



**AIN SHAMS UNIVERSITY**  
**FACULTY OF ENGINEERING**  
**Mechatronics Department**

## **Toward a Real Time Vehicle Detection System**

A Thesis submitted in partial fulfillment of the requirements of  
Master of Science (M.Sc.) Degree in Mechanical Engineering  
(Mechatronics Engineering Department)

by

**Mohamed Ashraf Ali Yousuf**

Bachelor of Science

(Mechatronics Engineering)

Faculty of Engineering, Ain Shams University, 2015

Supervised By

**Prof. Sherif Ali Mohamed Hammad**

**Prof. Ahmed H. Yousef**

**Dr. Hossam Eldin Hassan Abdelmunim**

Cairo, May 2019





**AIN SHAMS UNIVERSITY**  
**FACULTY OF ENGINEERING**  
**Mechatronics Engineering Department**

## **Toward a Real Time Vehicle Detection System**

by

**Mohamed Ashraf Ali Yousuf**

Bachelor of Science in Mechanical Engineering

(Mechatronics Engineering)

Faculty of Engineering, Ain Shams University, 2019

### **Examiners' Committee**

**Name and affiliation**

**Signature**

**Dr. Mohamed Abdallah Gaber**

Computer Science Department

Faculty of computer science, Nile University.

.....

**Prof. Hazem Mahmoud Abbas**

Computer Engineering and Systems Department

Faculty of Engineering, Ain Shams University.

.....

**Prof. Sherif Ali Hammad**

Mechatronics Department

Faculty of Engineering, Ain Shams University.

.....

**Dr. Hossam Eldin Hassan Abdelmunim**

Computer Engineering and Systems Department

Faculty of Engineering, University.

.....

Date:



# Statement

This thesis is submitted as a partial fulfillment of Choose Degree in Mechanical Engineering, Faculty of Engineering, Ain shams University. The author carried out the work included in this thesis, and no part of it has been submitted for a degree or a qualification at any other scientific entity.

**Mohamed Ashraf Ali Yousuf**

Signature

.....

**Date:**



# Researcher Data

**Name:** Mohamed Ashraf Ali Yousuf

**Date of Birth:** 28/3/1992

**Place of Birth:** Giza, Egypt

**Last academic degree:** Bachelor of Science

**Field of specialization:** Mechatronics Engineering

**University issued the degree :** Ain Shams University

**Date of issued degree :** July 2015

**Current job :** Demonstrator at Mechatronics Department, Faculty of Engineering,  
AIN SHAMS University



# Thesis Summary

Autonomous vehicles are one of the international future goals. Streets that have autonomous vehicles lead with efficiently developed software, can reduce the number of accidents caused by human factors. An efficient vision system that can identify the vehicle surrounding environments (Such as vehicles, traffic signs, road signs, etc.) is essential for the success of autonomous vehicles. Vehicle detection is one of the basic elements in the autonomous vehicle vision system. A lot of attention has been paid for this topic, but optimization is needed to reach a reliable detection system.

The objective of this thesis is to develop and implement a reliable vehicle detection system using computer vision techniques on Graphical Processing Units (GPUs), making a comparison between a classic vehicle detection system "Compass HOG" and a modern vehicle detection system "YOLO V2", making a comparison between vehicle detection systems on CPU and on GPU , implementing a GPU-based computer vision approach for vehicle detection by studying different techniques based on deep learning to make image classifications, in order to choose the best one that fits inside the different vehicle detection techniques that based on deep learning. These CNN architectures for vehicle classifications are ResNet, InceptionResnetV2, InceptionV3, NASNet, MobileNetV2, and PNASNet architectures. Also, there are two different datasets have been trained in these architectures to evaluate them. These datasets are Kitti dataset to train on car detection only, in additions to MIO-TCD dataset to detect various types of vehicles. For vehicle detection based on deep learning, Faster R-CNN, YOLO "You Only Look Once", and SSD "Single Shot Detector" have been compared in this thesis in speed an accuracy perspectives, in order to choose the best suitable one among the others.

Keywords: Deep-learning; CNN; ResNet; InceptionV3; Inception-ResnetV; MobileNetV2; NASNet; PNASNet; Vehicle-detection; classication; Kitti; MIO-TCD dataset; Faster R-CNN; YOLO; SSD.



# Acknowledgment

First of all, I am so grateful to Allah Almighty for his blessing and grace in giving me the ability to prepare this thesis.

I am really thankful to my thesis supervisor Dr. Hossam Eldin Hassan for his guidance and continuous support over the years, he is the most person caring about guide me to finish the thesis with excellent results. I am also immensely grateful to Prof. Ahmed Hassan, and Prof. Sherif Hammad, who are also my thesis supervisors for supporting me throughout my research with their effort and knowledge.

Without any doubt, this work would not have been possible without the aid of my colleagues; I would like to show my gratitude to Eng. Ahmed Abdelshakoor for their assistance and advice in the beginning part of the thesis. I also would like to thank Eng Yehia Zakaria, and Eng. Hanan Khairy for our successful association to make any needed comparisons. I would like also to thank Eng. Barihan Amr for helping me in my thesis and my publications to make them in the best shape. Lastly, I would like to especially thank my mother and my father for their love, care, support, and understanding during this work.

# Contents

<b>Contents</b>	<b>xi</b>
<b>List of Figures</b>	<b>xiii</b>
<b>List of Tables</b>	<b>xiv</b>
<b>Abbreviations</b>	<b>xv</b>
<b>Symbols</b>	<b>xvi</b>
<b>1 Introduction and Survey</b>	<b>1</b>
1.1 Problem Definition and Challenges	1
1.2 History and iterations of the automated and autonomous driving	5
1.3 History and iterations of the object detection, deep learning, and CNN architectures	8
1.4 Problem statement	11
1.5 Thesis contribution	13
1.6 The content of the thesis	15
<b>2 Neural Networks and Deep Learning Classification</b>	<b>1</b>
2.1 Artificial intelligence	17
2.2 Machine learning	18
2.2.1 Unsupervised learning and supervised learning	19
2.2.2 Classification	20
2.3 Neural Networks	21
2.4 Deep learning	23
2.4.1 Loss functions "Selective functions"	24
2.4.1.1 Mean Squared Error	25
2.4.1.2 Mean Squared Logarithmic Error	25
2.4.1.3 KL Divergence loss function	26
2.4.2 Activation functions	26
2.4.2.1 Step function	26
2.4.2.2 Linear function	27
2.4.2.3 Sigmoid function	28
2.4.2.4 Tanh function	29
2.4.2.5 ReLU function	29
2.4.2.6 Leaky ReLU function	30
2.4.2.7 Softmax function	31

2.4.3	Optimization Algorithms	31
2.4.3.1	Gradient Descent	32
2.4.3.2	Stochastic gradient descent	33
2.4.3.3	Mini Batch Gradient Descent	34
2.4.3.4	Momentum optimization function	34
2.4.3.5	Adagrad optimization function	35
2.4.3.6	Adam optimization function	35
2.4.3.7	Newton's method optimization function	36
2.4.3.8	Quasi-Newton method	37
2.4.4	OverFitting and Model regularization	37
2.4.4.1	L1 and L2 regularization	38
2.4.4.2	Early Stopping	39
2.4.4.3	Dropout	39
<b>3</b>	<b>CNN Architectures For Image Classification</b>	<b>42</b>
3.1	Convolutional Neural Networks	42
3.1.1	Convolutional Neural Networks vs Neural Networks in images	43
3.1.2	Convolution process	44
3.1.3	Pooling process	46
3.2	CNN architectures for image classification	47
3.2.1	InceptionV3 architecture	47
3.2.2	Inception-ResnetV2 architecture	49
3.2.3	ResNet architecture	51
3.2.4	MobileNetV2 architecture	52
3.2.5	NASNet architecture	54
3.2.6	PNASNet architecture	57
<b>4</b>	<b>Object Detection Algorithms</b>	<b>59</b>
4.1	Classification versus detection	59
4.2	Conventional object detection algorithms	60
4.2.1	HOG + SVM algorithm	60
4.2.2	Compass HOG	62
4.3	The generalized object detection algorithm based on deep learning	63
4.4	Faster R-CNN	64
4.4.1	Anchors	66
4.4.2	Region Proposal Network	68
4.4.3	Training process and loss functions	68
4.4.4	Post processing	69
4.4.5	Region of Interest Pooling	69
4.4.6	The base network for classification	70
4.5	Single Shot MultiBox Detector "SSD"	70
4.5.1	MultiBox	71
4.5.2	MultiBox with IOU	72
4.5.3	Hard Negative Mining	74
4.6	YOLO system "You only look once"	75
4.6.1	YOLO's Network design	79
4.6.2	Loss function	79

---

4.6.3	Non-maximal suppression	81
4.6.4	YOLOv2	81
4.6.5	YOLO V3	85
<b>5</b>	<b>Deep Learning Using GPU</b>	<b>87</b>
5.1	The needs of GPU	87
5.2	How GPU works	88
5.3	Why GPUs for CNN applications is better than CPUs?	89
5.4	GPUs features that are related to the deep learning	89
5.5	Tensorflow using GPU	90
5.6	Vehicle detection on CPU vs GPU	90
5.7	Implementing CNN architecture using TensorFlow	92
5.7.1	Convolution process using TensorFlow	92
5.7.2	Pooling process using TensorFlow	93
5.7.3	Dropout process using TensorFlow	93
5.7.4	Convolutional layer	94
5.7.5	Loss function and optimization function	94
<b>6</b>	<b>Experiments and Results</b>	<b>96</b>
6.1	Hardware and Software Requirements	96
6.2	Training on Kitti dataset	97
6.2.1	Compass HOG versus YOLOV2 on Kitti	98
6.2.2	CNN architectures on Kitti	98
6.3	Training on MIO-TCD dataset	101
6.4	Vehicle detection algorithms results	104
6.4.1	Detection results on challenges	104
6.4.2	Accuracy comparison results for different scales	111
6.4.3	Video comparison results	130
6.4.4	False positive samples	130
<b>7</b>	<b>Conclusion and Future Works</b>	<b>136</b>
	<b>Bibliography</b>	<b>138</b>

# List of Figures

1.1	The object detection inside the car cockpit to reveal how the object detection should work in real time inside the car. . . . .	2
1.2	The different challenges that will face the detection process. (A) Raining, (B) Night driving, (C) Driving through fog, (D) Noise on image, (E) Detect vehicle from side, (F) Vehicles on different scales. . . . .	3
1.3	The difference between the 5 levels of the Autonomous driving . . . . .	4
1.4	The Citroen DS car which has been made in the 1960s[2]. . . . .	6
1.5	The ParkShuttle at Netherlands[3]. . . . .	7
1.6	Google and Lexus driverless car [4]. . . . .	7
1.7	The components of the LeCan network to make the classification process [9]. . . . .	9
1.8	The components and the layers distribution of the AlexNet architecture[14].	10
1.9	The winners of the ImageNet competition for object detection and classification from 2011 to 2017. . . . .	10
1.10	The meaning of the height H and the width W inside the image. The H represents the number of columns and the W represents the number of rows inside the imported image. . . . .	11
1.11	The meaning of the height and width of the boundary box X,Y and shows the coordinates of the centroid of the drawn box $X_c$ and $Y_c$ . . . . .	13
2.1	The relationship between A.I and deep learning. As shown, the Deep learning is a part of the machine learning and AI. . . . .	18
2.2	The components of the neural network neuron. it shows how the neurons of the neural networks connect together to make the output . . . . .	22
2.3	An example shows the connections and the components of the deep neural network. . . . .	23
2.4	The role of the weights in the deep learning model. . . . .	24
2.5	The role of the loss function in the deep learning model. . . . .	25
2.6	shows the output of linear function on the left and its derivative on the right. . . . .	27
2.7	shows the output of sigmoid function on the left and its derivative on the right. . . . .	28
2.8	shows the output of Tanh function on the left and its derivative on the right. . . . .	29
2.9	shows the output of ReLU function on the right and its derivative on the left. . . . .	30
2.10	shows the output of leaky ReLU function on the right and its derivative on the left. . . . .	30
2.11	The role of the optimizer in the deep learning model. . . . .	32