Yolo method because it detects only one time, has more accuracy depending on the result, and is more intelligent.

V. CONCLUSIONS

In conclusion, vehicle detection helps to surveillance the traffic, reducing accidents and controlling traffic conditions, detecting vehicles using Haar classifier gives excepted result where the accuracy is 86.9%, Completeness is 90.9%, Quality is 86.9%, and the error rate of the detection is 13% in time 17 sec. While detecting vehicles using Yolo provides improved detection results compared to the Haar classifier, the most important thing when training yolov3 is a comprehensive data set of cars in all kinds of positions, views, image resolutions, etc. In this way, the model learns the concept of the car in all its forms that may appear in real life, then the idea of the car will be good and will be able to detect vehicles in the new images that were not used to train the model. It is better to train the network in the first steps with low-resolution images (608 × 608) and then randomly choose the network resolution values and change the network configuration to get a model capable of detecting compounds in the high and low-resolution images. An attempt was made to reduce training time using raw weights obtained from previously trained weights for the COCO and image net dataset. The accuracy of yolov3 is 91.31 %, the error rate is 8.69 % in time 10 sec. Finally, the Yolo method gives better results than the Haar classifier method.

REFERENCES

[1] Y. Liu, B. Tian, S. Chen, E. Zhu, and K. Wang, "A survey of vision-based vehicle detection and tracking techniques in ITS," in Conf. vehicular electronics and safety, IEEE, Dongguan, China, pp. 72–77, July 2013.

[2] A.S. Kalaki, and R. Safabakhsh, "Current and adjacent lanes detection for an autonomous vehicle to facilitate obstacle avoidance using a monocular camera," in Conf. Intelligent Systems, Bam, Iran, February 2014, pp. 1–6.

[3] L.W. Tsai, J.W. Hsieh, and K.C. Fan, "Vehicle detection using normalized color and edge map," IEEE transactions on Image Processing, vol. 16, issue 3, pp. 850–864, 2017.

[4] N.A. Mandellos, I. Keramitsoglou, and C.T. Kiranoudis, "A background subtraction algorithm for detecting and tracking vehicles," Expert Systems with Applications, vol. 38, issue 3, pp. 1619–1631, March 2011.

[5] L. Unzueta, M. Nieto, A. Cortes, J. Barandiaran, O. Otaegui, and P. Sanchez, "Adaptive Multicue Background Subtraction for Robust Vehicle Counting and Classification," IEEE Transactions on Intelligent Transportation Systems, vol. 13, issue 2, pp. 527–540, 2012.

[6] Z. Chen, T. Ellis, and S.A. Velastin, "Vehicle detection, tracking and classification in urban traffic," in Conf. Intelligent Transportation Systems, IEEE, Anchorage, AK, USA, pp. 951–956, 15th September 2012.

[7] S. Sivaraman, and M.M. Trivedi, "Looking at vehicles on the road: A survey of vision-based vehicle detection, tracking, and behavior analysis," IEEE transactions on intelligent transportation systems, vol. 14, issue 4, pp. 1773–1795, 2013.

[8] P. Menezes, J.C. Barreto, and J. Dias, "Face tracking based on haar-like features and Eigen faces," IFAC Proceedings Volumes, vol. 37, issue 8, pp. 304–309, July 2004.

[9] K. Noor, E.A. Siddiquee, D. Sarma, A. Nandi, S. Akhter, Sohrab Hossain, et al., "Performance analysis of a surveillance system to detect and track vehicles using Haar cascaded classifiers and optical flow method," in Conf. Industrial Electronics and Applications, IEEE, Siem Reap, Cambodia, pp. 258–263, June 2017.

[10] K.G. Derpanis, "Integral image-based representations," Department of Computer Science and Engineering York University Paper, vol. 1, issue 2, pp. 1–6, 2017.

[11] L. Shengyu, W. Beizhan, W. Hongji, C. Lihao, L. Ma, and Z. Xiaoyan, "A real-time object detection algorithm for video," Computers & Electrical Engineering, vol. 77, pp. 398–408, July 2019.

[12] M.M. Moghimi, M. Nayeri, M. Pourahmadi, and M.K. Moghimi, "Moving vehicle detection using AdaBoost and haar-like feature in surveillance videos," International Journal of Imaging and Robotics, vol. 18, issue 1, pp. 94–106, 2018.

Received 27/November/2021; Accepted 03/January/2022