GROUP ASSIGNMENT

Vehicle Detection & Classification

By

Wahidul Alam Riyad (TP043338)

Sylvia Chan Poh Yee (TP050862)

UC3F2002IS

B.Sc. (Hons) in Intelligent Systems

Lecturer Name : Dr Hamam Mokayed

Hand-Out Date : Week 9

Hand-In Date : 29th May 2020

Acknowledgement

I would also like to thank Dr Hamam Mokayed, supervisor of the report, who has shown great appreciation for the important role he has played in supporting the best result of the work of the research. Dr. Hamam Mokayed has made every effort to guide me towards achieving my goal, as well as his motivation to keep moving along the right track. I am also thankful for the support and guidance I have provided on all the various matters. He does not hesitate to help in the light of his contributions, and the extra time given to help me achieve my goals. His passion, motivation, and confidence helped. He was always available to answer my questions and gave me his time and immense knowledge generously. I confirm that this work is my task, and I have received an information resource acknowledgment.

Table of Contents

Abstract

This student project will have practical knowledge in programming and in computing regarding image processing , computer vision and pattern recognition. The aim will not be to teach all programming specifics under different platforms but to provide an excellent understanding of modular algorithms and the use of coding techniques for designing applications across various platforms and genres as well as to provide realistic implementation experiences. The goal is to provide an automatic identification and classification solution of the vehicle (car / bus, truck / motorcycle).

Introduction

Object detection and classification are two important tasks for image processing. The individual features of each object are special. During object detection, the object's classification provides information on the position of the object, which class belongs to. Object classification is the process of finding specific real world objects such as faces, buildings and vehicles, and is a technology based on computer science image processing (Kul, Eken and Sayar, 2017).

Proposed Algorithms

Saliency

The purpose of applied saliency function is to signify conspicuity of the image at each place in graphical pitch with amount of scalar as well as to monitor the range of concentrated places, established on spatial allocation for saliency. It also represents to make the image better as the image able to show the most important part and focus it automatically to improve the performance which is the item need to be detected.

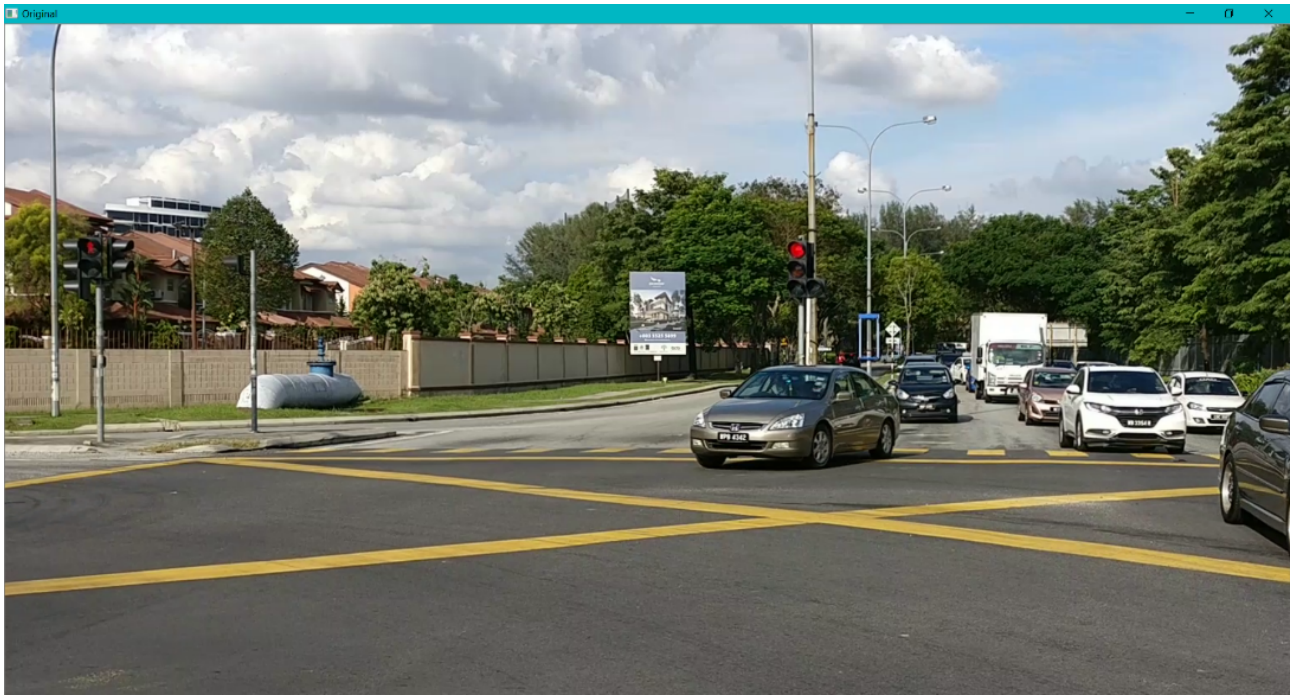

Figure 1 shows original image

Figure 2 shows the image after saliency

Background Subtraction

To apply the function of background subtraction in the image is to remove the objects from the background and keep all the objects in the foreground. The background subtraction is method for isolating out front components from back components and it is finished by producing a front mask. This method is utilized for distinguishing progressively moving items as of fixed cameras. Background subtraction method is significant used for any system that needs to track objects. Here background subtraction used to remove the object in the background to ensure the items in the foreground are clear to be detected. The white spots are the main thing needed to be focus as foreground.

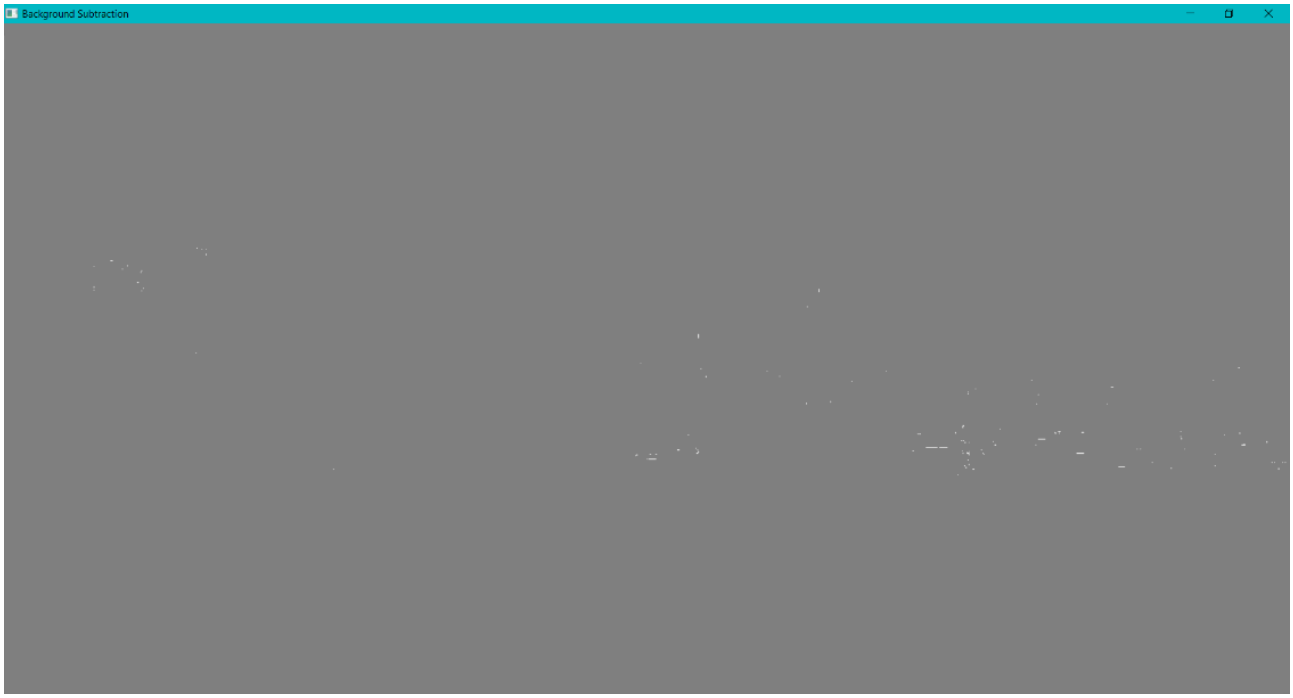

Figure 3 shows the image after apply background subtraction

K-means

Clustering is the extremely popular exploratory information evaluation method that used to find impulse regarding the formation of the information. It tends to describe as mission of recognizing subsections in information with the end goal that information focuses in a similar subsection are fundamentally the same as while information focuses in various clusters are totally unique.

K-means algorithm is repetitive calculation that attempts to parcel the dataset to Kpre-defined particular non-covering subsections wherever every data point has a place with one cluster. It attempts to create intra-group information focuses as comparative as it could be expected under the circumstances while equally keeping the bunches as various as could be expected under the circumstances. It allocates information focuses to a bunch to such an extent that the entirety of the squared separation between the information focuses and the group's centroid is at the base (Dabbura, 2018).

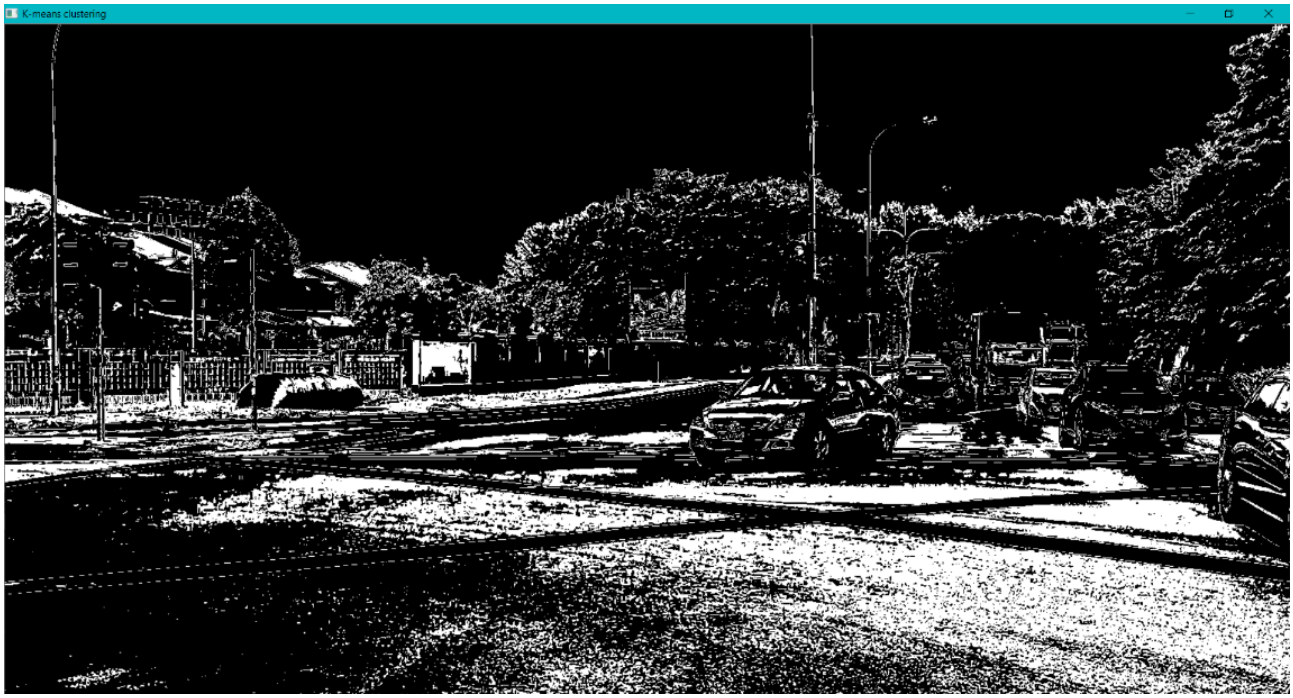

Figure 4 shows the image after applied K-means algorithm

Dilation

The reason of using dilation is to enhance the edge of the binary picture dependent on the picture of edge recognition. The scope of the brilliant segment will increase more as each and every pixel has been changed by the greatest value. In here, the dilation is implemented to compound the k-means clusters group by group so it can be used for image segmentation.

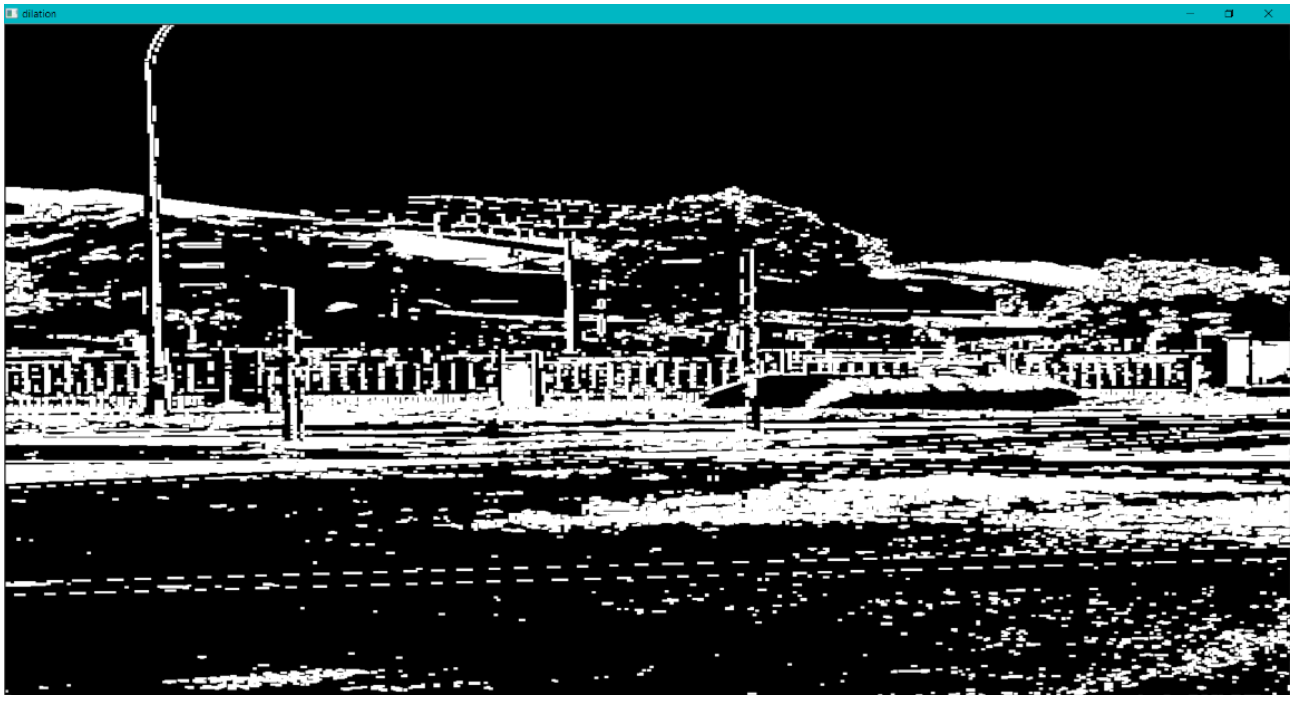

Figure 5 shows the image after applied the function of dilation

Erosion

The explanation of actualized the use for erosion is to diminish the pixel on the picture with the binarized picture. The number of pixels remove is relies upon the outcome of picture. The goals to use disintegration on the tag is to cut the size of characters and numbers on the picture so it tends to be independent. Moreover, it is similarly as to limit from the picture to remain independently with other excess sections after expansion which can cause it sifted through the filtration technique. Beside these methods, the estimation of the yield pixel is the most minimal estimation of the neighbor pixel esteem.

Figure 6 shows after applied the function of erosion

Blob

Blob describes as segments that planned for recognizing focuses and additionally areas in the picture that change in properties as intensity or shading contrasted with the comprising. One of the primary purposes is to give mutual data about environments, which not acquired as of edge indicators or else angle finders. It is utilized to acquire areas of excitement for additional preparing. These areas could mark the nearness of items or sections of items in the picture space with application to demonstrate recognition as well as item tracing.

Blob identification is normally performed after shading recognition and noise decrease to completely locate the necessary items from the picture. A great deal of immaterial blobs of necessary shading might be available in the image and furthermore need the particulars of the blobs

and the blob angles and just for which needed that the specific directions of the considerable number of pixels where the necessary blob is available in (Technology Robotix Society, 2019).

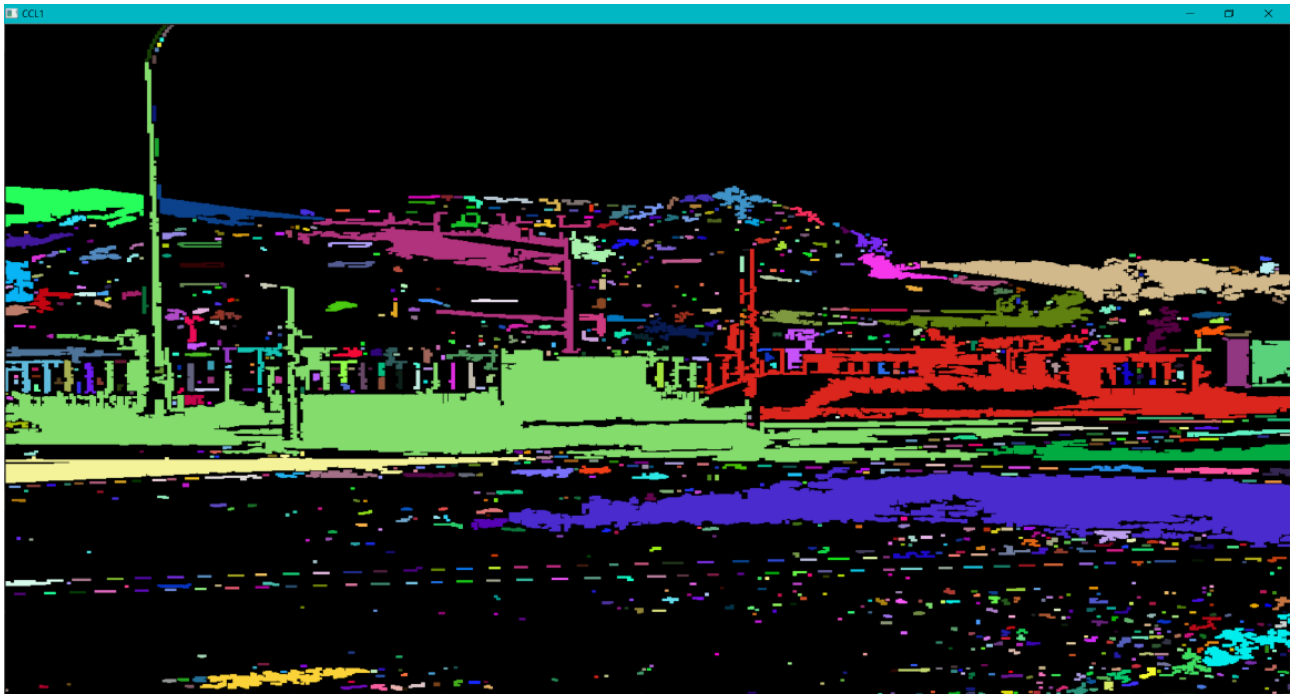

Figure 7 shows the image after the blob segmentation is applied

YOLO

Object recognition is one of the common issues in the work to perceive what is the items are inside a given picture and furthermore where all are in picture. The issue of item identification is more confusing than categorization, which additionally can perceive issues however it does not demonstrate where the item is placed in picture. Also, categorization does not deal with pictures including additional than one item.

YOLO utilizes a extremely unique methodology. YOLO defines as intelligent convolutional neural network (CNN) for executing item identification progressively. The calculation operates a solitary neural network to full picture, and afterward separates the picture into areas and forecasts jumping boxes and possibilities for every single area. These jumping boxes are biased by anticipated possibilities.

YOLO is famous in light of the fact that it accomplishes high precision while additionally having the option to run constantly. The calculation that refers to “you only look once” at following picture as in it needs exactly one further generating go across neural network for creating forecasts.

Once the item ensure finding estimate just identifies each item once, it will return perceived items along with the jumping boxes (Open Data Science, 2018).

Figure 8 shows the image that after applied YOLO algorithm

```
// Remove the bounding boxes with low confidence using non-maxima suppression
void postprocess(Mat& frame, const vector<Mat>& out);

// Draw the predicted bounding box
void drawPred(int classId, float conf, int left, int top, int right, int bottom, Mat&
frame);

// Get the names of the output layers
vector<String> getOutputsNames(const Net& net);

// Remove the bounding boxes with low confidence using non-maxima suppression
void postprocess(Mat& frame, const vector<Mat>& outs)
{
	vector<int> classIds;
	vector<float> confidences;
	vector<Rect> boxes;

	for (size_t i = 0; i < outs.size(); ++i)
	{
		// Scan through all the bounding boxes output from the network and keep
only the
		// ones with high confidence scores. Assign the box's class label as the
class
		// with the highest score for the box.
		float* data = (float*)outs[i].data;
		for (int j = 0; j < outs[i].rows; ++j, data += outs[i].cols)
		{
			Mat scores = outs[i].row(j).colRange(5, outs[i].cols);
			Point classIdPoint;
			double confidence;
			// Get the value and location of the maximum score
			minMaxLoc(scores, 0, &confidence, 0, &classIdPoint);
			if (confidence > confThreshold)
			{
				int centerX = (int)(data[0] * frame.cols);
				int centerY = (int)(data[1] * frame.rows);
				int width = (int)(data[2] * frame.cols);
				int height = (int)(data[3] * frame.rows);
				int left = centerX - width / 2;
				int top = centerY - height / 2;

				classIds.push_back(classIdPoint.x);
				confidences.push_back((float)confidence);
				boxes.push_back(Rect(left, top, width, height));
			}
		}
	}
```

Figure 9 shows part of YOLO algorithm

Experimental Results

The Vehicle Detection and Classification system had able to classify the vehicles in the given video. It can successfully detect and classify the following objects in the image as car, bus, truck, and motorbike. All vehicles had been successfully detected and has been labelled with following categories that it belongs to.

Figure 10 shows sample of the car been detected and labelled in the video

Figure 11 shows sample of the truck and bus been detected and labelled in the video

Figure 12 shows samples of the motorbike has been detected and labelled in the video

Class	Score
Bus	0.74, 0.66, 0.77
Car	0.86, 0.89, 0.98, 0.99, 1.00
Truck	0.94, 0.98
Motorbike	0.81, 1.00

Table 1 shows the output for a set number of detection results

Results Discussion

Object detection is computer knowledge connected with computer vision as well as image handling that recognizes and describes items, for example, people, constructions and vehicles from advanced pictures and recordings video. This innovation has the ability to organize only one or else a few items inside a computerized picture without a moment delay. The object detection works well to recognize the objects like car, truck, bus, motorbike, traffic light and person.

The accuracy of YOLO v3 implemented in the project is faster and more accurate based on the coco dataset. It contains with 53 convolutional layers that is efficiency to recognize the objects in the video while the YOLO v2 only contain with 19 convolutional layers. The dilation applied in the Vehicle Detection and Classification system here to enhance the accuracy of detection in using the YOLO v3. There are high confidence score showing in the Vehicle Detection and Classification system. Confidence score is the possibility that handle box includes an item. It is normally expected by a classifier. It will only get the highest score to ensure the prediction is correct and accurate.

The score of the object detected between the range of 0 to 1, which objects detected with the score closer to 1 is more accurate depending on the probability. One of the benefits for applying YOLO v3 is the algorithm included the labelling for the object, it can clearly define with objects in the video belongs to which labelling so the output can easily understand by the user.

Analysis & Future Direction

Somehow it is slow and delay if implement the YOLO algorithm to the Vehicle Detection and Classification system without a GPU. It is 1,549% faster for the object detection with using a NVIDIA GPU. The programmer had experienced on CPU to implement YOLO algorithm but it still able to work. Before apply YOLO, darknet need to be setup. Darknet refers to a framework use for the training of neural networks, it also defines as public resource and authored in CUDA as well as operates as foundation in YOLO. Even though the programmer does not have a GPU but darknet also need to setup for running the YOLO algorithm.

Object identification is smashing into wide-ranging scope of enterprises, with utilize cases extending from individual protection to profitability in work environment. Facial recognition is one of it, which able to use as safety attempt to give specific individuals access to an exceptionally ordered territory of an administration working, for instance. It may be utilized to check the quantity

of individuals present inside a conference to consequently modify other specialized devices that will help smooth out the time committed to that specific discussion. It able to be utilized inside a visual internet searcher to assist customers with finding a particular thing they are on chasing for. Pinterest is one of it, as whole social and spending stage is worked around for this innovation. The late appeared of Intelligent Vision 2.0, a setup of AI elements that enhance coordinated effort by conveying exact information to support additional mechanize work processes. These highlights use individuals and object recognition to make huge information for a range of utilizations in the work environment.

Conclusion

In the conclusion, the programmer had learnt how to implement the YOLO algorithm for object detection and it has successfully show the accurate classes for detecting the vehicles. Object detection is useful for the vehicle detection as it can be used to detect the number of the car and type of the car passing by the highway. It can be enhanced to a system to reduce the accident happened on the road. Moreover, the programmers had met the necessity as the given task question and they had figured out how to executed all the calculations to get the outcomes that has been recorded in the report.

Workload Matrix

Tasks	Wahid (TP043338)	Sylvia (TP050862)
Acknowledgement	50%	50%
Table of Contents	50%	50%
Abstract	50%	50%
Introduction	50%	50%
Proposed Algorithms	50%	50%
Experimental Results	50%	50%
Results Discussion	50%	50%
Analysis & Future Direction	50%	50%
Conclusion	50%	50%
Workload Matrix	50%	50%
References	50%	50%
Signature		

References

1. Dabbura, I., 2018. *K-means Clustering: Algorithm, Applications, Evaluation Methods, and Drawbacks*. [Online] Available at: https://towardsdatascience.com/k-means-clustering-algorithm-applications-evaluation-methods-and-drawbacks-aa03e644b48a
   [Accessed 22 May 2020].

2. Kul, S., Eken, S. and Sayar, A., 2017. A concise review on vehicle detection and classification. 2017 International Conference on Engineering and Technology (ICET),.

3. Open Data Science, 2018. *Overview of the YOLO Object Detection Algorithm.* [Online] Available at: https://medium.com/@ODSC/overview-of-the-yolo-object-detection-algorithm-7b52a745d3e0
   [Accessed 22 May 2020].

4. Technology Robotix Society, 2019. *Blob Detection.* [Online] Available at: https://medium.com/image-processing-in-robotics/blob-detection-309226a3ea5b
   [Accessed 22 May 2020].