

ระบบช่วยการขับขี่ยานยนต์ไฟฟ้าโดยใช้การประมวลผลภาพ  
ด้วยปัญญาประดิษฐ์  
Electric Vehicle Driving Assist System Using Image Processing  
by AI



ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

สาขาวิชาวิศวกรรมการวัดคุม

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2564

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# Electric Vehicle Driving Assist System Using Image Processing

by AI



BODINPHAT SIRITHANAWONGSA  
PHOTHISSET CHANTARAMONGKOL  
SIRIPHONG SAETAE

A THESIS UBMITTED IN PARTIAL FULFILLMENT  
OF THE REQUIREMENT FOR THE DEGREE OF  
BACHELOR OF ENGINEERING IN INSTRUMENTAION ENGINEERING  
FACLTY OF ENGINEERING  
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LARDKRABANG  
ACADEMIC YEAR 2021

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้


ปริญญาานิพนธ์ปีการศึกษา 2564  
คณะวิศวกรรมศาสตร์  
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง  
ใบรับรองปริญญาานิพนธ์

หัวข้อปริญญาานิพนธ์ ระบบช่วยการขับขี่ยานยนต์ไฟฟ้าโดยใช้การประมวลผลภาพด้วย  
ปัญญาประดิษฐ์

Electric Vehicle Driving Assist System Using Image  
Processing by AI

นักศึกษาผู้จัดทำ นายบดินทร์ภัทร์ ศิริธนาวงศา รหัสนักศึกษา 62015069  
นายโพธิ์เศรษฐ์ จันทรมงคล รหัสนักศึกษา 62015091  
นายศิริพงษ์ แซ่แต่ รหัสนักศึกษา 62015107

ปริญญา วิศวกรรมศาสตรบัณฑิต  
สาขาวิชา วิศวกรรมการวัดคุม  
ปีการศึกษา 2564

อาจารย์ผู้ควบคุมปริญญาานิพนธ์	ลายมือชื่อ
รศ.ดร.ทวีพล ชื่อสัตย์	
รศ.ร.ดร.วีระเชษฐ์ ชันเงิน	

หัวข้อวิทยานิพนธ์	ระบบช่วยการขับขี่ยานยนต์ไฟฟ้าโดยใช้การประมวลผลภาพด้วยปัญญาประดิษฐ์		
	Electric Vehicle Driving Assist System Using Image Processing by AI		
นักศึกษาผู้จัดทำ	นายบดินทร์ภัทร์	ศิริธนาวงศา	รหัสนักศึกษา 62015069
	นายโพธิเศรษฐ์	จันทรมงคล	รหัสนักศึกษา 62015091
	นายศิริพงษ์	แซ่แต่	รหัสนักศึกษา 62015107
อาจารย์ที่ปรึกษา	รศ.ดร.ทวีพล	ชื้อสตัย	
อาจารย์ที่ปรึกษาร่วม	รศ.ร.ดร.วีระเชษฐ์	ชันเงิน	
ปีการศึกษา	2564		

### บทคัดย่อ

วิทยานิพนธ์นี้นำเสนอระบบช่วยการขับขี่ยานยนต์ไฟฟ้าโดยใช้การประมวลผลภาพด้วยปัญญาประดิษฐ์ โดยสร้างระบบช่วยเหลือการขับขี่ยานยนต์ไฟฟ้าเพื่อการเกษตรกรรม หรือสามารถประยุกต์ไปใช้กับยานยนต์ไฟฟ้าชนิดอื่นได้ ระบบจะจำแนกชนิดของวัตถุบนท้องถนนผ่านกล้องและจะทำการส่งสัญญาณแจ้งเตือนพร้อมควบคุมเบรกที่กล้องควบคุมของยานยนต์ไฟฟ้า ซึ่งระบบนี้จะใช้บอร์ด CorgiDude ในการประมวลผลภาพ โดยนำข้อมูลที่รวบรวมได้ อัปโหลดขึ้นไป training ด้วยหลักการ Convolutional Neural Network หรือ CNN ที่ Google Colab สอนจนมีค่า Accuracy 0.94948 แล้วนำข้อมูลที่ training ได้ดาวน์โหลดไปยังตัวบอร์ด และใช้โปรแกรม MaixPy IDE ภายใตระบบปฏิบัติการ Windows 10 ตัวโปรแกรมจะทำการประมวลและแสดงผลภาพในรูปแบบของสัญลักษณ์รูปภาพ โดยใช้กล้อง OV2640 เป็นอุปกรณ์รับภาพ และมีชุดคำสั่งเพื่อให้รถหยุดเมื่อกล้องตรวจพบสิ่งกีดขวางหรือตรวจจับวัตถุที่ผ่านหน้ายานยนต์ไฟฟ้า เพื่อเสริมความปลอดภัยให้กับผู้ขับขี่ยานยนต์ไฟฟ้าเพื่อการเกษตรกรรม ผลการทดลองจากสิ่งกีดขวางจำนวน 4 รายการระบบทำงานได้อย่างถูกต้องแม่นยำและสามารถประยุกต์ใช้กับอุตสาหกรรมยานยนต์ไฟฟ้าได้เป็นอย่างดี

<b>Thesis Title</b>	Electric Vehicle Driving Assist System Using Image Processing by AI		
<b>Author</b>	Mr. Bodinphat	Sirithanawongsa	Student ID. 62015069
	Mr. Phothiset	Chantaramongkol	Student ID. 62015091
	Mr. Siriphong	Saetae	Student ID. 62015107
<b>Thesis Advisor</b>	Assoc.Prof.Dr.Taweepol	Suesut	
<b>Thesis Advisor</b>	Assoc.Prof.Dr.Werachet	Khanngern	
<b>Year</b>	2021		

## ABSTRACT

This project presents an electric vehicle driving assistance system using image processing by artificial intelligence as an electric vehicle driving assistance system for agriculture or can be applied to other types of electric vehicles. The vision system recognizes the type of objects on the road through the camera and sends a warning signal and brake control to the electric vehicle's control box. This system uses the CorgiDude board to process images by using the information collected Upload up to training with Convolutional Neural Network or CNN. The training set were taught by Google Colab until Accuracy 0.94948 and downloaded to the board using the MaixPy IDE program under the Windows 10 operating system. The camera OV2640 was used as an image acquisition device. The instructions for breaking the car to stop when the camera encounters an obstacle or detects an object passing in front of an electric vehicle has been developed to enhance the safety of drivers of electric vehicles for agriculture. Experimental results from 4 obstacle trials show the system works accurately and can be apply to the electric vehicle industry as well.

## กิตติกรรมประกาศ

ปริญญาบัตรฉบับนี้สำเร็จลุล่วงได้ด้วยความกรุณาเป็นอย่างสูงจากการให้คำปรึกษาจากอาจารย์ที่ปรึกษา ได้แก่ รองศาสตราจารย์ ดร.ทวีพล ชื้อสัตย์ รองศาสตราจารย์ร.ดร.วีระเชษฐ์ ชื่นเงิน และคณะอาจารย์ที่เป็นคณะกรรมการการสอบปริญญาบัตรทุกท่าน คำแนะนำในขั้นตอนที่ได้ทำการศึกษาปริญญาบัตร การวางแผนการศึกษา กระบวนการออกแบบผลงาน และการนำเสนอผลงานการออกแบบที่ถูกต้องตามหลักวิชา ตลอดจนการเขียนรายงานและรูปเล่มปริญญาบัตรเพื่อนำเสนอผลงาน

ขอขอบคุณคณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง อันเป็นสถานศึกษาที่ประสิทธิ์ ประสาทความรู้ ขอกราบขอบพระคุณคณาจารย์ประจำคณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ทุกท่านที่ได้ให้การสั่งสอนวิชาทางวิศวกรรม ให้คำแนะนำ และข้อคิดเห็นอันเป็นประโยชน์ต่อการทำปริญญาบัตร

ขอขอบคุณเจ้าหน้าที่และบุคลากรคณะวิศวกรรมศาสตร์ทุกท่านที่ให้ความช่วยเหลือและอำนวยความสะดวกในด้านต่างๆ ขอขอบคุณเพื่อนนักศึกษาคณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ทุกท่านที่ให้คำแนะนำและเป็นกำลังใจในการศึกษาตลอดมา

สุดท้ายผลอันจะเป็นประโยชน์ ความดีงามทั้งปวง ที่เกิดขึ้นจากการศึกษาปริญญาบัตรนี้ขอมอบให้แก่คุณพ่อและคุณแม่ที่เคารพยิ่งและหากมีข้อบกพร่องประการใดๆ ผู้วิจัยขอน้อมรับไว้ด้วยความขอบคุณยิ่ง

คณะผู้จัดทำ

# สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VII
สารบัญภาพ.....	IX
บทที่ 1 บทนำ.....	1
1.1 ความสำคัญของปริญญาโท.....	1
1.2 วัตถุประสงค์ของปริญญาโท.....	1
1.3 ขอบเขตของปริญญาโท.....	1
1.4 ขั้นตอนการศึกษา.....	2
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	2
บทที่ 2 ทฤษฎีและหลักการที่เกี่ยวข้อง.....	3
2.1 ยานยนต์ไฟฟ้า.....	3
2.1.1 ยานยนต์ไฟฟ้าไฮบริด.....	3
2.1.2 ยานยนต์ไฟฟ้าไฮบริดปลั๊กอิน.....	3
2.1.3 ยานยนต์ไฟฟ้าแบตเตอรี่.....	3
2.1.4 ยานยนต์ไฟฟ้าเซลล์เชื้อเพลิง.....	4
2.2 เทคโนโลยีนวัตกรรมยานยนต์.....	4
2.3 หลักการมองเห็นของมนุษย์กับการประมวลผลภาพดิจิทัล.....	5
2.3.1 ขั้นตอนการประมวลผลภาพดิจิทัล.....	6
2.3.2 การแปลงภาพดิจิทัล.....	7
2.3.3 ลักษณะของแฟ้มการเก็บภาพให้อยู่ในระบบข้อมูลแบบดิจิทัล.....	8
2.3.4 การเตรียมภาพก่อนการประมวลผลภาพ.....	9
2.3.5 วิธีการกรองแบบ Unsharp Filter.....	9
2.3.6 วิธีการแปลงภาพให้เป็นภาพไบนารี.....	10
2.3.7 การหาขอบภาพ.....	10

## สารบัญ(ต่อ)

2.3.8	การหาขอบภาพโดยใช้อุปกรณ์อันดับที่หนึ่ง .....	12
2.3.9	ประเภทของภาพ.....	15
2.3.9.1	ภาพระดับสีเทา .....	16
2.3.9.2	ภาพสี.....	16
2.1.9.3	ภาพสีเชิงซ้อน .....	17
2.1.9.4	ประเภทของไฟล์ภาพ (Image File).....	17
2.3.9.5	ตัวอย่างแนวคิดเครือข่ายความสัมพันธ์ระหว่างคำอธิบายความหมาย บนภาพ.....	18
2.3.9.6	การประมวลผลภาพดิจิทัล .....	19
2.3.9.7	การประมวลผลภาพระดับสูง.....	22
2.3.9.8	การจำแนกประเภทข้อมูลภาพ .....	22
2.4	การเรียนรู้ของเครื่อง .....	22
2.4.1	การเรียนรู้แบบมีผู้สอน .....	22
2.4.2	การเรียนรู้แบบไม่มีผู้สอน .....	23
2.5	การเรียนรู้เชิงลึก .....	23
2.5.1	หลักการทํางานโครงข่ายประสาทเทียม .....	24
2.5.1.1	ค่าน้ำหนัก.....	25
2.5.1.2	ค่าเบี่ยงเบน.....	25
2.5.1.3	ฟังก์ชันกระตุ้น .....	25
2.5.1.4	ฟังก์ชันสูญเสีย.....	25
2.5.1.5	แบ็คพรอพาเกชัน.....	26
2.5.1.6	เกรเดียนตีสเซนส์.....	26
2.6	โครงข่ายประสาทเทียมแบบคอนโวลูชัน.....	27
2.6.1	ชั้นคอนโวลูชันนอล.....	27
2.6.2	ชั้นพูลลิง.....	28
2.6.3	ชั้นการเชื่อมโยงสมบูรณ์ .....	28
2.7	การถ่ายโอนการเรียนรู้.....	29
2.8	การวัดประสิทธิภาพของการจำแนกประเภท.....	29

## สารบัญ(ต่อ)

<b>บทที่ 3</b>	<b>วิธีการดำเนินการวิจัย.....</b>	<b>32</b>
3.1	การออกแบบระบบ .....	32
3.2	ขั้นตอนการปฏิบัติงาน.....	32
3.3	ขั้นตอนการทำงานของ CorgiDude.....	34
3.4	โครงสร้างภายในของบอร์ด CorgiDude .....	34
3.4.1	การวางตำแหน่งอุปกรณ์ภายในกล่องควบคุม .....	35
3.5	โปรแกรม Colaboratory .....	36
<b>บทที่ 4</b>	<b>ผลการทดลอง .....</b>	<b>37</b>
4.1	ขั้นตอนการทดลอง.....	37
4.1.1	การเทรนนิ่งบอร์ด CorgiDude โดยใช้โปรแกรม Colaboratory .....	37
4.2	ผลการทดสอบโมเดลของเทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน .....	42
4.2.1	ผลการทดสอบประสิทธิภาพ .....	42
4.2.2	ผลการทดสอบหาค่าคะแนนเทรซโฮลด์ .....	43
4.2.3	ผลการทดสอบตารางคอนฟิวชันเมทริกซ์ ของเทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน .....	43
4.3	ผลการเทรนนิ่งตัวกล้องผ่านรูปภาพแต่ละประเภท .....	45
4.4	ทดสอบโดยนำอุปกรณ์มาติดตั้งที่รถแทรกเตอร์ไฟฟ้า .....	46
4.5	เปรียบเทียบผลการทดลอง .....	49
<b>บทที่ 5</b>	<b>สรุปผลการทดลอง.....</b>	<b>55</b>
5.1	สรุปผลการดำเนินงาน .....	55
5.2	ปัญหาและแนวทางในการแก้ไขปัญหา.....	55
5.2.1	ปัญหาที่พบ .....	55
5.2.2	แนวทางการแก้ปัญหา .....	56
5.3	ข้อเสนอแนะ.....	56

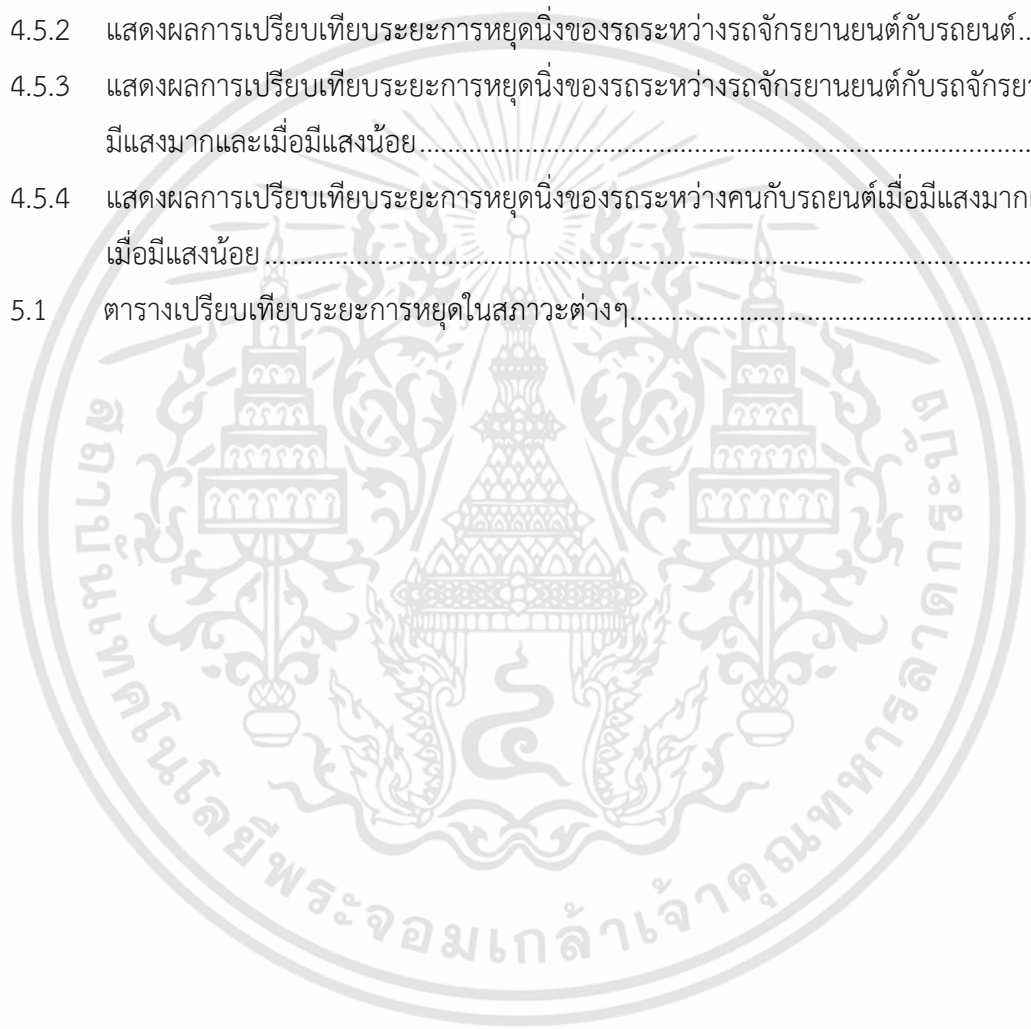
## สารบัญ(ต่อ)

บรรณานุกรม .....	58
ภาคผนวก.....	60



# สารบัญตาราง

ตารางที่	หน้า
2.3.9	แสดงจำนวนไบท์ที่ใช้ในแต่ละพิกเซลของภาพประเภทต่าง ๆ ..... 16
4.2a	คอนทราสต์ของเมทริกซ์ของโมเดลตรวจสอบความสมบูรณ์ของบรรจุภัณฑ์..... 44
4.2b	ผลลัพธ์ค่าความแม่นยำ, ความครบถ้วน, ค่าประสิทธิภาพโดยรวม ..... 44
4.5.1	แสดงผลการเปรียบเทียบระยะเวลาการหยุดนิ่งของรถระหว่างคนกับริดจักรยาน ..... 49
4.5.2	แสดงผลการเปรียบเทียบระยะเวลาการหยุดนิ่งของรถระหว่างรถจักรยานยนต์กับริดยนต์..... 49
4.5.3	แสดงผลการเปรียบเทียบระยะเวลาการหยุดนิ่งของรถระหว่างรถจักรยานยนต์กับริดจักรยานเมื่อมีแสงมากและเมื่อมีแสงน้อย..... 51
4.5.4	แสดงผลการเปรียบเทียบระยะเวลาการหยุดนิ่งของรถระหว่างคนกับริดยนต์เมื่อมีแสงมากและเมื่อมีแสงน้อย..... 52
5.1	ตารางเปรียบเทียบระยะเวลาการหยุดในสภาวะต่างๆ.....55



# สารบัญรูป

รูปที่		หน้า
2.3.1	อุปกรณ์พื้นฐานสำหรับระบบประมวลผลภาพ.....	6
2.3.3	เมตริกซ์ของรูปขนาด 256*256.....	8
2.3.5	วิธีการปรับปรุงคุณภาพโดยวิธีการ Unsharp.....	9
2.3.7	แสดงลักษณะขอบภาพแบบต่างๆ.....	11
2.3.7.1	แสดงขอบภาพชนิดต่างๆ.....	12
2.3.8	แสดงตัวดำเนินการหาขอบภาพของ Laplacian.....	15
2.3.9.2	แสดงระนาบที่บรรจุภาพสี.....	17
2.3.9.5	ตัวอย่างแนวคิดเครือข่ายความสัมพันธ์ระหว่างคำอธิบายความหมายบนภาพ.....	19
2.3.9.6	แบบจำลองของภาพด้วยระนาบ.....	19
2.3.9.7	การจัดเก็บและเรียกใช้ข้อมูลภาพดิจิทัลลงในเมตริกซ์.....	20
2.5ก	โครงสร้างของโครงข่ายประสาทเทียม.....	23
2.5ข	ความแตกต่างระหว่างการเรียนรู้ของเครื่องและการเรียนรู้เชิงลึก.....	24
2.5.1	ภาพตัวอย่างการทำงานของโครงข่ายประสาทเทียม.....	24
2.5.1.6	ภาพตัวอย่างกระบวนการเกรเดียนเดสเซนท์ สำหรับการหาค่าสูญเสียที่ต่ำที่สุด.....	26
2.6	โครงข่ายประสาทเทียมแบบคอนโวลูชัน.....	27
2.6.1	การทำงานของตัวกรองที่ใช้สำหรับการสร้างพีเจอร์แมพ.....	28
2.6.2	ตัวอย่างการลดขนาดของภาพด้วยวิธีพูลลิง.....	28
2.7	การใช้เทคนิคทรานเฟอร์เลิร์นนิ่ง.....	29
2.8	ตัวอย่างตารางคอนฟิวชันเมตริกซ์.....	29
3.1	บล็อกไดอะแกรมการทำงานของระบบ.....	32
3.2.1	การออกแบบโครงสร้างของรถโดยใช้โปรแกรม Sketch up.....	32
3.2.2	การตัดแปลงโครงสร้างของรถ.....	33
3.2.3	ติดตั้งระบบช่วยในการขับเคลื่อนยนต์ไฟฟ้า.....	33
3.2.4	ติดตั้งระบบ Inverter Controller และ ติดตั้งแบตเตอรี่.....	33
3.2.5	ติดตั้งระบบช่วยในการขับเคลื่อนยนต์ไฟฟ้าเสร็จสิ้นพร้อมทดสอบระบบ.....	33
3.2.6	Flow chart การทำงานของระบบ.....	34
3.3	การเชื่อมต่อวงจร.....	35
3.4	รายละเอียดภายในของบอร์ด Corgitude.....	36
3.4.1	การวางตำแหน่งของอุปกรณ์ภายในกล่อง.....	36

3.5	หน้าต่างของโปรแกรม Colaboratory.....	37
4.1.1	หน้าต่างเริ่มต้นของโปรแกรม.....	38
4.1.2	การติดตั้ง Library.....	39
4.1.3	การแตกไฟล์รูปภาพเข้าในโปรแกรม.....	39
4.1.4	การแบ่งไฟล์รูปภาพบนท้องถนนเพื่อไปทำการเทรนนิ่ง AI.....	40
4.1.5	ทำการเทรนนิ่งรูปภาพเพื่อให้ได้ out put.....	40
4.1.6	การ Resize ตัวข้อมูลในไฟล์ .kmodel.....	41
4.1.7	การนำข้อมูลที่เทรนใน Colab ออกมาเป็น ไฟล์ .Kmodel.....	41
4.1.8	นำไฟล์ที่เทรนและได้ Output ออกมา ทำการตรวจสอบข้อมูลในไฟล์.....	42
4.1.9	การสุ่มตัวอย่างข้อมูลในไฟล์มาตรวจสอบค่าความแม่นยำ.....	42
4.2ก	กราฟแสดงความแม่นยำเฉลี่ย (mAP).....	43
4.2ข	กราฟแสดงการสูญเสียทั้งหมด.....	43
4.2.2	กราฟความสัมพันธ์ระหว่างค่าร้อยละความสามารถในการทำนายและความถูกต้อง.....	44
4.3ก	ภาพคนที่นำมาเทรนนิ่ง.....	46
4.3ข	ภาพรถที่นำมาเทรนนิ่ง.....	46
4.3ค	ภาพรถจักรยานยนต์ที่นำมาเทรนนิ่ง.....	47
4.3ง	ภาพรถยนต์ที่นำมาเทรนนิ่ง.....	47
4.4.1	ทดสอบโดยนำรถจักรยานมาวางหน้ารถ.....	48
4.4.2	ทดสอบโดยให้มีคนมายืนมาวางหน้ารถ.....	48
4.4.3	ทดสอบโดยนำรถยนต์มาวางหน้ารถ.....	49
4.4.4	ทดสอบโดยนำรถจักรยานยนต์มาวางหน้ารถ.....	49
4.5.ก	กราฟแสดงผลระยะที่รถหยุดนิ่งเมื่อมีคนมาวางหน้ารถ.....	51
4.5.ข	กราฟแสดงผลระยะที่รถหยุดนิ่งเมื่อมีรถจักรยานมาวางหน้ารถ.....	51
4.5.ค	กราฟแสดงผลระยะที่รถหยุดนิ่งเมื่อมีรถจักรยานยนต์มาวางหน้ารถ.....	52
4.5.ง	กราฟแสดงผลระยะที่รถหยุดนิ่งเมื่อมีรถยนต์มาวางหน้ารถ.....	52
4.5.a	กราฟแสดงผลระยะที่รถหยุดนิ่งเมื่อมีคนมาวางหน้ารถในสภาวะแสงต่างๆ.....	53
4.5.b	กราฟแสดงผลระยะที่รถหยุดนิ่งเมื่อมีรถจักรยานมาวางหน้ารถในสภาวะแสงต่างๆ.....	54
4.5.c	กราฟแสดงผลระยะที่รถหยุดนิ่งเมื่อมีรถยนต์มาวางหน้ารถในสภาวะแสงต่างๆ.....	54
4.5.d	กราฟแสดงผลระยะที่รถหยุดนิ่งเมื่อมีรถจักรยานยนต์มาวางหน้ารถ ในสภาวะแสงต่างๆ .....	55

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
 X  
 ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# บทที่ 1

## บทนำ

### 1.1 ความสำคัญของปัญญาประดิษฐ์

เทคโนโลยียานยนต์ที่ขับเคลื่อนด้วยพลังงานไฟฟ้าหรือยานยนต์ไร้คนขับ กำลังเป็นที่สนใจในประเทศไทย และเป็นอีกหนึ่งนวัตกรรมที่กำลังได้รับความสนใจจากทั่วโลก จึงมีการนำ AI (Artificial Intelligence) หรือ ระบบปัญญาประดิษฐ์ เข้ามาเพิ่มประสิทธิภาพของรถยนต์ที่ขับเคลื่อนด้วยไฟฟ้า เพื่อให้เกิดความสะดวกในการใช้งานรวมถึงระบบป้องกันอุบัติเหตุเพื่อความปลอดภัยของผู้ขับขี่ โดยการใช้ คอมพิวเตอร์วิชั่น หรือระบบที่ทำหน้าที่เป็นหูเป็นตาให้กับรถยนต์ไฟฟ้าในการตรวจจับการเปลี่ยนแปลงของสภาพแวดล้อมเมื่อรถวิ่ง นำข้อมูลมาประมวลผลโดยวิธีการเรียนรู้เชิงลึกหรือระบบการจำลองรูปแบบการประมวลผลของสมองมนุษย์ โดยใช้โครงข่ายคล้ายเซลล์ประสาทในการประมวลผล เป็นกระบวนการจดจำ ทำความเข้าใจ ตัดสินใจ และแก้ไขปัญหา โดยอาศัยข้อมูลจำนวนมากที่มีลักษณะซ้ำ ๆ จากการประมวลผลข้อมูลที่ได้รับจากระบบ คอมพิวเตอร์วิชั่นนั่นเอง

ปัจจุบันได้มีการใช้รถไฟฟ้าอย่างแพร่หลาย ทั้งในด้านการขนส่งและในด้านอุตสาหกรรม การเกษตร ไม่ว่าจะนำไปขับขี่ในวัตถุประสงค์ไหนก็ล้วนใช้มนุษย์ในการขับขี่เพื่อบรรลุวัตถุประสงค์นั้นๆ ดังนั้นการเกิดอุบัติเหตุทางท้องถนนจึงมักเกิดจากมนุษย์ทั้งสิ้น ไม่ว่าจะเกิดจากความประมาทหรือความตั้งใจ ผลที่ตามมาคือความเสียหายทางวัตถุหรือเกิดการสูญเสียขึ้น ผู้ศึกษาได้เล็งเห็นถึงปัญหาความผิดพลาดที่เกิดจากมนุษย์ด้านการขับขี่ สามารถลดปัญหาเหล่านี้ได้ด้วยระบบคอมพิวเตอร์วิชั่นคู่กับปัญญาประดิษฐ์ (Artificial Intelligence) มาพัฒนาระบบที่ช่วยในการควบคุมรถยนต์ไฟฟ้าเพื่อความปลอดภัยในการขับขี่ แก้ไขปัญหาด้วยเทคโนโลยีที่ทันสมัย

### 1.2 วัตถุประสงค์ของปัญญาประดิษฐ์

1. เพื่อพัฒนาระบบเบรกยานยนต์ไฟฟ้าอเนกประสงค์เพื่อเกษตรกรรมแบบอัตโนมัติโดยทำการประมวลผลด้วยเทคนิคการเรียนรู้เชิงลึก
2. เพื่อลดการเกิดอุบัติเหตุจากการขับขี่บนท้องถนน โดยใช้ระบบคอมพิวเตอร์วิชั่น
3. เพื่อประยุกต์ใช้กับยานยนต์ไฟฟ้าอเนกประสงค์เพื่อเกษตรกรรมและปรับเปลี่ยนให้เข้ากับวิถีชีวิตเกษตรกรรมรูปแบบใหม่ (New Normal)

### 1.3 ขอบเขตของปัญญาประดิษฐ์

เป็นการใช้ความรู้เกี่ยวกับระบบคอมพิวเตอร์วิชั่น ใช้สำหรับการประมวลผลภาพโดยใช้กล้อง OV2640 เป็นอุปกรณ์รับภาพใช้ในการจับภาพระยะและรูปร่างวัตถุ นำไปประมวลผลในบอร์ด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

CorgiDude ที่ผ่านการ Training ด้วยหลักการ Convolutional Neural Network (CNN) ใน Google Colab จากรวบรวมรูปภาพวัตถุที่จำนวนมาก จนมีความแม่นยำในการแยกแยะสูงก่อนที่จะดาวน์โหลดไปยังบอร์ด เพื่อทำการแยกแยะประเภทของวัตถุ อาทิเช่น รถยนต์ รถจักรยาน รถจักรยานยนต์ รวมถึงมนุษย์ แล้ววิเคราะห์ของวัตถุ จากนั้นประมวลผลออกมาเป็นสัญญาณดิจิทัล นำสัญญาณไปควบคุมระบบการเบรกของตัวยานยนต์ไฟฟ้าอเนกประสงค์เพื่อเกษตรกรรมพร้อมกับส่งสัญญาณเตือนผู้ขับขี่ ระบบนี้จะพัฒนาร่วมกับยานยนต์ไฟฟ้าอเนกประสงค์เพื่อเกษตรกรรมโดยระบบนี้จะเข้าไปแก้ปัญหาเรื่องความปลอดภัยของยานยนต์ไฟฟ้าเพื่อลดอุบัติเหตุที่เกิดจากความประมาทของผู้ขับขี่

#### 1.4 ขั้นตอนการศึกษา

1. ศึกษาหลักการทำงานของกล้องกับการเชื่อมต่อข้อมูลคอมพิวเตอร์วิชั่น
2. ศึกษาการเรียนรู้เชิงลึกของระบบคอมพิวเตอร์วิชั่นในการประมวลผลรูปภาพ
3. ศึกษาหลักการ Convolutional Neural Network (CNN)
4. ศึกษาการเขียนโปรแกรม Python
5. ศึกษาการแยกแยะรูปภาพของบอร์ด CorgiDude

#### 1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. ควบคุมระบบเบรกของรถยนต์ไฟฟ้าอเนกประสงค์เพื่อเกษตรกรรมได้อย่างมีประสิทธิภาพ
2. ลดการเกิดอุบัติเหตุโดยตรงกับผู้ใช้งาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 2

# ทฤษฎีและหลักการที่เกี่ยวข้อง

### 2.1 ยานยนต์ไฟฟ้า (Electric Car)

รถยนต์ที่ขับเคลื่อนด้วยมอเตอร์ไฟฟ้าซึ่งเป็นพลังงานไฟฟ้าจะเก็บอยู่ในแบตเตอรี่หรืออุปกรณ์เก็บพลังงานไฟฟ้าแบบอื่นๆ และด้วยข้อดีของมอเตอร์ไฟฟ้าที่ให้แรงบิดได้ทันทีทำให้รถพลังงานไฟฟ้ามี อัตราเร่งที่เรียบและรวดเร็ว โดยจะมีชุดอุปกรณ์หลักคือ ตัวแบตเตอรี่ที่มีความจุพลังงานไฟฟ้าสูงซึ่งเก็บไว้ในรถ และตัวมอเตอร์ไฟฟ้าที่ทำหน้าที่ในการขับเคลื่อนโดยใช้พลังงานจากแบตเตอรี่ ส่วนมากเรื่องของระยะเวลาและระยะทางของรถไฟฟ้าจะมากขึ้นอยู่กับขนาดความจุของแบตเตอรี่เป็นสำคัญ ประเภทของรถยนต์ไฟฟ้าสามารถแบ่งยานยนต์ไฟฟ้าออกเป็น 4 ประเภท คือ

#### 2.1.1 ยานยนต์ไฟฟ้าไฮบริด (Hybrid Electric Vehicle, HEVs)

ประกอบด้วยเครื่องยนต์ลูกสูบเป็นต้นกำลังในการขับเคลื่อนหลัก ซึ่งใช้เชื้อเพลิงที่บรรจุในยานยนต์และทำงานร่วมกับมอเตอร์ไฟฟ้าเพื่อเพิ่มกำลังของยานยนต์ให้เคลื่อนที่ ทำให้เครื่องยนต์มีประสิทธิภาพสูงขึ้น จึงมีความสิ้นเปลืองเชื้อเพลิงต่ำกว่ายานยนต์ปกติ กำลังที่ผลิตจากเครื่องยนต์และมอเตอร์ไฟฟ้า ทำให้อัตราเร่งของยานยนต์สูงกว่ายานยนต์ที่มีเครื่องยนต์ลูกสูบขนาดเดียวกัน รวมทั้งยังสามารถนำพลังงานกลที่เหลือหรือไม่ใช้ประโยชน์เปลี่ยนเป็นพลังงานไฟฟ้าเก็บในแบตเตอรี่

#### 2.1.2 ยานยนต์ไฟฟ้าไฮบริดปลั๊กอิน (Plug-in Hybrid Electric Vehicle, PHEVs)

เป็นยานยนต์ไฟฟ้าที่พัฒนาต่อมาจากยานยนต์ไฟฟ้าไฮบริด โดยสามารถประจุพลังงานไฟฟ้าได้จากแหล่งภายนอก (Plug-in) ทำให้อานยนต์สามารถใช้พลังงาน พร้อมกันจาก 2 แหล่ง จึงสามารถวิ่งในระยะทางและความเร็วที่เพิ่มขึ้นด้วยพลังงานจากไฟฟ้าโดยตรง ยานยนต์ไฟฟ้า แบบ PHEV มีการออกแบบอยู่ 2 ประเภท ได้แก่ แบบ Extended range EV (EREV) และแบบ Blended PHEV โดย แบบ EREV จะเน้นการทำงานโดยใช้พลังงานไฟฟ้าเป็นหลักก่อน แต่แบบ Blended PHEV มีการทำงานผสมผสานระหว่างเครื่องยนต์และไฟฟ้า ดังนั้น ยานยนต์ไฟฟ้าแบบ EREV สามารถวิ่งด้วยพลังงานไฟฟ้าอย่างเดียวนานกว่าแบบ Blended PHEV

#### 2.1.3 ยานยนต์ไฟฟ้าแบตเตอรี่ (Battery Electric Vehicle, BEVs)

เป็นยานยนต์ไฟฟ้าที่มีเฉพาะมอเตอร์ไฟฟ้าเป็นต้นกำลังให้ยานยนต์เคลื่อนที่ และใช้พลังงานไฟฟ้าที่อยู่ในแบตเตอรี่เท่านั้น ไม่มีเครื่องยนต์อื่นในยานยนต์ ดังนั้นระยะทางการวิ่งของยานยนต์จึงขึ้นอยู่กับการออกแบบขนาดและชนิดของแบตเตอรี่ รวมทั้งน้ำหนักบรรทุก อย่างไรก็ตาม ปัจจุบันบริษัทรถยนต์ได้มีการผลิตและจำหน่ายยานยนต์ไฟฟ้าแบตเตอรี่ขึ้น ในประเทศพัฒนาแล้วเช่น เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ญี่ปุ่น ยุโรป และ สหรัฐอเมริกา เป็นต้น ทำให้เทคโนโลยีรถยนต์ไฟฟ้าแบตเตอรี่มีความเป็นไปได้มากขึ้น

#### 2.1.4 ยานยนต์ไฟฟ้าเซลล์เชื้อเพลิง (Fuel Cell Electric Vehicle, FCEVs)

เป็นยานยนต์ไฟฟ้าที่มีเซลล์เชื้อเพลิง (Fuel cell) ที่สามารถผลิตพลังงานไฟฟ้าได้โดยตรง รถยนต์ไฟฟ้าเซลล์เชื้อเพลิงมีข้อดีหลายๆ ประการ ข้อดีที่สำคัญที่สุดคือ ประสิทธิภาพของเซลล์เชื้อเพลิงมีค่าสูงถึง 60% และความจุพลังงานจำเพาะที่สูงกว่าแบตเตอรี่ที่มีอยู่ในปัจจุบัน รถยนต์ไฟฟ้าเซลล์เชื้อเพลิงจึงเป็นเทคโนโลยีที่บริษัทรถยนต์เชื่อว่าเป็นคำตอบที่แท้จริงของพลังงานสะอาดในอนาคต อย่างไรก็ตามก็ยังมีข้อจำกัดในเรื่องการผลิตไฮโดรเจนและโครงสร้างพื้นฐาน

## 2.2 เทคโนโลยีนวัตกรรมยานยนต์อัจฉริยะ

ปัจจุบันและในอนาคตจะเต็มไปด้วยความปลอดภัยอย่างสูงสุด ด้วยเทคโนโลยีรถยนต์อัจฉริยะ ที่ไม่ต้องการคนขับ โดยรถอัจฉริยะจะมีระบบตรวจวัดหลายประเภทติดที่ตัวรถ และส่งผ่านไปเครื่องสื่อสาร แล้วนำมาประมวลผลคำนวณเส้นทางและควบคุมการเคลื่อนที่ของรถยนต์อัจฉริยะ ซึ่งได้พัฒนามาจาก Artificial Intelligence หรือ AI มนุษย์เราทุกวันนี้สามารถออกแบบและสร้างให้รถยนต์มีความสามารถที่จะทำการขับเคลื่อนได้อย่างชาญฉลาด ทำการเลี้ยวซ้ายเลี้ยวขวาได้ดังใจต้องการ แถมด้วยมีการติดตั้งระบบอัจฉริยะป้องกันอุบัติเหตุและป้องกันการสูญเสียนั้น คงจะไม่ใช่ว่าเรื่องยากเกินจินตนาการของมนุษย์ในยุคสมัยนี้อีกต่อไป แต่ความเป็นจริงแล้วกว่าจะสร้างฝันสู่ความเป็นจริงก็ไม่ใช่ง่ายนัก หากแต่ความพยายามของคนไม่เคยหยุดนิ่ง ดังนั้นแล้วจึงเกิดเทคโนโลยีและนวัตกรรมใหม่ รถยนต์อัจฉริยะ (Intelligent Vehicle) ขึ้นมา ซึ่งมีระบบหลักการทำงาน และการควบคุม โดยเริ่มต้นตั้งแต่ รับข้อมูลจากอุปกรณ์ตรวจวัดประเภทต่างๆ ที่ติดตั้งมาบนรถ (Various Sensors) เช่น อุปกรณ์ตรวจวัดตำแหน่งปัจจุบันของรถ, อุปกรณ์ตรวจวัดตำแหน่งของรถคันอื่นที่เคลื่อนที่ไปในทิศทางเดียวกันหรือสวนทางมา, อุปกรณ์ตรวจวัดสิ่งกีดขวางบนเส้นทางการเคลื่อนที่ ทั้งที่อยู่กับที่และเคลื่อนที่, อุปกรณ์ตรวจวัดเส้นแบ่งบนเส้นทาง สัญลักษณ์ หรือสัญญาณจราจรอื่นๆ, อุปกรณ์การสื่อสารระหว่างรถอัจฉริยะด้วยกัน ต่อจากนั้นจะนำข้อมูลจากอุปกรณ์เหล่านี้มาประมวลผล (Data Processing) เพื่อกำหนดเส้นทางการเคลื่อนที่ และควบคุมการเคลื่อนที่ของรถยนต์โดยอัตโนมัติ และมีความปลอดภัย

ระดับของรถไร้คนขับ มาจากที่ทาง Society of Automotive Engineering (SAE) หรือ สมาคมวิศวกรยานยนต์ ได้มีการแบ่งระดับการขับอัตโนมัติตามมาตรฐาน J3016 ไว้ซึ่งการแบ่งนี้จัดทำเป็น 6 ระดับ

เริ่มจากระดับ 0

ในระดับนี้ การควบคุมตัวรถ และการตอบสนองต่อสภาวะฉุกเฉิน ผู้ขับยังคงต้องรับผิดชอบเองทั้งหมด แต่จะมีเทคโนโลยีเข้ามาช่วยให้การทำงานของระบบต่างๆ ดีขึ้น ตัวอย่างเช่น ระบบเบรค

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ABS ที่เรายังคงต้องเหยียบเบรคเอง แต่การทำงานของเบรคจะดีขึ้นกว่าเดิม เพื่อป้องกันไม่ให้เราเหยียบเบรคจนล้อล๊อค เหมือนการเบรคธรรมดานั่นเอง

#### ระดับ 1

ในระดับนี้ การควบคุมรถจะมีระบบเข้ามาช่วย โดยใช้การใช้ระบบประมวลผล ซึ่งเก็บข้อมูลจากอุปกรณ์ต่างๆ เช่น กล้องหน้ารถ หรือเซนเซอร์ ก่อนที่จะแปรเป็นสัญญาณเตือน ตัวอย่างที่ชัดเจนก็เช่น ระบบแจ้งเตือนรถออกนอกเลน หรือ Lane-keep Assist ที่จะทำงานโดยการสั่นพวงมาลัย หรือหักพวงมาลัยเล็กน้อย เพื่อเตือนเมื่อรถเราออกนอกเลน

#### ระดับที่ 2

ในระดับนี้รถของเราจะสามารถเคลื่อนไปข้างหน้าได้เอง โดยอาศัยการทำงานของระบบ แต่เรายังต้องนั่งอยู่ด้านหลังของพวงมาลัย เพื่อคอยระวังเมื่อเกิดเหตุฉุกเฉิน เหตุฉุกเฉินที่ว่าก็อย่างการมีรถพุ่งเข้ามาหา หรือคนตัดหน้า รวมทั้งการทำงานผิดพลาดของระบบด้วย ซึ่งตัวอย่างของระดับนี้ก็เช่น ระบบควบคุมความเร็วแบบ Adaptive Cruise Control หรือ บางค่ายอาจเรียกว่า Active Cruise Control

#### ระดับที่ 3

เป็นระดับเริ่มต้นของ ระบบขับเคลื่อนอัตโนมัติ นั่น ซึ่งทั้งการเคลื่อนที่ของรถ และการตอบสนองต่อวัตถุภายนอกนั้น ระบบจะทำหน้าที่จัดการทั้งหมด ผู้ขับขี่มีหน้าที่แค่นั่งอยู่หลังพวงมาลัย และต้องเข้ามาขับเองเมื่อระบบแจ้งเตือนเข้ามา ซึ่งตัวอย่างก็เช่น ระบบ Autopilot ของ Tesla หรือ CoPilot ของ BMW

#### ระดับที่ 4

เป็นระดับที่ค่ายรถหลายค่าย กำลังจะก้าวไปให้ถึง และมีกระแสว่าบางค่ายกำลังทุ่มพัฒนา ระบบขับเคลื่อนอัตโนมัติ ในระดับนี้ให้สำเร็จอยู่ ซึ่งในระดับที่ 4 นี้ จะเหมือนกับระดับที่ 3 ทุกอย่าง แต่สิ่งที่แตกต่างกันคือ รถจะสามารถวิ่งได้เองในเส้นทางที่ตั้งไว้ โดยไม่ขอความช่วยเหลือจากเราเลย ตัวอย่างเช่น แท็กซี่ไร้คนขับ หรือรถที่วิ่งได้เองโดยไม่ต้องมีพวงมาลัย คันเร่ง และเบรค

#### ระดับที่ 5

ที่สุดของระบบขับเคลื่อนอัตโนมัติ นั่น คือมันสามารถเคลื่อนที่ไปที่ไหนก็ได้ โดยที่ไม่ต้องมีคนขับ แถมยังสามารถเรียนรู้เส้นทางได้เองรถจะทำงานเหมือนไมโคร แต่เป็นผู้ช่วยที่รู้ใจของเรา เพียงแค่เราเอ่ยปากว่าจะไปจุดนั้นจุดนี้ (หรือมากกว่านั้นคือแค่คิดว่าจะไปจุดไหน) มันก็สามารถพาเราไปได้โดยอัตโนมัติประสิทธิภาพด้านการเรียนรู้ของรถยนต์ไร้คนขับระดับนี้ จะเหมือนกับคนที่มีการประสบการณ์ขับนับล้านชั่วโมง เพื่อกลายเป็นนักขับที่เก่งกาจและหลีกเลี่ยงการเกิดอุบัติเหตุมากที่สุด

### 2.3 หลักการมองเห็นของมนุษย์กับการประมวลผลภาพดิจิทัล

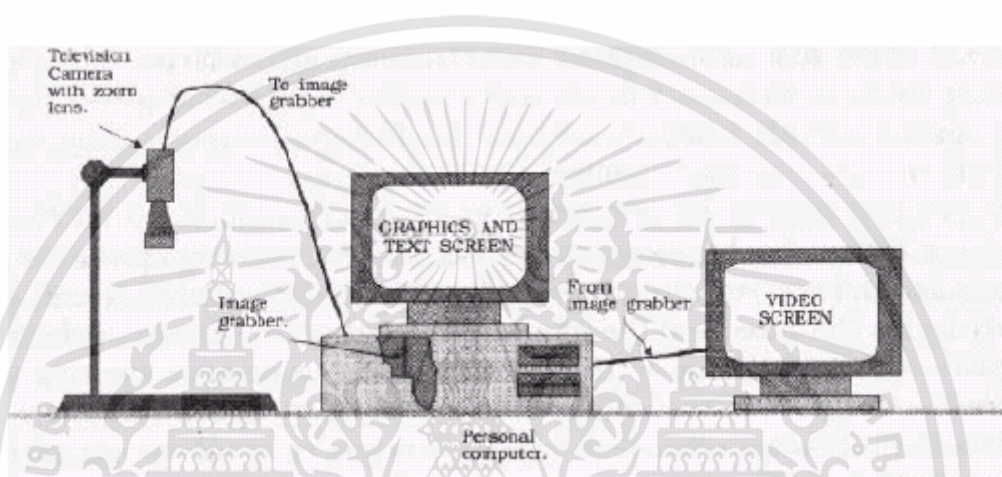
ระบบการมองเห็นของมนุษย์ หากเปรียบเทียบกับปัจจุบันก็คือกล้องถ่ายภาพดิจิทัล และ กล้องวิดีโอ อุปกรณ์เหล่านี้ทำหน้าที่ในการแปลงข้อมูลภาพให้เป็นสัญญาณทางไฟฟ้า ในการเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประมวลผลภาพดิจิทัลนิยมใช้กล้องถ่ายภาพดิจิทัล หรือ กล้องวิดีโอ เนื่องจากให้ภาพที่คมชัด และเหมือนจริงจากนั้นนำข้อมูลภาพที่ได้มาทำการวิเคราะห์ต่อไป

### 2.3.1 ขั้นตอนการประมวลผลภาพดิจิทัล

ขั้นตอนการประมวลผลภาพจะประกอบไปด้วย 3 ส่วน

1. ขั้นตอนการนำข้อมูลภาพเข้าสู่เครื่องคอมพิวเตอร์
2. ขั้นตอนการนำคอมพิวเตอร์มาพัฒนาอัลกอริทึมเพื่อประมวลผลภาพ
3. ขั้นตอนการแสดงผลภาพ



รูปที่ 2.3.1 อุปกรณ์พื้นฐานสำหรับระบบประมวลผลภาพ

โดยทั่วไปอุปกรณ์พื้นฐานในการประมวลผลภาพตามขั้นตอนที่กล่าวมาข้างต้นจะประกอบด้วย เครื่องไมโครคอมพิวเตอร์ที่มีอุปกรณ์ตรวจจับภาพ (frame grabber card) ซึ่งจะต่อกับกล้องวิดีโอตัวจับภาพ ดังรูปที่ 2.3.1 ขั้นตอนการนำภาพเข้าสู่เครื่องคอมพิวเตอร์เป็นการแปลงสัญญาณภาพทั่วไปเป็นสัญญาณอนาล็อก (analog signal) ให้เป็นสัญญาณภาพดิจิทัล เพื่อให้เครื่องคอมพิวเตอร์สามารถประมวลผลภาพได้ ทั้งนี้สามารถทำได้นำข้อมูลภาพจากตัวจับภาพ ซึ่งโดยทั่วไปแล้วภาพดิจิทัลที่ได้จากการแปลงสัญญาณจะถูกเก็บในหน่วยความจำที่มีอยู่ใน อุปกรณ์ตัวจับภาพ และสามารถนำมาประมวลผลได้โดยการเขียนโปรแกรมภาษาระดับสูง อาทิเช่นภาษาซีอุปกรณ์ตัวจับภาพที่นิยมใช้จะสามารถเก็บภาพขนาด  $512 * 512$  จุดภาพ(pixel) และ แต่ละจุดภาพสามารถแสดงระดับเทาได้อย่างน้อย 256 ระดับ ส่วนกล้องวิดีโอที่ใช้ในการเก็บภาพ มักจะเป็นกล้องซีซีดี (charge coupled device (CCD) camera)

ก่อนการประมวลผลภาพต้องเตรียมข้อมูลภาพซึ่งมีขั้นตอนดังนี้

1. ปรับปรุงข้อมูลภาพในส่วนที่ไม่คมชัด
2. กำจัดสัญญาณรบกวน
3. การปรับคอนทราสต์ หรือปรับปรุงคุณสมบัติของการมองเห็น เช่น การปรับค่าความ

เข้ม ของจุดภาพ การกลับข้อมูลภาพ เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. การแบ่งแยกข้อมูลภาพออกจากสีพื้น โดยใช้ค่าเทรชโฮลด์
5. การแปลงข้อมูลภาพในทางเรขาคณิต เช่น การหมุนภาพ การเปลี่ยนแปลงขนาดภาพ และการแก้ไขตำแหน่งของจุดภาพ
6. การแก้ไขข้อมูลภาพในส่วนที่มีการผิดเพี้ยน

### 2.3.2 การแปลงภาพดิจิทัล

สิ่งมองเห็นที่เป็นภาพทั่วไปจะเป็นลักษณะสัญญาณแบบอนาล็อก ซึ่งจะสามารถกำหนดได้เป็นฟังก์ชันความเข้ม 2 มิติ (two-dimensional light intensity function)  $f(x,y)$  โดยที่  $X$  และ  $y$  เป็นตำแหน่งพิกัดของภาพ ส่วนค่าของ  $f$  ที่ตำแหน่ง  $(x,y)$  นั้นจะเป็นสัดส่วนกับความเข้ม (brightness) หรือ ระดับเทา (gray level) ของภาพที่ตำแหน่งนั้น ซึ่งค่า  $f(x,y)$  จะมีความมากกว่า ศูนย์ กล่าวคือ

$$0 < f(x,y) < \infty \quad (2-1)$$

โดยทั่วไปค่าระดับความเข้มของแสง  $f(x,y)$  จะขึ้นกับองค์ประกอบสองส่วนคือ 1) ค่า illumination ซึ่งเป็นค่าระดับแสงที่มองเห็น และ 2) ค่า reflection ซึ่งค่าระดับที่สะท้อนมาจากวัตถุ ซึ่งค่าทั้งสองนี้อาจกำหนดให้เป็นฟังก์ชัน  $i(x,y)$  และ  $F(x,y)$  ตามลำดับ และค่า  $f(x,y)$  จะเป็นผลคูณของค่าทั้งสอง

$$f(x,y) = i(r,y)r(x,y) \quad (2-2)$$

โดยที่

$$0 < i(x,y) < \infty \quad (2-3)$$

และ

$$0 < r(x,y) < 1 \quad (2-4)$$

ค่าโดยเฉลี่ยสำหรับสภาพสิ่งแวดล้อมต่างๆ  $i(x,y)$  และ  $F(x,y)$  สำหรับวัตถุ(วัสดุ) ต่างๆ จะเป็นดังนี้

0.01	black velvet
0.65	stainless steal

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

	$r(x,y) =$	}	0.80	flat white paint
0.90	silver plated metal			
0.93	snow			

และ

9000	clear sunny day	}	0.10	full moon night
1000	cloudy day			
100	inside office			

โดยทั่วไปเราจะกำหนดค่าระดับความเข้มของแสงในกรณีของภาพโมนอโครม (monochrome image)  $f$  ที่พิกัด  $(x,y)$  เป็นค่าระดับเทา ซึ่งจากสมการที่ (3-1) ถึง (3-4) เราจะได้ค่า ดังนี้  $L_{min} \leq L \leq L_{max}$

$$L_{min} \leq L \leq L_{max} \quad (2-5)$$

โดยที่  $L_{min}$  จะมีค่าเป็น  $L_{min} = i_{min} r_{min} \approx 0.005$  และ  $L_{max}$  จะมีค่าเป็น  $L_{max} = i_{max} r_{max} \approx 100$

### 2.3.3 ลักษณะของแฟ้มการเก็บภาพให้อยู่ในระบบข้อมูลแบบดิจิทัล

ลักษณะของแฟ้มข้อมูลภาพจะมีการกำหนดตำแหน่งเหมือนกับเมทริกซ์โดยที่จำนวนคอลัมน์ (Column) และแถว (Row) จะแทนจำนวนจุดภาพของภาพเช่น ภาพขนาด  $256 \times 256$  หมายถึงแฟ้มข้อมูลของภาพดังกล่าวมีขนาด 256 คอลัมน์และ 256 แถว

$$A = \begin{bmatrix} f(1,1) & f(1,2) & f(1,3) & \dots & f(1,256) \\ f(2,1) & f(2,2) & f(2,3) & \dots & f(2,256) \\ f(3,1) & f(3,2) & f(3,3) & \dots & f(3,256) \\ \dots & \dots & \dots & \dots & \dots \\ f(256,1) & f(256,2) & f(256,3) & \dots & f(256,256) \end{bmatrix}$$

รูปที่ 2.3.3 เมทริกซ์ของรูปขนาด  $256 \times 256$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 2.3.3 เมทริกซ์ของรูปขนาด  $256 \times 256$  จากรูปที่ 2.1.3 จะเห็นว่าเพิ่มข้อมูลภาพจะประกอบไปด้วยสมาชิกของเมทริกซ์ ทั้งหมด  $i, j$  ตัว และที่ตำแหน่งในเมทริกซ์ต่าง ๆ ก็จะมี  $f(t,j)$  ซึ่งเป็นค่าระดับความเข้ม (Gray Level) ของภาพที่จุดนั้น

### 2.3.4 การเตรียมภาพก่อนการประมวลผลภาพ

การเตรียมภาพเป็นขั้นตอนก่อนการแปลงภาพให้เป็นภาพไบนารี ส่วนใหญ่จะเป็นกรองภาพ เพื่อกำจัดการรบกวน (Noise) หรือทำให้ภาพมีคุณภาพที่ดีขึ้นเมื่อไปเป็นภาพไบนารี

- การเปลี่ยนแปลงฮิสโตแกรม (Histogram Equalize)

เป็นการแปลงฮิสโตแกรมของภาพให้มีค่ากว้างขึ้นผลของการแปลงทำให้ภาพมีความเข้มที่ดีขึ้น เทคนิคเช่นนี้นิยมมาใช้กับภาพที่ต้องการดูรายละเอียดของภาพที่ซ่อนอยู่โดยไม่ คำนึงถึงสีของภาพ

- การเปลี่ยนแปลงคุณภาพอัตโนมัติ (Auto Level)

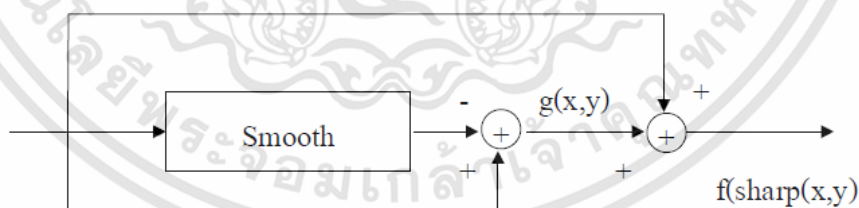
เป็นเทคนิคหนึ่งที่ยอดนิยมอีกวิธีหนึ่ง เทคนิคดังกล่าวนี้เป็นการเพิ่มความสว่างและความเข้มแสงของภาพ โดยไม่มีผลกระทบต่อสีของภาพเทคนิคดังกล่าวนี้มนำมาใช้กับภาพที่ถ่ายแล้วมืด

- การกรองภาพ (Filtering)

การกรองภาพมีอยู่หลายวิธีเช่น Average, Gaussian, Laplacian, LOG, Prewitt Sobel, Unsharp ในนี้จะกล่าวเฉพาะวิธีการกรองภาพ

### 2.3.5 วิธีการกรองแบบ Unsharp Filter

วิธีการกรองแบบนี้การเตรียมภาพก่อนการประมวลผลภาพอีกวิธีหนึ่งโดยการเพิ่มความคมชัดให้กับภาพ ความเร็วในการทำงานในการประมวลผลภาพรวดเร็วเนื่องจากการทำงานที่ไม่ซับซ้อน ลักษณะการทำงานแสดงได้ดังรูปที่ 2.3.5



รูปที่ 2.3.5 วิธีการปรับปรุงคุณภาพโดยวิธีการ Unsharp

จากรูปที่ 2.3.5 เมื่อข้อมูลภาพที่ต้องการปรับโดยใช้วิธีการนี้ต้องผ่านฟังก์ชันการปรับภาพให้ เรียบจากนั้นนำข้อมูลภาพที่ผ่านการปรับมาบวกกับข้อมูลภาพเดิมก็จะทำให้ได้ภาพที่คมชัดขึ้น ตามสมการที่ (2-6)

$$F_{\text{sharp}}(x,y) = f(x,y) + k * g(x,y) \quad (2-6)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อค่า  $k$  เป็นค่าในการปรับความคมชัดมีค่าระหว่าง 0.2-0.7 ในกรณีค่า  $k$  มีค่ามากจะถือ ได้ว่าการกรองจะมีความคมชัดขึ้น

### 2.3.6 วิธีการแปลงภาพให้เป็นภาพไบนารี

การแปลงภาพให้เป็นไบนารีเพื่อแปลงค่าภาพก่อนนำไปสู่กระบวนการ ประมวลผลภาพแบบดิจิทัล (Digital Image Processing) ที่ใช้ในงานวิจัยคือ การใช้ค่าเทรชโฮลด์ (Threshold) การที่ภาพมีทั้งสีและลำดับชั้นของสี (Shade) ทำให้เกิดลวดลาย ซึ่งระดับความเข้มที่น้อย ที่สุดที่ทำให้เกิดลวดลายได้คือความเข้มสองระดับหรือที่เรียกว่าไบนารี (Binary) ซึ่งส่วนใหญ่คือสี ขาวดำและระดับความเข้มของสีขาวดำที่สามารถเก็บรายละเอียดของภาพได้หมดคือระดับ 256 ระดับ ดังนั้นการที่เราสร้างอัลกอริทึมจะสามารถจำแนกลวดลายได้รวดเร็วเราควรแปลงภาพให้มี ระดับความเข้มสองระดับ เพื่อให้คอมพิวเตอร์ทำการประมวลผลภาพได้ และมีวิธีการที่ใช้กันทั่วไป คือการแปลงภาพโดยใช้ค่าเทรชโฮลด์ (Threshold) ซึ่งมีสมการดังนี้ต่อไป

สมมุติให้  $B[i,j]$  คือภาพไบนารีที่เกิดจากภาพต้นแบบ  $f[i,j]$  ดังนี้

$$f_T[i,j] = \begin{cases} 1 & \text{if } f[i,j] > T \\ 0 & \text{otherwise} \end{cases} \quad (2-7)$$

โดยที่  $F_T[i,j]$  = ค่าระดับความสว่างที่ตำแหน่ง  $(i,j)$   $T$  = ค่าเทรชโฮลด์

จากสมการข้างต้นภาพวัตถุ (Object) จะเป็นระดับความเข้มที่สว่างส่วนที่มีดีกว่าจะเป็นพื้น (Background) แต่ถ้าวัตถุมีความเข้มอยู่ในระดับกลางๆเราจะหาภาพไบนารี ได้จาก

$$f_T[i,j] = \begin{cases} 1 & \text{if } T1 \leq f[i,j] < T2 \\ 0 & \text{otherwise} \end{cases} \quad (2-8)$$

โดยที่  $T1$  = ค่าเทรชโฮลด์ของความเข้มระดับที่ 1

$T2$  = ค่าเทรชโฮลด์ของความเข้มระดับที่ 2

จากสมการเราจะได้ภาพที่มีความเข้มในระดับกลางๆ เป็น 1 ส่วนนอกนั้นจะเป็น 0 การที่ในภาพเดียวกันอาจมีส่วนภาพในแต่ละส่วนไม่เท่ากันเราอาจแบ่งภาพออกเป็นส่วนๆแล้วใช้ค่าเทรชโฮลด์อัตโนมัติ (Automatic Threshold)

### 2.3.7 การหาขอบภาพ

การหาขอบภาพเป็นการหาขอบเขตของวัตถุภายในภาพซึ่งขอบเขตของวัตถุจะเป็นที่เด่น และมีความสำคัญมากที่สุดที่จะนำไปสู่การรู้จักวัตถุนั้นๆ โดยคอมพิวเตอร์จะเห็นได้ว่าการหาขอบ ภาพเป็นการหาจุดเด่นของวัตถุตัวเอง ลักษณะที่เด่นของวัตถุที่เรามองเห็นโดยทั่วไปก็คือ ส่วนที่

เป็นเส้น หรือส่วนที่เป็นขอบของวัตถุ เมื่อมีแสงมากระทบก็จะมีแสงสว่างมากกว่าส่วนอื่น โดยทั่วไป ลักษณะของขอบภาพที่ดีจะต้องมีลักษณะดังต่อไปนี้

- มีความบางคือ จุดที่เป็นขอบภาพจะต้องมีความกว้างเพียงจุดเดียวเท่านั้น
- มีความต่อเนื่อง โดยจุดของขอบภาพในวัตถุเดียวกันมีความต่อเนื่องกัน ถ้าจุดที่เป็นขอบภาพมีอยู่เพียงจุดเดียว โดยไม่ต่อเนื่องกับจุดใดในย่านใกล้เคียงเลยอาจจะเป็น ขอบภาพที่ไม่สมบูรณ์ หรืออาจจะเป็นสัญญาณรบกวนได้

- ตำแหน่งของจุดที่เป็นขอบภาพควรมีความถูกต้องโดยขอบที่มีความถูกต้องนั้นจะต้องอยู่ในส่วนที่มีค่าความเข้มของจุดภาพที่สูงกว่า

จากรูปที่ 2.3.7 ได้แสดงขอบภาพ โดยใช้ตัวเลขในลักษณะต่างๆ โดยรูปที่ 2.3.7 (ก) แสดงขอบ ภาพที่มีความหนา และรูปที่ 2.3.7 (ข) เป็นขอบภาพที่ไม่ต่อเนื่อง

0	0	0	1	1	0	0	0	0
0	0	0	1	1	0	0	0	0
0	0	0	1	1	0	0	0	0
0	0	0	1	1	0	0	0	0
0	0	0	1	1	0	0	0	0
0	0	0	1	1	0	0	0	0
0	0	0	1	1	0	0	0	0

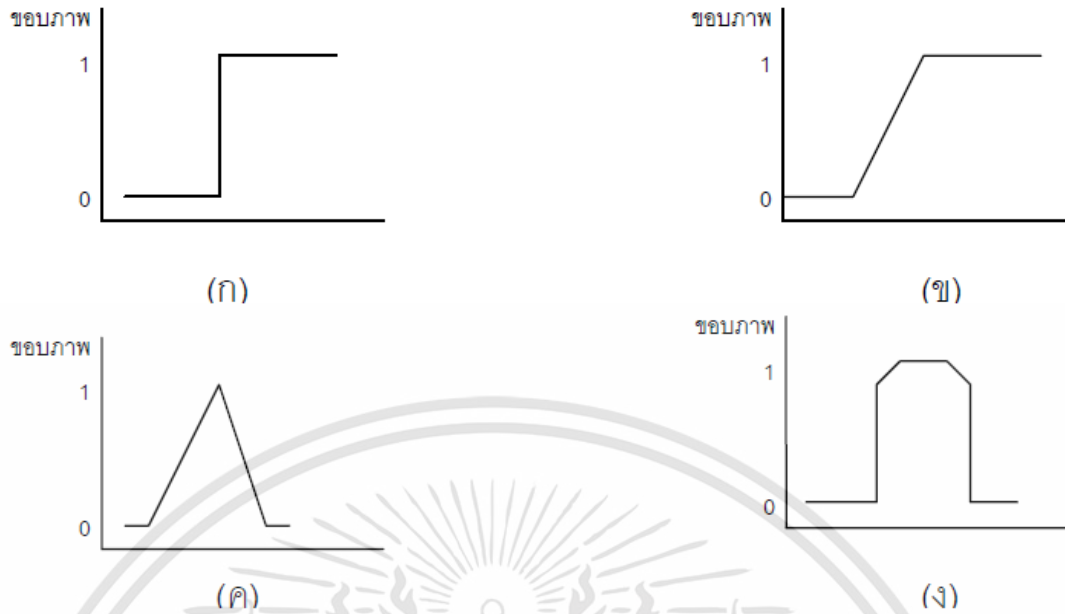
(ก) ขอบภาพหนา(กว้าง 2 จุด)

0	0	0	1	0	0	0	0	0
0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0
0	0	0	1	0	0	0	0	0

(ข) ขอบภาพที่ไม่ต่อเนื่อง

### รูปที่ 2.3.7 แสดงลักษณะขอบภาพแบบต่างๆ

รูปแบบของขอบภาพที่ปรากฏในรูปภาพโดยทั่วไปจะประกอบไปด้วยขอบภาพหลายชนิด ดังแสดงในรูป 2.3.7.1 (ก) เป็นขอบภาพที่เปลี่ยนแปลงทันทีทันใดขอบภาพแบบนี้ส่วนใหญ่เป็นภาพที่ มนุษย์จำลองขึ้นมาส่วนภาพโดยทั่วไปจะเป็นขอบภาพที่มีการเปลี่ยนแปลงความเข้มทีละน้อย ดัง รูป 2.3.7.1 (ข) รูป 2.3.7.1 (ค) ขอบภาพที่ค่าความเข้มที่ค่าความเข้มค่อยๆ เพิ่มขึ้น หรือลดลงทีละน้อย ในลักษณะของรูปหน้าจั่ว และรูป 2.3.7.1 (ง) แสดงขอบภาพที่เป็นเส้น



รูปที่ 2.3.7.1 แสดงขอบภาพชนิดต่างๆ

การเปลี่ยนแปลงค่าความเข้มแสงมีผลต่อการหาขอบภาพ หรือลักษณะทางกายภาพตัววัตถุเอง (เช่น ลักษณะของพื้นผิว และรูปทรง) คุณสมบัติเหล่านี้จะเป็นปัจจัยพื้นฐานที่สำคัญของรูปภาพที่จะบอกขอบเขตและลักษณะทางกายภาพของวัตถุที่อยู่ในภาพ

### 2.3.8 การหาขอบภาพโดยใช้อนุพันธ์อันดับที่หนึ่ง

เป็นการหาขอบภาพโดยแปลงเกรเดียนต์แบบไม่ต่อเนื่องบนข้อมูลภาพเชิงตัวเลข ซึ่งเป็น การหาตำแหน่งค่าความเข้มของจุดภาพที่มีการเปลี่ยนแปลง และเกรเดียนต์  $\nabla f(x,y)$  ที่เป็นฟังก์ชันต่อเนื่องในระนาบสองมิติซึ่งสามารถกำหนดได้โดย

$$\nabla f(x, y) = \left( \frac{\partial f}{\partial x} \right) + \left( \frac{\partial f}{\partial y} \right) \quad (2-9)$$

สำหรับขนาดของมันสามารถกำหนดได้ดังนี้

$$|\nabla f(x, y)| = \sqrt{\left( \frac{\partial f}{\partial x} \right)^2 + \left( \frac{\partial f}{\partial y} \right)^2} \quad (2-10)$$

ส่วนทิศทางของเวกเตอร์เกรเดียนต์คือ

$$\theta = \tan^{-1} \left( \frac{\partial f}{\partial x} / \frac{\partial f}{\partial y} \right) \quad (2-11)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เนื่องจากการหาขอบภาพเป็นการประมวลผลแบบไม่ต่อเนื่องดังนั้นจึงใช้อนุพันธ์ย่อยแบบไม่ต่อเนื่องทิศทางที่ตั้งฉากกับแกน  $x$  และ  $y$  ซึ่งสามารถกำหนดได้ดังนี้

$$\begin{aligned} \nabla_x g(x,y) &= g(x,y) - g(x-1,y) \\ \nabla_y g(x,y) &= g(x,y) - g(x,y-1) \end{aligned} \begin{bmatrix} x-1, y & x, y \\ x, y & x, y-1 \end{bmatrix} \quad (2-12)$$

และสำหรับทิศทาง  $\theta$  หาได้จากสมการที่ (2-11)

$$\nabla g(x, y) = g(x, y) \cos \theta + g(x, y) \sin \theta \quad (2-13)$$

ส่วนขนาดโดยประมาณของเกรเดียนต์  $g(x,y)$  สามารถกำหนดได้ดังนี้

$$|\nabla g(x, y)| = \sqrt{(\nabla_x g(x, y))^2 + (\nabla_y g(x, y))^2} \quad (2-14)$$

การหาขอบภาพโดยใช้เกรเดียนต์ในทางปฏิบัติจะมีลักษณะที่แตกต่างกันไป เช่น วิธี Roberts สามารถเขียนเป็นสมการได้ดังนี้

$$g(x, y) = \text{Max}(|g(x, y) - g(x + 1, y + 1)|, |g(x + 1, y) - g(x, y + 1)|) \quad (2-15)$$

โดยพิกัด  $x, y$  มีลักษณะดังนี้

$$\begin{bmatrix} x, y + 1 & x + 1, y + 1 \\ x, y & x + 1, y \end{bmatrix}$$

และฟังก์ชัน  $\text{Max}(A, B)$  จะเป็นค่าที่มากกว่าระหว่าง  $A$  กับ  $B$  วิธีของ Prewitt และ Sobel จะใช้จุดภาพในบริเวณใกล้เคียงที่จะอยู่ในกรอบหรือวินโดว์ขนาด  $3 \times 3$  ซึ่งอยู่ในรูปแบบดังนี้

$$\begin{bmatrix} x, y + 2 & x + 1, y + 2 & x + 2, y + 2 \\ x, y + 1 & x + 1, y + 1 & x + 2, y + 1 \\ x, y & x + 1, y & x + 2, y \end{bmatrix}$$

วิธีของ Prewitt

$$g(x, y) = |\sum_y^{y+2} g(x, y) - \sum_y^{y+2} g(x + 2, y)| + |\sum_x^{x+2} g(x, y) - \sum_x^{x+2} g(x, y + 2)| \quad (2-16)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิธีของ Sobel จะมีการใช้ค่าถ่วงน้ำหนัก ซึ่งสามารถกำหนดได้ดังนี้

$$g(x, y) = \left( |g(x, y) + 2g(x, y + 1) + g(x, y + 2) - (g(x + 2, y) + g(x + 2, y + 1) + g(x + 2, y + 2))| - |g(x, y) + 2g(x, y + 1) + g(x + 2, y) - (g(x, y + 2) + 2g(x + 1, y + 2) + g(x + 2, y + 2))| \right)$$

จากสมการข้างต้น (2-12)-(2-13) สามารถนำมาเขียนใหม่ในรูปของวินโดว์ได้ดังนี้

Roberts

$$w_1 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad w_2 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

Prewitt

$$w_1 = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} \quad w_2 = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

Sobel

$$w_1 = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \quad w_2 = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

จะเห็นได้ว่าในวินโดว์ของแต่ละวิธีจะมีค่าถ่วงน้ำหนักต่างไปจากการหาเกรเดียนต์ของสมการเริ่มต้น (2-11) ที่มีเพียง 2 ระดับ เช่น ค่าถ่วงน้ำหนักของ Roberts และ Prewitt จะได้ค่าถ่วงน้ำหนักของเกรเดียนต์ 3 ระดับ (คือ -1,0,1) และของ Sobel มี 5 ระดับ (คือ -2,-1,0,1,2) 3.10 การหาขอบภาพโดยใช้ออนุพันธ์อันดับที่สอง

การหาขอบภาพโดยวิธีนี้จะมีการพิจารณาทิศทางของขอบภาพ แต่จะสนใจเฉพาะขนาดของความเป็นขอบภาพ เกรเดียนต์) เท่านั้น ดังนั้นวินโดว์ที่ใช้หาขอบภาพโดยวิธีนี้จะมีเพียงวินโดว์เดียว ซึ่งวิธีต่างๆ ที่นำมาใช้ในงานวิจัยมี ดังนี้

- ตัวดำเนินการ Laplacian

ตัวดำเนินการหาขอบภาพ Laplacian ของฟังก์ชันสอง มิติ  $g(x,y)$  ซึ่งเป็นฟังก์ชันไม่ต่อเนื่องบนช่วง  $[x,y]$  ที่ได้มาจากการประมาณอนุพันธ์เชิงเส้นของฟังก์ชันต่อเนื่องดังสมการ(2-14)

$$\nabla^2 f(x, y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (2-17)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับในกรณีของเกรเดียนต์ของ  $\nabla^2 g(x, y)$  ที่ได้จากการประมาณสมการที่ (2-14) เพื่อใช้ในการหาขอบภาพในพื้นที่  $3 \times 3$  จุดสามารถกำหนดได้ดังนี้

$$\nabla^2 g(x, y) = 4g(x, y) - (g(x - 1, y) + g(x, y + 1) + g(x, y - 1)) \quad (2-18)$$

ในการกำหนด Laplacian เชิงตัวเลขสัมประสิทธิ์ที่คุณอยู่กับพิกัดที่ต้องการหาขอบภาพจะต้องมีค่าเป็นบวกเสมอส่วนสัมประสิทธิ์ในตำแหน่งอื่นๆอาจจะมีค่าเป็นบวกหรือลบก็ได้ เนื่องจากอนุพันธ์ของ Laplacian ผลรวมของสัมประสิทธิ์ทั้งหมดต้องเป็นศูนย์เสมอซึ่งในกรณีที่ถ้าจุดภาพที่อยู่ในพื้นที่  $3 \times 3$  มีค่าเท่ากันหมดเมื่อคอนโวลูชันกับตัวดำเนินการนี้ผลลัพธ์ที่ได้ก็คือศูนย์ซึ่งแสดงได้ว่าจุดภาพไม่ใช่ขอบภาพอย่างแน่นอน ในรูปที่ 2.3.8 แสดงวินโดว์ของ Laplacian ในรูปแบบต่างๆ ทั้งที่ใช้จุดภาพใกล้เคียง 4 จุดดังรูปที่ 2.3.8(ก) และ 8 จุดดังรูปที่ 2.3.8(ข) กับ(ค) ซึ่งจะมีผลต่อจุดภาพต่างกัน

$$\begin{array}{ccc} \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} & \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} & \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix} \\ \text{(ก)} & \text{(ข)} & \text{(ค)} \end{array}$$

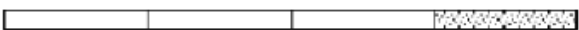


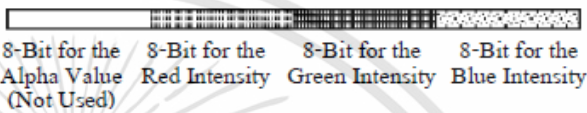
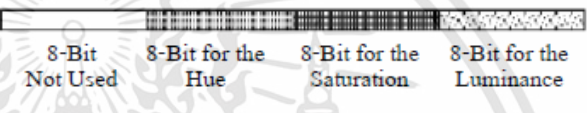

รูปที่ 2.3.8 แสดงตัวดำเนินการหาขอบภาพของ Laplacian

### 2.3.9 ประเภทของภาพ (Image Types)

ใน CorgiDude นั้นจะสามารถประมวลผลภาพได้ 3 ประเภท คือ ภาพระดับสีเทา ภาพสี และภาพสีเชิงซ้อน (Complex Image) อย่างไรก็ตาม แม้ว่า CorgiDude จะสามารถรองรับรูปแบบของ ภาพทั้ง 3 รูปแบบได้ แต่ก็จะมีกระบวนการบางกระบวนการที่ไม่สามารถใช้กับภาพบางประเภทได้ เช่น จะไม่สามารถใช้ Logic AND กับรูปภาพประเภทภาพสีเชิงซ้อนได้ทั้งนี้ขึ้นอยู่กับลักษณะพื้นฐานในการเข้ารหัสภาพแต่ละแบบ

จากตารางที่ 2.3.9 แสดงให้เห็นภาพประเภทต่างๆและจำนวนไบนารี (Byte) ที่ใช้ใน แต่ละประเภททั้งภาพที่เป็นแบบภาพระดับสีเทา ภาพสี และภาพสีเชิงซ้อน ดังนั้นสำหรับรูปที่มีความละเอียดเท่ากัน รูปสีจะใช้หน่วยความจำมากกว่ารูปที่เป็นแบบระดับสีเทา 8 บิต อยู่ถึง 4 เท่าและรูปแบบภาพสีเชิงซ้อนจะใช้หน่วยความจำมากกว่ารูปแบบระดับสีเทา 8 บิตอยู่ถึง 3 เท่า

ตารางที่ 2.3.9 แสดงจำนวนไบนารีที่ใช้ในแต่ละพิกเซลของภาพประเภทต่าง ๆ

ประเภทของภาพ	จำนวนไบนารีที่ใช้ในแต่ละพิกเซล
ภาพระดับสีเทา แบบ 8 บิต เลขจำนวนเต็ม	 8-Bit for the Grayscale Intensity
ภาพระดับสีเทา แบบ 16 บิต เลขจำนวนเต็ม	 16-Bit for the Grayscale Intensity
ภาพระดับสีเทา แบบ 32 บิต เลขทศนิยม	 32-Bit for the Grayscale Intensity
ภาพสีจริง RGB แบบ 32 บิต	 8-Bit for the Alpha Value (Not Used)   8-Bit for the Red Intensity   8-Bit for the Green Intensity   8-Bit for the Blue Intensity
ภาพสีจริง HSL แบบ 32 บิต	 8-Bit Not Used   8-Bit for the Hue   8-Bit for the Saturation   8-Bit for the Luminance
ภาพสีจริงเชิงซ้อน แบบ 64 บิต	 32-Bit Floating for the Real Part 32-Bit for the Imaginary Part

### 2.3.9.1 ภาพระดับสีเทา

ภาพที่เข้ารหัสแบบระดับสีเทานั้นจะประกอบด้วยพิกเซลที่อยู่ในระนาบ 1 ระนาบ โดยแต่ละพิกเซลอาจเข้ารหัสโดยใช้เลขจำนวนเต็มใดๆ ได้ตามวิธีการต่อไปนี้

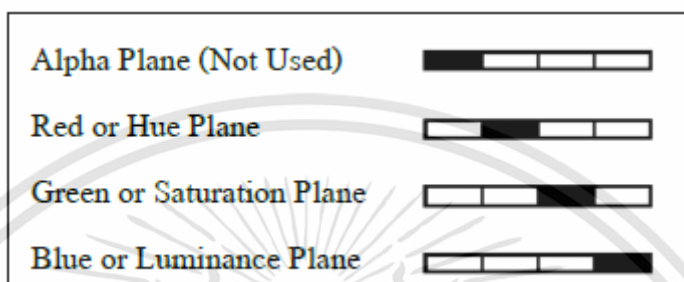
- ใช้ 8 บิต แบบไม่มีเครื่องหมาย (8-Bit Unsigned Integer) เป็นเลขจำนวนเต็ม ระหว่าง 0 ถึง 255 แทนค่าระดับความเข้มของสีดำ
- ใช้ 16 บิต แบบมีเครื่องหมาย (16-Bit Signed Integer) มีค่าระดับความเป็นสีดำระหว่าง - 32,768 ถึง 32,767
- ใช้เลขทศนิยมแบบความแม่นยำเดียว (Single-Precision Floating Point Number) เข้ารหัสโดยใช้ 4 ไบนารี หรือ 32 บิตเป็นค่าระดับความดำซึ่งจะแสดงค่าในช่วง 0 ถึง +50

### 2.3.9.2 ภาพสี

การเข้ารหัสเพื่อให้ได้ภาพสีนั้นจะมีการเข้ารหัสระดับความเข้มของสี 3 สี คือ สีแดง สีเขียว และสีน้ำเงิน ในระบบ RGB หรือในอีกวิธีการหนึ่งจะเป็นการแสดงด้วยค่า Hue Saturation และ Luminance หรือภาพแบบ HSL โดยในการเข้ารหัสของภาพสีทั้งสองแบบจะเป็นการเข้ารหัสที่ให้แต่ละพิกเซลมีค่าลำดับ 4 ค่า สำหรับภาพ RGB จะเก็บค่าโดยใช้เลข 8 บิต เพื่อเก็บค่าสี R-G-B ในแต่ละระนาบ ส่วน HSL ก็จะใช้เลข 8 บิตสำหรับค่า H-S-L ของแต่ละพิกเซล เช่นกัน ดังแสดงในรูปที่ 2.3.9.2 สำหรับภาพเข้ารหัสแบบ RGB-U64 จะเป็นการเก็บข้อมูลรหัสโดยใช้ เลข 16

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บิต สำหรับสี R-GB ในแต่ละระนาบ อย่างไรก็ตามภาพสีที่ใช้อยู่โดยทั่วไปจะมีการเตรียม ระนาบของสีไว้จำนวนทั้งสิ้น 4 ระนาบ ดังนั้นจะมีระนาบที่ไม่ได้ใช้งานในกรณีภาพสีบางแบบอยู่ หนึ่งระนาบ ดังแสดงในตารางที่ 2.3.9 และเนื่องจากการเตรียมระนาบไว้ 4 ระนาบ จึงมีการเรียกสีเหล่านี้ว่าเป็นแบบ 4x 8 บิต หรือการเข้ารหัสแบบ 32 บิต ส่วนภาพที่มีเฉพินิชั้น 16 บิต ก็จะเรียกว่า ภาพ 4 x 16 บิตหรือการเข้ารหัสแบบ 64 บิต



รูปที่ 2.3.9.2 แสดงระนาบที่บรรจุภาพสี

### 2.3.9.3 ภาพสีเชิงซ้อน

ในการเข้ารหัสของภาพสีเชิงซ้อนจะเป็นการบอกค่าความถี่ (Frequency) ของ ภาพระดับสีเทาซึ่งการสร้างภาพสีจะสามารถทำได้โดยการเปลี่ยนโดเมนที่ภาพนั้นอยู่ในสามารถสร้าง Complex Image ได้ โดยการใช้ Fast Fourier Transform (FFT) เข้ากับรูปที่ต้องการ และเมื่อเรา เปลี่ยนค่าของภาพระดับสีเทา ให้ไปอยู่ในรูปของภาพสีเชิงซ้อน แล้วเราจะสามารถใช้การคำนวณ หรือดำเนินการต่างๆ ในโดเมนความถี่เชิงซ้อนกับภาพที่เรากำลังวิเคราะห์ที่ได้สำหรับในแต่ละพิกเซลของภาพเชิงซ้อน ภาพจะถูกเข้ารหัสด้วยตัวเลข ทศนิยมความแม่นยำเดียว 2 จำนวน ซึ่งแต่ละจำนวนก็จะแทนค่าของส่วนจริงและอีกจำนวนหนึ่งจะ แทนส่วนของจำนวนจินตภาพตามลำดับ เราสามารถทราบค่าที่สำคัญ 4 องค์ประกอบออกจากภาพเชิงซ้อน นั่นคือส่วนจริง ส่วนจินตภาพ ขนาด (Magnitude) และ เฟส (Phase)

### 2.3.9.4 ประเภทของไฟล์ภาพ (Image File)

ประเภทของไฟล์ภาพที่นิยมใช้ในการเก็บภาพสำหรับคอมพิวเตอร์นั้นมีหลายประเภท ซึ่งจะมีลักษณะการจัดเก็บขนาดของไฟล์ที่แตกต่างกันแต่ในไฟล์รูปภาพจะมีส่วนประกอบสำคัญที่ เหมือนกันหลายจุด ในส่วนแรกไฟล์จะประกอบด้วยหัวไฟล์ (Header) หรือหัวเรื่อง ซึ่งบอกลักษณะเฉพาะของไฟล์นั้น จากนั้นจะตามด้วยค่าพิกเซลซึ่งรายละเอียดการกำหนดค่าต่างๆ จะขึ้นอยู่กับประเภทหรือรูปแบบของไฟล์ (File Format) แต่โดยทั่วไป หัวเรื่องของไฟล์จะบรรจุข้อมูลเกี่ยวกับไฟล์ เช่น จำนวนพิกเซลในแต่ละแถวอนและแถวตั้ง ค่าเฉพินิชั้นของพิกเซล และค่าอื่นๆ นอกเหนือจากนั้นไฟล์รูปอาจจะบรรจุข้อมูลที่เกี่ยวข้องกับการเปรียบเทียบ (Calibration) รูปแบบการจับคู่

(Pattern Matching Templates) และ Overlay ซึ่งนิยามของข้อมูลเหล่านี้จะกล่าวในภายหลัง ส่วนรูปแบบไฟล์ดิจิทัลที่นิยมใช้กันมีดังต่อไปนี้

- Bitmap (BMP)
- Tagged Image File Format (TIFF)
- Portable Network Graphic (PNG) โดยไฟล์นี้จะให้ข้อมูลที่สำคัญหลาย

แบบ เช่น การย่อ - ขยายรูปภาพ Calibration Pattern Matching และ Overlay

- Joint Photographic Experts Group Format (JPEG)
- National Instrument Internal Image File Format (AIPD) ใช้ในการ

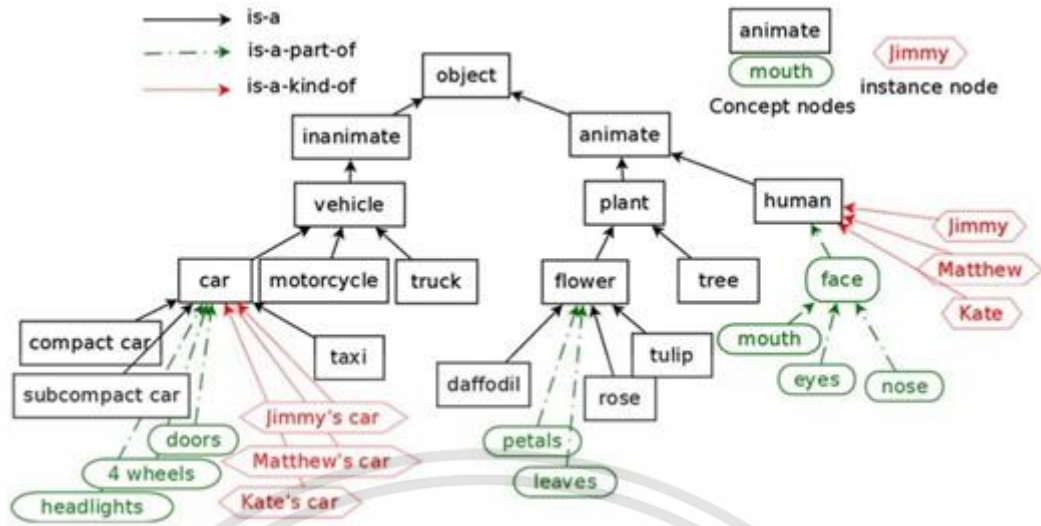
เก็บไฟล์ แบบเลขทศนิยม ภาพสีแบบเชิงซ้อน และภาพสีแบบ HSL

สำหรับโปรแกรมที่ใช้นี้จะเป็นโปรแกรมที่สามารถรับรูปแบบไฟล์ได้หลายแบบ ประกอบด้วย รูปแบบไฟล์มาตรฐาน แบบระดับสีเทา 8 บิต ส่วนไฟล์ภาพสีแบบ RGB จะมีรูปแบบ ได้หลายแบบคือ BMP TIFF PNG JPEG และ AIPD ส่วนมาตรฐานของระดับสีเทา 16 บิตระบบสี RGB แบบ 64 บิตและภาพเชิงซ้อน จะใช้ไฟล์ในรูปแบบของ PNG และ AIPD

### 2.3.9.5 ตัวอย่างแนวคิดเครือข่ายความสัมพันธ์ระหว่างคำอธิบายความหมายบน

ภาพ

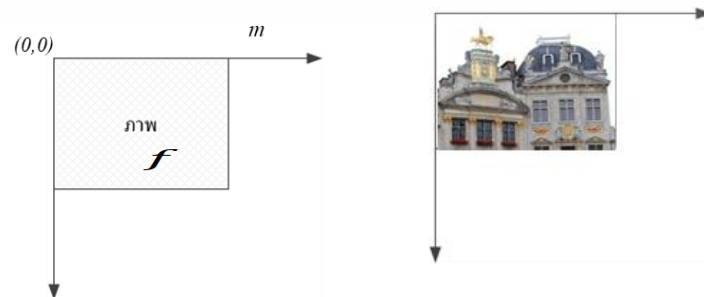
ออนโทโลยีเป็นแนวคิดแบบจำลองสำหรับรูปแบบคำอธิบาย ซึ่งจะมีการกำหนดประเภทวัตถุ ความสัมพันธ์ระหว่างวัตถุ รวมทั้งความสัมพันธ์ระหว่างประเภท ซึ่งจะมีการนำส่วนของเครือข่าย ความสัมพันธ์เข้ามาใช้บนออนโทโลยีดังแสดงในภาพที่ 2.3.9.5 ได้แสดงเครือข่ายความหมายที่ถูกแทน ด้วยประเภทของความสัมพันธ์ โดยมีโหนด 2 ประเภท ที่เรียกว่า โหนดคอนเซปต์ (concept) จะ แทนกลุ่มของวัตถุ เช่น car เป็นต้น และ โหนดอินสแตน (instance) แทนวัตถุ เช่น Jimmy's car เป็นต้น จะเห็นว่า การใช้โครงสร้างความหมาย (semantic structure) จะแทนความสัมพันธ์แบบ ระดับขึ้นระหว่างคอนเซปต์ แต่อย่างไรก็ตามการวิเคราะห์คำบรรยายเพื่อให้มาถึงความหมายของภาพ นั้น อาจจะไม่สามารถควบคุมการรู้จำวัตถุในบางกรณีได้เนื่องจาก ขนาดของวัตถุ เล็กเกินไป และ คำศัพท์ที่แทนนั้นยากในการควบคุม ทุกปัญหาที่เกิดขึ้นจะมีความสัมพันธ์ที่เกี่ยวข้องกันกับ อินสแตน และกลุ่มของวัตถุสำหรับการใส่คำอธิบาย



รูปที่ 2.3.9.5 ตัวอย่างแนวคิดเครือข่ายความสัมพันธ์ระหว่างคำอธิบายความหมายบนภาพ

### 2.3.9.6 การประมวลผลภาพดิจิทัล

ภาพจะถูกจัดเก็บเป็นสัญญาณดิจิทัล (digital) และทำการจัดเก็บข้อมูลภาพในหน่วยความจำ ในรูปของตัวแปรแบบอะเรย์ (array) โดยค่า ในแต่ละช่องของอะเรย์แสดงถึงค่าความเข้มของแสง (intensity) ของแต่ละจุดของภาพที่เรียกว่า พิกเซล (pixel) และตำแหน่งของช่องอะเรย์เป็นตัวกำหนดตำแหน่งของจุดภาพ เมื่อนำภาพมาแสดงในรูปของ 2 มิติ ตามแนวแกน x และ y สามารถเขียนในรูปของฟังก์ชันภาพ สามารถแสดงในรูปของ  $i = f(x,y)$  ในรูปแบบ 2 มิติโดย i หมายถึง ความสว่างหรือความเข้มของแสงและ (x,y) เป็นพิกัดในแบบจำลองของภาพ (image model) ค่าของฟังก์ชัน  $f(x,y)$  ดังกล่าวเป็นค่าความเข้มซึ่งเปลี่ยนไปตามตำแหน่งหรือพิกัดแบบจำลองภาพ ค่าความเข้มของแสงนี้ เป็นค่าดิจิทัลหรือเชิงตัวเลข ดังนั้น เมื่อแสดงเป็นตัวเลขของค่าความเข้มในแต่ละพิกเซลเรียงกันทั้ง ภาพจะเห็นว่าเป็นเมทริกซ์สี่เหลี่ยมที่แบบจำลองต่างจากแบบจำลองทางคณิตศาสตร์ คือ จุดกำเนิด (Coordinate origin) ที่ตำแหน่ง (0,0) ซึ่งอยู่ที่จุดซ้ายบนสุดของสเกลต่างจากจุดกำเนิดของแบบจำลองทางคณิตศาสตร์ที่อยู่จุดซ้ายล่างสุด ดังนั้นการนำเอาแบบจำลองทางคณิตศาสตร์มา ประยุกต์ในการประมวลผลภาพ จำเป็นต้องเปลี่ยนจุดกำเนิดให้ถูกต้องด้วย ดังแสดงในภาพที่ 2.3.9.6



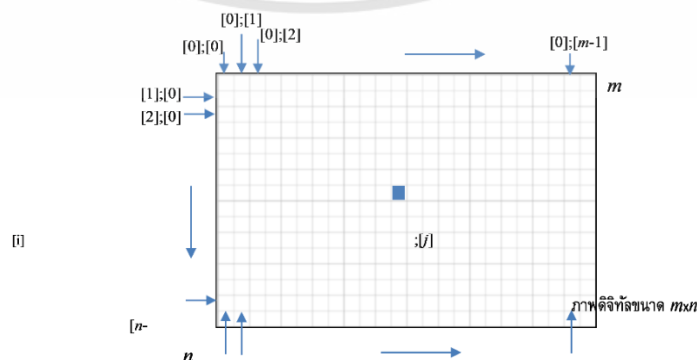
รูปที่ 2.3.9.6 แบบจำลองของภาพด้วยระนาบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในหน่วยความจำ จะทำการจัดเนื้อที่ในการเก็บภาพ สามารถคำนวณได้จาก  $m \times n \times b$  เมื่อ  $b$  เป็นจำนวนเต็มที่แทนจำนวนบิตของข้อมูลในแต่ละจุดภาพ ตัวอย่างถ้า  $b$  มีค่าเท่ากับ 8 บิต จะสามารถเก็บความแตกต่างของระดับสีที่เป็นไปสูงสุด 256 ระดับ ค่า  $m$  และ  $n$  จะเป็นตัวบอกถึงความละเอียดของภาพ สำหรับคอมพิวเตอร์ทั่วไปในระบบ VGA (Video Graphic Array) จะมีขนาดความละเอียดของภาพ สำหรับคอมพิวเตอร์ทั่วไปในระบบ VGA (Video Graphic Array) จะมีขนาด 640 x 480, 800 x 600 และ 1024 x 768 จุด เป็นต้น การกำหนดความละเอียดจะขึ้นอยู่กับงานที่จะใช้ในงานบางอย่างใช้ความละเอียดเพียง 30 x 50 จุด ก็พอซึ่งความละเอียดนั้นจะขึ้นอยู่กับงานที่จะใช้ ในบางงานจะใช้ความละเอียดถึง 1000 x 1000 จุด ก็ยังไม่พอ จากภาพที่ 2.1 สมมติให้ภาพแทนเป็นตัวแปรชื่อ : เป็นตัวแปรแบบอะเรย์ขนาด  $m \times n$  ( $m$  แทน แถว และ  $n$  แทน คอลัมน์) ที่ใช้เก็บภาพขนาด  $m \times 1$  จุด และ ค่าของความเข้มของแสง (ค่าความสว่าง) ของจุดภาพในแถวที่ 5 คอลัมน์ที่ 4 จะตรงกับค่าของข้อมูล : เป็นรูปของ (54) จะเห็นว่าใช้ตำแหน่งของจุดภาพทั้งสองแกนเป็นตัวชี้ค่าข้อมูลในอะเรย์ ดังแสดงการเรียงตัวของข้อมูลบนอะเรย์ และการกำหนด ความละเอียดของภาพ (image resolution) จากการกำหนดขนาดของพิกเซลตัวอย่างเช่น 1 ไมครอนต่อพิกเซล (um/pix) 1 มิลลิเมตรต่อพิกเซล (mm/pix) เป็นต้น ในงานที่ต้องการทราบ ตำแหน่งหรือขนาดของวัตถุที่วัดเป็นค่าจริง เราสามารถที่จะคำนวณได้จาก

$$\text{Resolution} = \frac{\text{Field of vision in Y direction (mm)}}{\text{Number of pixels in Y direction}}$$

โดยปกติแล้วในการเก็บข้อมูลภาพโดยเครื่องมือต่างๆ จะเก็บตามมาตรฐานของโทรทัศน์ซึ่งมี อัตราส่วน 1 ต่อ  $y$  เท่ากับ 4:3 สำหรับเครื่องมือเก็บข้อมูลภาพที่ไม่เป็นไปตามอัตราส่วน 4:3 เมื่อนำ ภาพนี้ไปแสดงในจอภาพมาตรฐาน จะทำให้ภาพที่แสดงนั้นมีขนาดของจุดภาพไม่เป็นที่เหลี่ยมจัตุรัส เช่น ในบางระบบอาจใช้ความละเอียดในการแสดง เท่ากับ 640 x 580 ซึ่งจะทำให้ขนาดของ จุดภาพที่ได้มีขนาดของด้านกว้างมีความยาวมากกว่าด้านสูง เมื่อมีการกำหนดให้ขนาดของบิตต่อจุด มากขึ้น จะทำให้จำนวนของสีมากขึ้นด้วย ตัวอย่างเช่น 1 บิต = 21 จะได้ 4 สี 2 บิต = 22 จะได้ 4 สี 4 บิต = 24 จะได้ 16 สี 8 บิต = 28 จะได้ 256 สี 16 บิต = 216 จะได้ 65536 สี เป็นต้น



รูปที่ 2.3.9.7 การจัดเก็บและเรียกใช้ข้อมูลภาพดิจิทัลลงในเมมทริกซ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากคุณลักษณะของภาพดิจิทัลที่กล่าวมาข้างต้นนั้น เป็นการเก็บข้อมูลภาพเป็นแบบเมทริกซ์ ที่มีถึง 3 ระนาบ (dimension) การประมวลผลภาพดิจิทัล (digital image processing) เป็น การเรียกใช้ขั้นตอนหรือกระบวนการที่มากกระทำบนภาพ โดยมีวัตถุประสงค์เพื่อปรับปรุงคุณภาพของ ภาพ ให้ได้ภาพใหม่ที่มีคุณสมบัติตามวัตถุประสงค์ที่นำไปใช้งาน เช่น การปรับให้ภาพมีความคมชัด มากขึ้น (enhancement) หรือการบีบอัดข้อมูลภาพ (compression) เพื่อประหยัดพื้นที่ในการจัดเก็บ ข้อมูล หรือการฝังลายน้ำ (watermark) เพื่อป้องกันการลักลอบการใช้ภาพที่ไม่ได้รับอนุญาตเป็นต้น สำหรับการประมวลผลภาพระดับสูงด้วยคอมพิวเตอร์ สามารถทำได้โดย นำภาพที่ได้มาจากกล้องหรือ image source ต่างๆ ซึ่งเป็นสัญญาณอนาล็อก แล้วนำมาแปลงเป็นสัญญาณดิจิทัลที่มีลักษณะเป็น รหัสเชิงตัวเลขฐานสอง (binary) ประกอบด้วยตัวเลข 0 และ 1 ที่สามารถใช้รูปแบบทางคณิตศาสตร์ เข้ามาช่วยในการคำนวณและการประมวลผลข้อมูลภาพ การดำเนินงานวิจัย สำหรับการจัดกลุ่ม ความหมายของภาพได้มีผู้วิจัยจำนวนมากที่ศึกษาและทำการทดลองโดยทั่วไป แบ่งเป็น 2 ระดับ ดังนี้ สำหรับขั้นตอนการประมวลผลภาพเบื้องต้น (image preprocessing) ในรูปแบบของการหาบริเวณที่ต้องการในรูปแบบอัตโนมัติ ยังคงเป็นงานวิจัยที่ยังหาข้อยุติไม่ได้ โดยเฉพาะที่เป็นลักษณะ ของ ระบบเรียลไทม์ (real time) ด้วยแล้วนั้นจะต้องคำนึงถึงเวลาในการคำนวณของ อัลกอริทึม (computational cost of algorithm) ที่เป็นสิ่งที่จำเป็นค่อนข้างมาก สำหรับระบบเรียลไทม์ ที่มีการ แปรเปลี่ยนรูปแบบของพื้นหลัง (modeling of background) จะทำให้เกิดกระบวนการในการสกัด วัตถุที่ต้องการขึ้นมา ซึ่งเทคนิคที่เรียกว่า Gaussians Mixture Model เป็นเทคนิคที่ค่อนข้างประสบความสำเร็จมาก แต่อย่างไรก็ตามยังคงต้องมีส่วนภาพพื้นหลังที่ไม่มีการเปลี่ยนแปลงมากนัก เพราะฉะนั้นปัญหาที่เกิดขึ้นยังคงเกิดขึ้นเมื่อพื้นหลังมีการแปรเปลี่ยนและยังคงเป็นปัญหาที่ยังคงต้องมีการแก้ไข ฉะนั้นในงานวิจัยนี้จึงข้ามในส่วนการวิเคราะห์การแปรเปลี่ยนของพื้นหลังที่ไม่คงที่ และได้มีการนำเทคนิคของฮิสโตแกรม (Histogram matching) ที่ไม่ได้มีการนำส่วนของพื้นหลังเข้ามา เกี่ยวข้อง เพื่อทำการเปรียบเทียบแต่อย่างไรก็ตาม ด้วยคุณลักษณะของเทคนิคฮิสโตแกรมไม่สามารถ ที่จะรองรับตำแหน่งของพิกเซล (pixel location) ทำให้ยังคงมีปัญหาเมื่อมีวัตถุที่ทับซ้อนกัน จนไม่สามารถที่จะประมวลผลได้อย่างถูกต้องเทคนิคสหสัมพันธ์ (Correlation) จะมีการทำงานที่ไม่สูญเสียในข้อมูลส่วนของสเปเชียลทำให้สามารถทดแทนวิธีการของฮิสโตแกรมได้ และได้มีการใช้เทคนิคสหสัมพันธ์เพื่อทำการหาขอบของวัตถุซึ่งวิธีการนี้ถูกเรียกว่า Edge-Enhanced Normalized Correlation (EENC) เพื่อทำการแก้ไขปัญหาต่างๆ ที่เกิดจากการหาวัตถุที่เกิดจากการทับซ้อนกัน หรือมีปัญหาจากสัญญาณรบกวนหรือวัตถุที่มีการเปลี่ยนแปลงทิศทางจากการหมุน เป็นต้น และการใช้เทคนิค EENC สามารถทำงานได้ดีสำหรับการแมชชีนซึ่งส่วนของพื้นที่ (matching region) พร้อมทั้งยังเป็นเทมเพลต (template) ที่ใช้ทำงานได้อย่างรวดเร็ว

### 2.3.9.7 การประมวลผลภาพระดับสูง

การประมวลผลภาพระดับสูง (high-level image processing) เป็นการใช้ข้อมูลพีเจอร์หรือ ผลลัพธ์จากการประมวลผลข้างต้นเพื่อผ่านกระบวนการ หรืออัลกอริทึมทำให้คอมพิวเตอร์รู้จักและเข้าใจภาพ (image understanding) รวมทั้งการหาความหมายของภาพ (Semantic image) ได้ เพราะฉะนั้นการประมวลผลภาพระดับสูงจำเป็นต้องใช้ข้อมูลที่ได้มาจากการประมวลผลภาพระดับต่ำ ดังนั้นจะเห็นได้ว่าการประมวลผลภาพระดับต่ำมีความสำคัญมากสำหรับการทำให้คอมพิวเตอร์รู้จักและเข้าใจ

### 2.3.9.8 การจำแนกประเภทข้อมูลภาพ

การจำแนกประเภทข้อมูลภาพ (image classification) เป็นการนำเอาข้อมูลทั้งหมดที่เก็บ รวบรวมมาได้ มาผ่านกระบวนการขั้นตอนการแยกแยะข้อมูลลงในแต่ละกลุ่มที่จัดไว้ โดยในแต่ละกลุ่มของข้อมูลนั้นจะมีคุณลักษณะเด่นของแต่ละกลุ่มที่แตกต่างกัน ขึ้นกับข้อมูล หรือพีเจอร์ ที่เก็บรวบรวมมาได้ รวมทั้งกระบวนการ หรือวิธีการที่ใช้แยกแยะข้อมูล ซึ่งในงานวิจัยนี้ได้ใช้เครือข่าย แบบเบย์ เพื่อเปรียบเทียบข้อมูลวัตถุดิบภาพเบื้องต้น และใช้วิธีการแยกประเภทข้อมูลภาพแบบ ลำดับชั้น ด้วยซอฟต์แวร์เวกเตอร์แมชชีน มีทฤษฎีดังนี้

## 2.4 การเรียนรู้ของเครื่อง

การเรียนรู้ของเครื่อง (Machine learning) เป็นวิธีการที่คอมพิวเตอร์สามารถเรียนรู้ได้ด้วยตนเอง เพื่อที่สามารถทำนาย หาผลลัพธ์ของปัญหาได้เอง จากความสัมพันธ์ของข้อมูลที่ป้อนเข้ากับผลลัพธ์ ตัวอย่างที่นำการเรียนรู้ของเครื่องมาประยุกต์ใช้ ได้แก่ การวิเคราะห์ตลาดหุ้น, ระบบอัจฉริยะสำหรับขับเคลื่อนโดยไร้คนขับ, การวินิจฉัยโรคในทางการแพทย์ เป็นต้น ซึ่งการเรียนรู้ของเครื่องสามารถแบ่งได้ 2 ประเภทหลัก ดังนี้

### 2.4.1 การเรียนรู้แบบมีผู้สอน

การเรียนรู้แบบมีผู้สอน (Supervised Learning) เป็นการเรียนรู้ของเครื่องโดยมีมนุษย์มาช่วย สอนคอมพิวเตอร์ก่อน ด้วยวิธีการระบุความสัมพันธ์ระหว่างข้อมูลที่ป้อนเข้ากับผลลัพธ์ สำหรับการ เรียนรู้ เพื่อที่สร้างโมเดลขึ้นมาสำหรับการทำนายผลลัพธ์ ยิ่งถ้าหากมีข้อมูลที่มีมาก ก็ สามารถสร้าง โมเดลที่มีความแม่นยำขึ้นได้มากขึ้นการเรียนรู้แบบมีผู้สอนของเครื่องสามารถแบ่งได้ 2 ประเภท ได้แก่ การแบ่งแยกประเภท (Classification) เป็นการทำการแยกประเภทสิ่งของต่างๆ จากการ สอนมนุษย์ด้วยการติดป้ายชื่อกำกับ (Label) เพื่อที่โมเดลสามารถเรียนรู้จากการทำป้ายชื่อกำกับของภาพที่กำหนดและดึงคุณลักษณะของภาพที่สำคัญสำหรับการเรียนรู้ในการสร้างโมเดลขึ้นมาซึ่งประเมินผลของการทำนายจากความน่าจะเป็นที่มากที่สุดที่ใกล้เคียงกับวัตถุที่ได้ติดป้ายชื่อกำกับไว้ การถดถอย (Regression) เป็นการหาความเกี่ยวข้องกันของข้อมูลที่มีความต่อเนื่อง โดยที่จะ

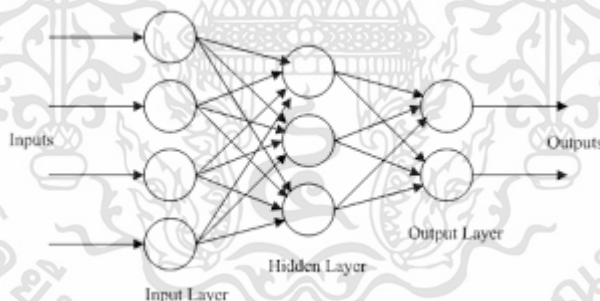
ดึง คุณลักษณะที่สำคัญจากข้อมูล เพื่อที่จะสามารถสร้างโมเดลและทำนายผลลัพธ์ออกมา เช่น การทำกรอบบอกตำแหน่งของการตรวจจับวัตถุ, การทำนายพยากรณ์อากาศ

#### 2.4.2 การเรียนรู้แบบไม่มีผู้สอน

การเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning) เป็นการเรียนรู้ของเครื่องโดยไม่มีมนุษย์ มาช่วยสอน ในการระบุนความสัมพันธ์ระหว่างข้อมูลที่ป้อนเข้ากับผลลัพธ์ ในการเรียนรู้ของเครื่อง จะ เรียนรู้จากข้อมูลที่ป้อนเข้า และหาความสัมพันธ์จากการดึงคุณลักษณะที่สำคัญของข้อมูลด้วยตนเอง สำหรับใช้งานในลักษณะการแบ่งกลุ่มการเรียนรู้ (Clustering) ที่ใช้ในการจัดโครงสร้างหรือแบ่งกลุ่ม ของข้อมูลที่ป้อนเข้า

### 2.5 การเรียนรู้เชิงลึก

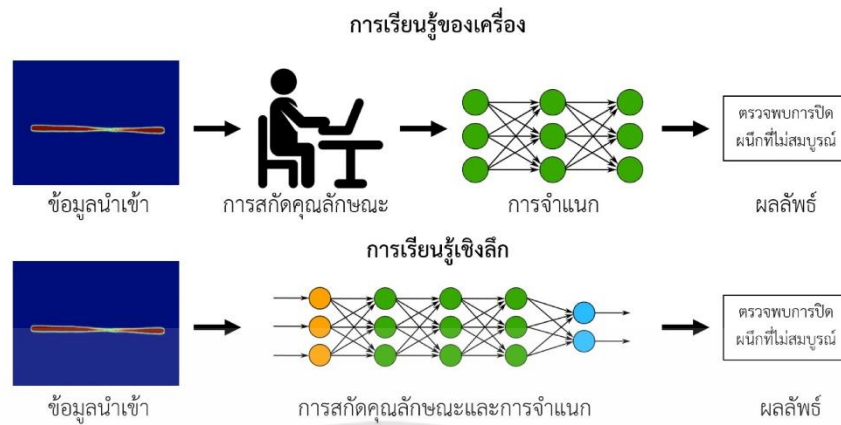
การเรียนรู้เชิงลึก (Deep learning) เป็นหนึ่งในโครงข่ายประสาทเทียม (Artificial Neuron Networks) ที่เป็นการเลียนแบบการทำงานของระบบโครงข่ายประสาทของสมอง ที่มีการเชื่อมต่อเส้นประสาทและเซลล์ประสาท โครงข่ายประสาทเทียมนั้นประกอบด้วย 3 เลเยอร์หลัก ได้แก่ เลเยอร์นำเข้า (Input layer), เลเยอร์ซ่อน (Hidden layer), เลเยอร์ผลลัพธ์ (Output layer) ดังรูปที่ 2.5ก ซึ่งเลเยอร์ซ่อนที่มากกว่า 2 ชั้นขึ้นไป สามารถเรียกว่าการเรียนรู้เชิงลึก และถือเป็นส่วนหนึ่งของการเรียนรู้ของเครื่อง



รูปที่ 2.5ก โครงสร้างของโครงข่ายประสาทเทียม

จุดประสงค์ของการเรียนรู้ของเครื่อง และการเรียนรู้เชิงลึก คือความสามารถในการทำนาย ผลลัพธ์ของข้อมูล ด้วยโมเดลที่สร้างจากการหาความสัมพันธ์ระหว่างข้อมูลป้อนเข้ากับผลลัพธ์ สำหรับในการสร้างโมเดลขึ้นมา ซึ่งความแตกต่างระหว่างการเรียนรู้ของเครื่อง และการเรียนรู้เชิงลึกอยู่ในส่วนของขั้นตอนการสกัดคุณลักษณะ (Feature Extraction) ของข้อมูลโดยการเรียนรู้ของเครื่องต้องสกัดคุณลักษณะด้วยมนุษย์เสียก่อน ถูกส่งต่อไปยังอัลกอริทึมสำหรับการจำแนกประเภท รวมทั้งผู้สกัดคุณลักษณะจำเป็นต้องมีความเข้าใจ สำหรับการเลือกข้อมูลที่จำเป็นที่ต้องใช้ในการสอน ของอัลกอริทึม แต่การเรียนรู้เชิงลึกสามารถสกัดคุณลักษณะได้ด้วยตนเองและจำแนกประเภทได้ โดยผู้สอนไม่การเรียนรู้ของเครื่อง

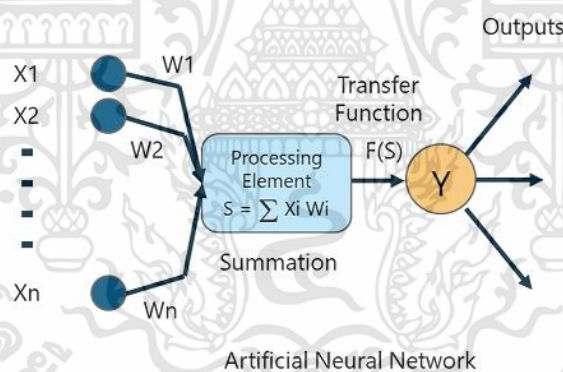
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.5ข ความแตกต่างระหว่างการเรียนรู้ของเครื่องและการเรียนรู้เชิงลึก

### 2.5.1 หลักการทำงานโครงข่ายประสาทเทียม

โครงข่ายประสาทเทียมเป็นการเชื่อมโยงถึงกันของเซลล์ประสาททั้งหมด สำหรับประมวลผลและส่งข้อมูลให้แก่กันโดยนิวรอน (Neuron) หรือเพอร์เซปตรอน (Perceptron) ซึ่งเป็นส่วนที่เล็กที่สุดของในโครงข่ายประสาทเทียม ดังรูปที่ 2.5.1 โดยทฤษฎีที่เกี่ยวข้องและหลักการทำงานโครงข่ายประสาทเทียมมีดังต่อไปนี้



รูปที่ 2.5.1 ภาพตัวอย่างการทำงานของโครงข่ายประสาทเทียม

การทำงานของนิวรอนเริ่มจากได้รับข้อมูลป้อนเข้า ซึ่งคำนวณผลรวมของค่าป้อนเข้าทั้งหมด ค่าน้ำหนัก และค่าเบี่ยงเบน จากนั้นผ่านการคำนวณของฟังก์ชันกระตุ้น ก่อนที่จะคำนวณผลลัพธ์ ออกมา กระบวนการนี้เรียกว่ารูปแบบป้อนไปข้างหน้า (Feedforward) ดังรูปที่ 2.5 ในการเรียนรู้ แบบมีผู้สอน เช่น การแบ่งแยกประเภทภาพในส่วนเลเยอร์ผลลัพธ์นั้นมีฟังก์ชันกระตุ้น สำหรับแปลง ผลลัพธ์เป็นความน่าจะเป็น โดยความแม่นยำของโมเดลสามารถตรวจสอบได้จากฟังก์ชันสูญเสีย จาก การคำนวณค่าความผิดพลาดของการทำนายกับผลลัพธ์จริง สำหรับใช้ในการปรับแก้ด้วยกระบวนการ แบ็คพรอพาเกชัน จากการคำนวณของเกรเดียนตสเซนส์ทำหน้าที่ในการปรับค่าน้ำหนัก และค่า เบี่ยงเบน เพื่อลดค่าความผิดพลาดของโมเดล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 2.5.1.1 ค่าน้ำหนัก

ค่าน้ำหนัก (Weight) มีหน้าที่ในการเชื่อมแต่ละชั้นของนิวรอนเข้าด้วยกัน เมื่อมีตัวแปร ป้อนเข้ามาจะคูณกับค่าน้ำหนักนั้นๆ ซึ่งค่าน้ำหนักที่มากจะบ่งบอกถึงความสำคัญของค่าตัวแปรที่ป้อนเข้าที่มีผลต่อค่าของผลลัพธ์

### 2.5.1.2 ค่าเบี่ยงเบน

ค่าเบี่ยงเบน (Bias) เป็นค่าตัวแปรที่ช่วยช่วยเพิ่มค่าของผลลัพธ์ หลังจากผลรวมทั้งหมดของผลคูณค่าป้อนเข้ากับค่าน้ำหนักทำให้กราฟของฟังก์ชันกระตุ้นเลื่อนไปทางซ้ายหรือขวา เพื่อให้ได้ผลลัพธ์ที่ออกมามีความเหมาะสมที่สุด ซึ่งค่าเบี่ยงเบนมีการสอนและปรับเปลี่ยนค่าคล้ายกับค่าน้ำหนักด้วยอัลกอริทึมเกรเดียนเดสเซนส์

### 2.5.1.3 ฟังก์ชันกระตุ้น

ฟังก์ชันกระตุ้น (Activation function) เป็นฟังก์ชันที่ช่วยในการตัดสินใจที่อยู่ในแต่ละ ชั้นของเลเยอร์ซ่อน หลังจากการคำนวณของค่าป้อนเข้าทั้งหมด ค่าน้ำหนัก และค่าเบี่ยงเบน โดยทำหน้าที่ในการกำจัดการกระตุ้นที่ไม่เกี่ยวข้อง เพื่อช่วยลดการคำนวณของคอมพิวเตอร์ ถ้าหากไม่มีการ กำจัดการกระตุ้นในนิวรอนเน็ตเวิร์กจะมีจำนวนมากซึ่งส่งผลต่อการคำนวณ และถ้าอยู่ในชั้นเลเยอร์ ผลลัพธ์ ทำหน้าที่แปลงเป็นค่าความน่าจะเป็นสำหรับการจำแนกภาพโดยใช้ฟังก์ชัน Softmax และ ฟังก์ชันกระตุ้นยังมีความสำคัญสำหรับการเรียนรู้เชิงลึกเนื่องจากเป็นฟังก์ชันที่หาอนุพันธ์ได้ (Differentiable Function) เพราะการเรียนรู้เชิงลึกต้องมีการปรับค่าน้ำหนักและค่าเบี่ยงเบนให้โมเดลมีความแม่นยำสูงขึ้น ด้วยวิธีการแบ็คพรอพาเกชัน จากการคำนวณของเกรเดียนเดสเซนส์ที่ต้อง ใช้การคำนวณหาอนุพันธ์ของฟังก์ชัน ฟังก์ชันกระตุ้นที่นิยม ได้แก่ Sigmoid, Tanh, Rectified Linear Unit (ReLU) ซึ่งแต่ละฟังก์ชันมีการกำหนดขอบเขตการตัดสินใจที่แตกต่างกัน

### 2.5.1.4 ฟังก์ชันสูญเสีย

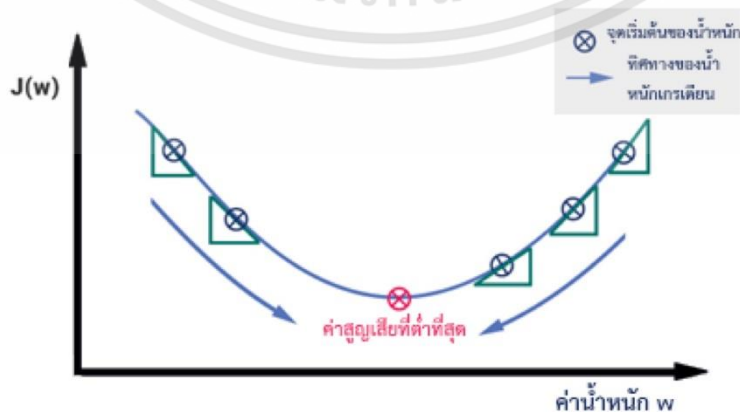
ฟังก์ชันสูญเสีย (Loss function) เป็นฟังก์ชันสำหรับเป็นค่าชี้วัดความแม่นยำของโมเดล จากการเปรียบเทียบความผิดพลาดระหว่างค่าการทำนายกับค่าที่เกิดขึ้นจริง โดยเป้าหมายของโมเดล ต้องการเกิดความผิดพลาดให้น้อยที่สุด ซึ่งฟังก์ชันสูญเสียมีหลากหลายประเภท ขึ้นอยู่กับการใช้งาน และสามารถนำฟังก์ชันสูญเสียที่หลากหลายมาผสมรวมกันเป็นฟังก์ชันสูญเสียใหม่ได้ เช่น งาน ประเภทตรวจจับวัตถุ นั้น เป็นการผสมผสานระหว่าง ฟังก์ชันสูญเสียของในงานการจำแนกประเภท เช่น ฟังก์ชันครอสเอนโทรปี (Cross Entropy Loss) ร่วมกับ ฟังก์ชันสูญเสียของในงานหาตำแหน่ง วัตถุจาก การถดถอย เช่น ฟังก์ชันแอลวันลอส (L1 Loss) และ ฟังก์ชันแอลทูลอส (L2Loss) เป็นต้น

### 2.5.1.5 แบ็คพรอพาเกชัน

แบ็คพรอพาเกชัน (Back propagation) เป็นกระบวนการสำหรับปรับแก้ค่าน้ำหนัก และค่าเบี่ยงเบน ให้โมเดลมีความแม่นยำมากขึ้น ด้วยวิธีการคำนวณจากค่าความผิดพลาดจากทำนาย ผลลัพธ์จากนิเวรอนเน็ตเวิร์คนำมาเปรียบเทียบกับค่าความจริง (Ground Truth) และนำค่าความผิดพลาดนี้ไปปรับค่าน้ำหนักและค่าเบี่ยงเบนของแต่ละนิเวรอน ที่มีอิทธิพลส่งผลกระทบต่อ การเปลี่ยนแปลง ของโมเดลที่แตกต่างกัน ด้วยหลักการคำนวณ สมการเชิงอนุพันธ์ย่อย (Partial Differential Equation) ของค่าความผิดพลาดเทียบกับค่าน้ำหนักและค่าเบี่ยงเบน ของแต่ละนิเวรอนของแต่ละตัว สำหรับในการหาพารามิเตอร์ของแต่ละค่าน้ำหนักและค่าเบี่ยงเบน เพื่อใช้อัตราการเปลี่ยนแปลง ของแต่ละค่าน้ำหนักและค่าเบี่ยงเบน สำหรับที่ในการลดค่าความผิดพลาดให้ต่ำที่สุด ด้วยวิธีการปรับ ให้กับค่าน้ำหนัก และค่าเบี่ยงเบนจากอัลกอริทึมเกรเดียนเดสเซนส์

### 2.5.1.6 เกรเดียนเดสเซนส์

เกรเดียนเดสเซนส์ (Gradient descent) เป็นกระบวนการที่อยู่ในการคำนวณของแบ็ค พรอพาเกชัน ในการปรับค่าน้ำหนักและค่าเบี่ยงเบน สำหรับที่จะลดค่าฟังก์ชันสูญเสียให้เกิดขึ้นน้อย ที่สุด เกรเดียนเดสเซนส์เปรียบเสมือนอยู่บนภูเขา เพื่อหาทางลงจากภูเขา ซึ่งบริเวณด้านล่างของภูเขานั้นเปรียบเสมือนฟังก์ชันสูญเสียที่เกิดขึ้นน้อยที่สุด การคำนวณเบื้องต้นของเกรเดียนเดสเซนส์ ประกอบด้วย 3 ส่วน จุดเริ่มต้น (ตำแหน่งหรือของค่าตอบปัจจุบัน), ความชันที่หาจาก (Partial Derivative) และอัตราการเรียนรู้ (Learning Rate) เมื่อเข้าสู่ใกล้ฟังก์ชันสูญเสียที่ต่ำได้ผลลัพธ์ต่ำที่สุด ขนาดการเปลี่ยนแปลงจะลดลงเรื่อยๆ เนื่องจากค่าความชันที่ลดลง และหยุดการคำนวณต่อเมื่อความ ชันมีค่าเท่ากับ 0 หรือมีค่าเข้าใกล้ 0 จนแทบไม่มีการเปลี่ยนแปลง (ดังรูปที่ 2.6) อัตราการเรียนรู้เป็น ค่าคงที่และส่งผลต่อการกำหนดขนาดก้าวเนื่องจาก ถ้าหากอัตราการเรียนรู้มีค่าที่มากอาจไม่พบจุดที่ ต่ำที่สุด เพราะจะก้าวเลยจุดต่ำที่สุด และถ้าหากอัตราการเรียนรู้มีค่านี้น้อย จะใช้เวลานานสำหรับคำนวณ ซึ่งอาจทำให้ไม่สามารถพบจุดต่ำที่สุดได้



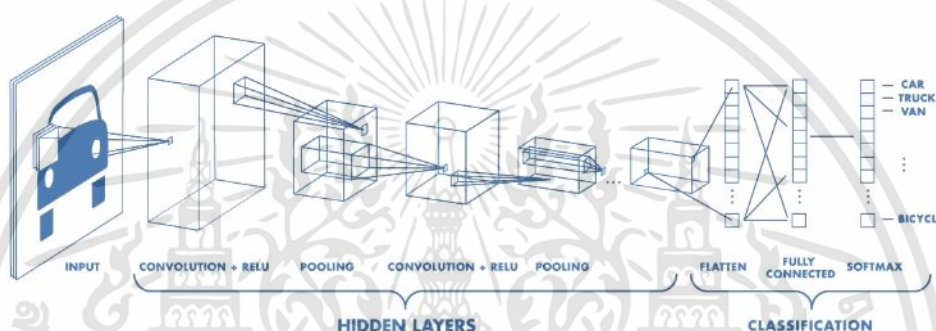
รูปที่ 2.5.1.6 ภาพตัวอย่างกระบวนการเกรเดียนเดสเซนส์ สำหรับการหาค่าสูญเสียที่ต่ำที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บางครั้งระหว่างการสอนอัลกอริทึมของเกรเดียนต์เดสเซนท์อาจพบปัญหาระหว่างการสอนได้แก่ แวนิชชิงเกรเดียนต์ (Vanishing Gradient) เป็นการปรับค่าน้ำหนัก และค่าเบี่ยงเบนที่น้อยมาก ที่ไม่เห็นความแตกต่างของการเปลี่ยนแปลง ส่งผลให้โมเดลไม่สามารถมีการสอนต่อได้ และเอ็กซ์พลอดิงเกรเดียนต์ (Exploding Gradient) เป็นการปรับค่าน้ำหนัก และค่าเบี่ยงเบนเป็นค่าอนันต์

## 2.6 โครงข่ายประสาทเทียมแบบคอนโวลูชัน

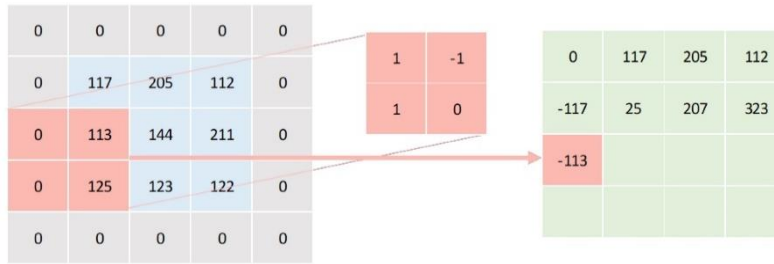
โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network) เป็นเครือข่ายประสาทเทียมชนิดเชิงลึกประเภทหนึ่งที่ใช้สำหรับในการจำแนกภาพ



รูปที่ 2.6 โครงข่ายประสาทเทียมแบบคอนโวลูชัน

### 2.6.1 ชั้นคอนโวลูชันนอล

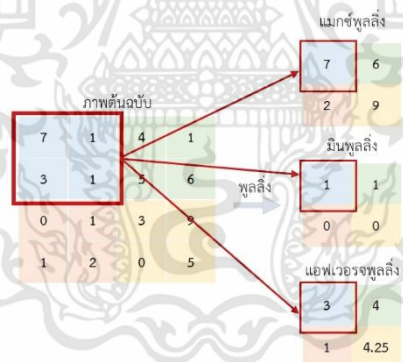
ชั้นคอนโวลูชันนอล (Convolutional Layer) เป็นชั้นสำหรับกระบวนการสร้างพีเจอร์แมพ (Feature map) จากภาพที่ป้อนเข้ามา ซึ่งกระบวนการสร้างพีเจอร์แมพ ประกอบด้วย ตัวกรอง (Filter) หรือ เคอร์เนล (Kernel) ที่ช่วยในการสกัดคุณลักษณะที่เด่นของภาพ เช่น รูปทรง, ขอบภาพ หรือสี เป็นต้น ซึ่งตัวกรองสามารถมีมากกว่าหนึ่งตัวกรองสำหรับการสร้างพีเจอร์แมพ การเลื่อนของตัวกรองเรียกว่าการก้าวข้าม (Stride) ทำหน้าที่ในการเลื่อนตัวกรองในภาพสำหรับสร้างพีเจอร์แมพ ถ้าหากการก้าวข้ามที่มากขึ้น การสกัดคุณลักษณะเด่นของภาพจะสกัดได้น้อยลงรวมและพีเจอร์แมพมีขนาดที่เล็กลง ในบางครั้งภาพที่นำมาใช้มีรายละเอียดที่สำคัญอยู่บริเวณขอบภาพ จึงมีกระบวนการแพดดิ้ง (Padding) เพื่อป้องกันข้อมูลที่สำคัญจากบริเวณขอบภาพ โดยทำการเพิ่มข้อมูลที่เสริมเกินออกไปจะแทนค่าข้อมูลช่องๆนั้นด้วยค่า 0 ดังรูปที่ 2.6.1 แสดงตัวอย่างการสร้างพีเจอร์แมพ จากภาพสี่ขนาด  $3 \times 3 \times 1$  พิกเซลของซาแนลสีแดง และมีเพิ่มกระบวนการแพดดิ้งกลายเป็น ขนาด  $5 \times 5 \times 1$  พิกเซล ตัวกรองมีขนาด  $2 \times 2 \times 1$  พิกเซล และการก้าวข้ามครั้งละ 1 พิกเซล



รูปที่ 2.6.1 การทำงานของตัวกรองที่ใช้สำหรับในการสร้างพีเจอร์แมพ

### 2.6.2 ชั้นพูลลิ่ง

ชั้นพูลลิ่ง (Pooling Layer) จุดประสงค์สำหรับการย่อขนาดภาพให้มีขนาดข้อมูลภาพเล็กลง โดยยังคงลักษณะเด่นของภาพไว้สำหรับในการคำนวณขั้นถัดไป ซึ่งการทำงานมีตัวกรองในลักษณะ เดียวกับการสร้างพีเจอร์แมพของชั้นคอนโวลูชันนอล และสามารถลดปัญหาโอเวอร์ฟิตติ้ง (Overfitting) และเพิ่มประสิทธิภาพการคำนวณให้รวดเร็วยิ่งขึ้นการพูลลิ่งสามารถจำแนกได้ 3 ประเภทได้แก่ แมกซ์พูลลิ่ง (Max pooling) มินพูลลิ่ง (Min pooling) และ แอฟเวอร์จพูลลิ่ง (Average pooling) ตัวอย่างกรณีต้องการแมกซ์พูลลิ่ง มินพูลลิ่ง และแอฟเวอร์จพูลลิ่ง จากเมทริกซ์ของพีเจอร์ แมพขนาด 4x4 พิกเซล โดยกำหนดตัวกรองที่มีขนาด 2x2 พิกเซล และการก้าวข้ามครั้งละ 2 พิกเซล โดยตัวกรองเลือกค่าพิกเซลตามแต่ละประเภท ดังตัวอย่างรูปที่ 2.6.2



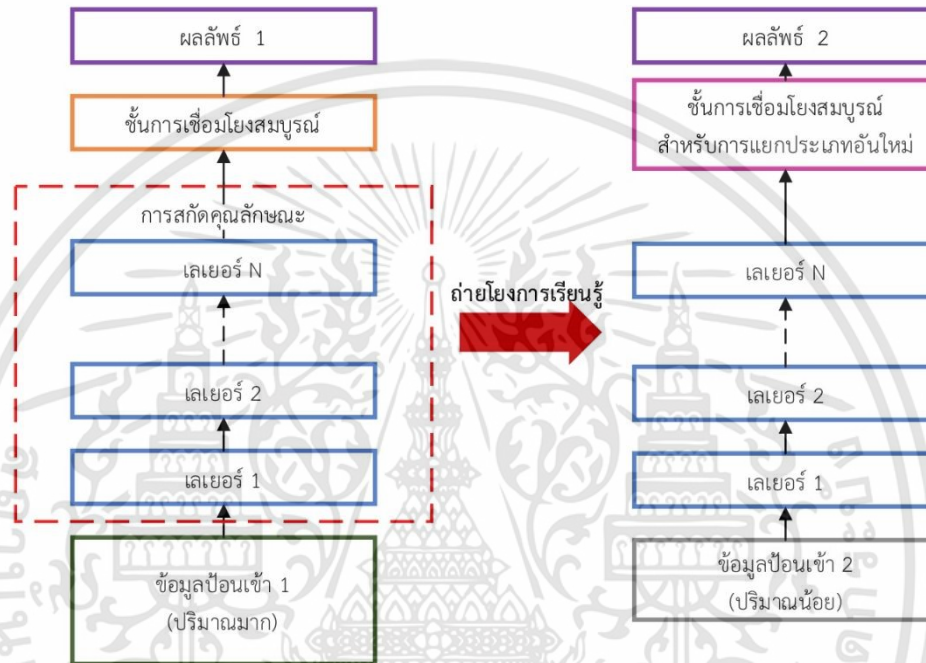
รูปที่ 2.6.2 ตัวอย่างการลดขนาดของภาพด้วยวิธีพูลลิ่ง

### 2.6.3 ชั้นการเชื่อมโยงสมบูรณ์

ชั้นการเชื่อมโยงสมบูรณ์ (Fully connected layer) เป็นส่วนสุดท้ายที่ทำหน้าที่ในการเชื่อมต่อของจุดของแต่ละชั้นก่อนหน้าที่ผ่านการคอนโวลูชันนอล พูลลิ่ง จากการดึงคุณลักษณะสำคัญของภาพ ที่ใช้สำหรับการสร้างโครงข่ายประสาทเทียมทำการเชื่อมต่อกันเป็น 1 มิติ ในการเรียนรู้และทำนายประเภทของรูปภาพ แสดงผลลัพธ์ในรูปแบบแสดงค่าความมั่นใจ

## 2.7 การถ่ายโยงการเรียนรู้

การถ่ายโยงการเรียนรู้ (Transfer Learning) เป็นการนำโมเดลที่ผ่านการฝึกฝนมาจากฐานข้อมูลขนาดใหญ่มาเรียบร้อยแล้ว หลังจากนั้นทำการแก้ไขบางส่วนในชั้นการเชื่อมโยงสมบูรณ์เพื่อที่สร้างโมเดลขึ้นมาใหม่ในลักษณะงานรูปแบบเฉพาะที่ต้องการ ซึ่งข้อดีสามารถช่วยประหยัดเวลาสำหรับในการสร้างโมเดลและใช้ปริมาณฐานข้อมูลปริมาณไม่มาก



รูปที่ 2.7 การใช้เทคนิคทรานเฟอร์เลิร์นนิง

## 2.8 การวัดประสิทธิภาพของการจำแนกประเภท

การวัดประสิทธิภาพของการจำแนกประเภท (Performance Evaluation of Classification) หลังจากสร้างโมเดลสำหรับจำแนกประเภทข้อมูล จำเป็นต้องมีการวัดประสิทธิภาพของโมเดล โดยสามารถทำได้ด้วยการชี้วัดจากตารางคอนฟิวชันเมทริกซ์ (Confusion Matrix) ดังตัวอย่างตารางคอนฟิวชันเมทริกซ์รูปที่ 2.8

		Actual	
		Positives	Negatives
Predicted	Positives	TP	FP
	Negatives	FN	TN

รูปที่ 2.8 ตัวอย่างตารางคอนฟิวชันเมทริกซ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยตัวอย่างที่อยู่ในตารางคอนฟิวชันเมทริกซ์มีความหมายดังต่อไปนี้

1) True Positive (TP) คือ สิ่งที่ทำนายว่ามันมีจริง และ เฉลยบอกว่ามีจริง ดังเช่น การทำนายว่าเป็นบริเวณปิดผนึกบรรจุภัณฑ์ที่สมบูรณ์ และผลลัพธ์คือเป็นบริเวณปิดผนึกบรรจุภัณฑ์ที่สมบูรณ์

2) True Negative (TN) คือ สิ่งที่ทำนายว่ามันไม่มีจริง และ เฉลยก็บอกว่ามันไม่มีจริง ดังเช่นการทำนายว่าเป็นบริเวณปิดผนึกบรรจุภัณฑ์ที่ไม่สมบูรณ์ และผลลัพธ์คือเป็นบริเวณปิดผนึกบรรจุภัณฑ์ที่ไม่สมบูรณ์

3) False Positive (FP) คือ สิ่งที่ทำนายว่ามันมีจริงแต่เฉลยก็บอกว่ามันไม่จริง ดังเช่น การทำนายว่าเป็นบริเวณปิดผนึกบรรจุภัณฑ์ที่สมบูรณ์ และผลลัพธ์คือเป็นบริเวณปิดผนึกบรรจุภัณฑ์ที่ไม่สมบูรณ์

4) False Negative (FN) คือ สิ่งที่ทำนายว่ามันไม่มีจริงแต่เฉลยก็บอกว่ามันมีจริงดังเช่น การทำนายว่าเป็นบริเวณปิดผนึกบรรจุภัณฑ์ที่ไม่สมบูรณ์ และผลลัพธ์คือเป็นบริเวณปิดผนึกบรรจุภัณฑ์ที่สมบูรณ์

จากคอนฟิวชันเมทริกซ์ที่กล่าวมาข้างต้นสามารถนำค่าในตารางไปคำนวณ เพื่อหาค่าวัดประสิทธิภาพที่เป็นตัวชี้วัดได้ความแม่นยำ, ความครบถ้วน, ค่าประสิทธิภาพโดยรวมและความถูกต้อง ดังต่อไปนี้

1) ความแม่นยำ (Precision) เป็นการวัดความแม่นยำโมเดลของสิ่งที่ทำนายที่สนใจ สามารถแยกกับสิ่งอื่นที่มารบกวนได้แม่นยำแค่ไหน โดยคิดเป็นอัตราส่วนระหว่าง การทำนายที่ ถูกต้องว่าจริง และที่เกิดขึ้นจริง กับ การทำนายว่าจริงแต่สิ่งที่เกิดขึ้นไม่จริงรวมกับการ ทำนายที่ถูกต้องว่าจริงและที่เกิดขึ้นจริง

$$precision = \frac{TP}{TP+FP} \quad (2.4)$$

2) ความครบถ้วน (Recall) เป็นการวัดความถูกต้องโมเดลของการสิ่งที่ทำนายที่สนใจกับผลลัพธ์สิ่งที่น่าสนใจที่เกิดขึ้นจริงทั้งหมด โดยคิดเป็นอัตราส่วนระหว่าง การทำนายที่ถูกต้อง ว่าจริงและที่เกิดขึ้นจริงกับการทำนายว่าไม่จริงแต่สิ่งที่เกิดขึ้นจริงกับรวมกับการทำนายที่ ถูกต้องว่าจริงและที่เกิดขึ้นจริง

$$Recall = \frac{TP}{TP+FN} \quad (2.5)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3) ค่าประสิทธิภาพโดยรวม (F - Measure) เป็นการวัดประสิทธิภาพโมเดลโดยรวมทั้งสองค่าระหว่าง ความแม่นยำ และความครบถ้วนพร้อมกันโดยใช้วิธีการหาค่าเฉลี่ยฮาร์โมนิก (Harmonic mean) ในการคำนวณ

$$F - Measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (2.6)$$

4) ความถูกต้อง (Accuracy) เป็นการวัดความถูกต้องของโมเดล ที่ทำนายจำนวนครั้งที่ทำนาย ถูกเทียบกับคำตอบทั้งหมด

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (2.7)$$



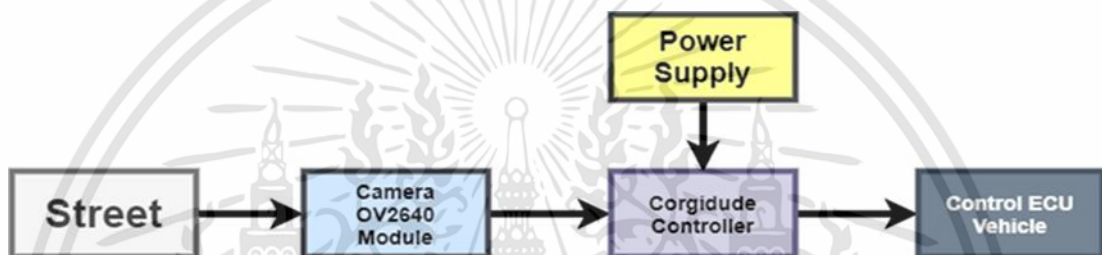
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### บทที่ 3

## วิธีการดำเนินการวิจัย

#### 3.1 การออกแบบระบบ

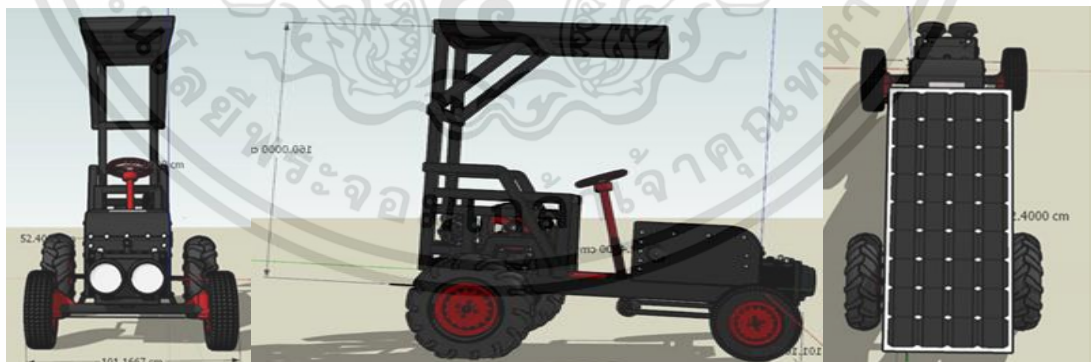
หลักการออกแบบระบบช่วยการขับเคลื่อนยนต์ไฟฟ้าโดยใช้การประมวลผลภาพ และแสดงผลบนโปรแกรม Colaboratory สามารถแบ่งการทำงานเป็น 2 ส่วนคือส่วนของกล้องที่ใช้ตรวจจับวัตถุที่ผ่านหน้ารถแทรกเตอร์ไฟฟ้า และส่วนของระบบควบคุมและขับเคลื่อนรถแทรกเตอร์ไฟฟ้าเพื่อทำให้รถแทรกเตอร์ไฟฟ้าสามารถเคลื่อนที่ได้ การออกแบบระบบควบคุมยานยนต์ไฟฟ้าแสดงดังรูป



รูปที่ 3.1 บล็อกไดอะแกรมการทำงานของระบบ

#### 3.2 ขั้นตอนการปฏิบัติงาน

1.การออกแบบโครงสร้างของรถโดยใช้โปรแกรม Sketch up Pro2020



รูปที่ 3.2.1 การออกแบบโครงสร้างของรถโดยใช้โปรแกรม Sketch up

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

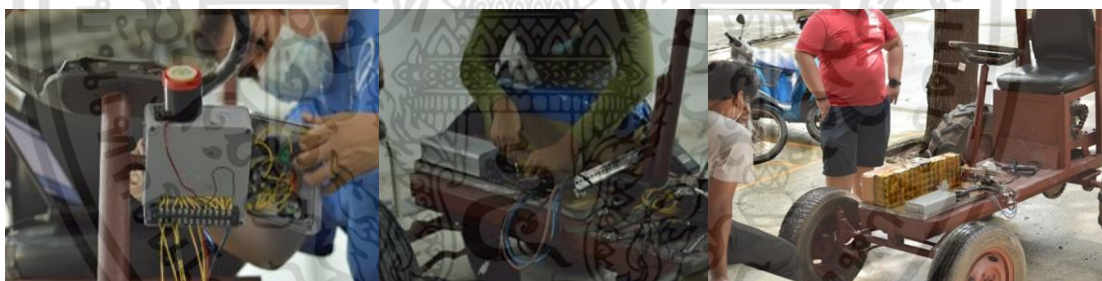
## 2. การดัดแปลงโครงสร้างเพื่อติดตั้งระบบกล่องควบคุม



รูปที่ 3.2.2 การดัดแปลงโครงสร้างของรถ



รูปที่ 3.2.3 ติดตั้งระบบช่วยในการขับเคลื่อนไฟฟ้า



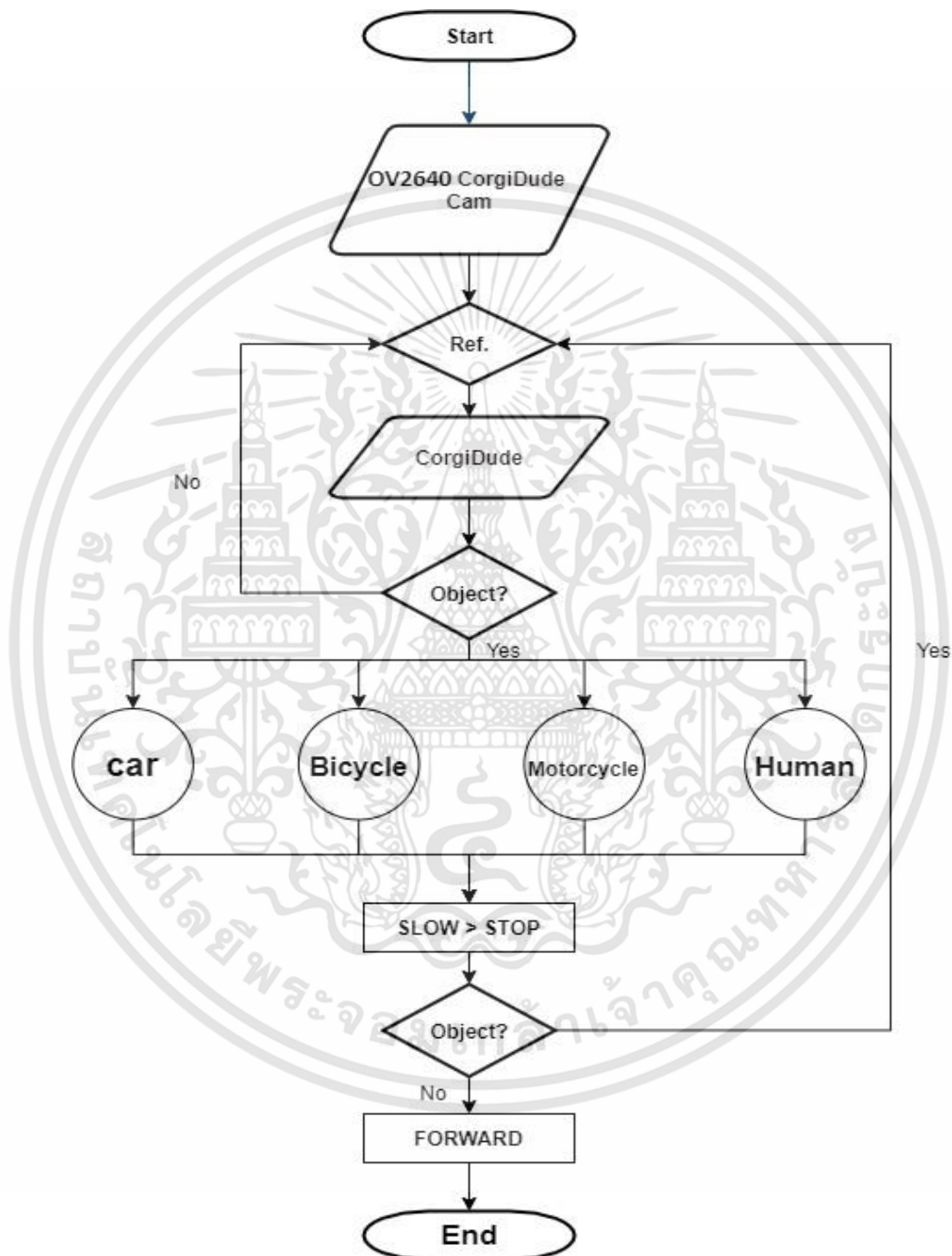
รูปที่ 3.2.4 ติดตั้งระบบ Inverter Controller และ ติดตั้งแบตเตอรี่



รูปที่ 3.2.5 ติดตั้งระบบช่วยในการขับเคลื่อนไฟฟ้าเสร็จสิ้นพร้อมทดสอบระบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยมีขั้นตอนการทำงานของระบบดังนี้

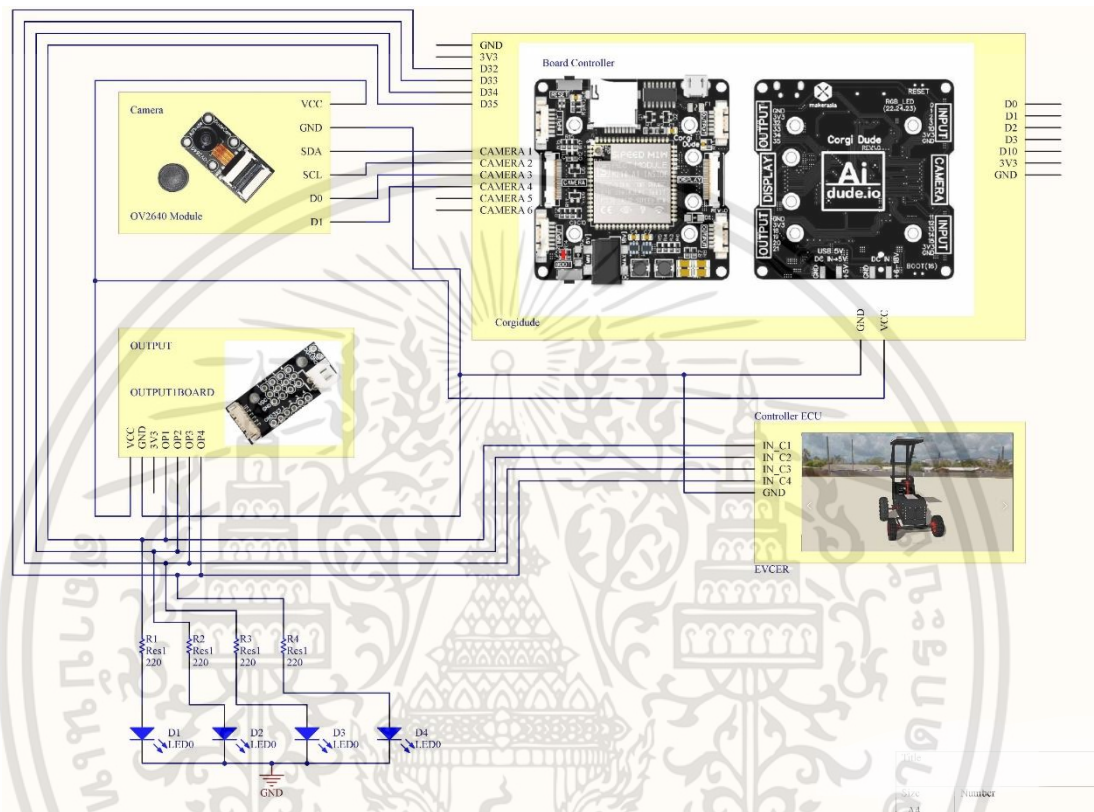


รูปที่ 3.2.6 Flow chart การทำงานของระบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 3.3 ขั้นตอนการทำงานของ CorgiDude

โดยในส่วนของ CorgiDude ออกแบบเพื่อทำการตรวจจับวัตถุบนถนนเพื่อสั่งการไปยังระบบควบคุมของรถแทรกเตอร์ไฟฟ้า โดยการติดตั้งกล่องควบคุม CorgiDude เพิ่มเติมไว้หน้ารถโดยภายในกล่องควบคุมมีการเชื่อมต่ออุปกรณ์ แสดงในรูปที่ 3.3



รูปที่ 3.3 การเชื่อมต่อวงจร

### 3.4 โครงสร้างภายในของบอร์ด CorgiDude ประกอบด้วย

1.ช่องต่อ Input พอร์ตสำหรับต่อเซนเซอร์ที่อ่านค่าได้ เช่นอุณหภูมิ ความชื้น ความเร็ว ความแรง แรงแดันไฟฟ้า โดยจะมีทั้งหมด 2 ช่อง (INPUT1 และ INPUT2) ช่องละ 5 GPIO นับรวมไฟเลี้ยง 3.3v รวมเป็น 7 Pin รวมกันแล้วสามารถใช้ GPIO ได้ทั้งหมด 10 Pin

2.ช่องต่อ Output มีลักษณะคล้ายกันกับ Input แต่ใช้สำหรับต่ออุปกรณ์ Output เช่น ควบคุมมอเตอร์ เซอร์โวมอเตอร์ ปิดเปิด Relay โดยช่อง Output จะมี 2 ช่อง คือ OUTPUT1 และ OUTPUT2 แต่ละช่องจะมี GPIO ทั้งหมด 4 ตัว และไฟเลี้ยง 2 ขา รวมเป็น 6 ขา รวมแล้วใช้ GPIO ได้ทั้งหมด 8 Pin ด้วยกัน

3.ช่องต่อกล้อง จะมีคำว่า “CAMERA” เขียนกำกับไว้ โดยกล่องที่รองรับจะเป็นรุ่น OV2640 24Pin

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.ช่องต่อจอภาพ จอภาพที่ต่อได้จะเป็นรุ่น ST7789 การต่อคล้ายกันกับการต่อกล้อง คือจะมีคอนเน็คเตอร์เขียนว่า “DISPLAY” กำกับเอาไว้

5.ช่องต่อไฟเลี้ยง ไฟเลี้ยงบอร์ดสามารถใช้ไฟจาก USB 5V หรือเลือกเสียบ Adaptor ที่จ่ายไฟเลี้ยงได้ตั้งแต่ 6-18V DC โดยใช้แจ๊คขนาดประมาณ 2x5.5mm

ปุ่มกดในบอร์ดจะมีทั้งหมด 2 ปุ่มคือ

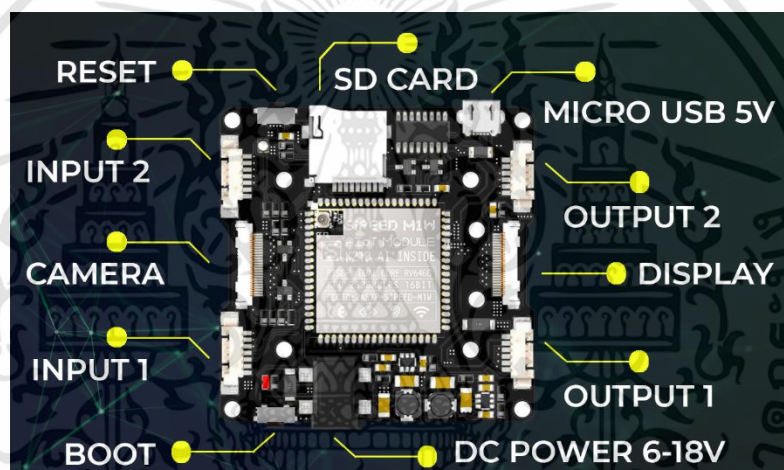
-RESET ใช้สำหรับกด reset บอร์ดให้เริ่มทำงานใหม่

-BOOT ใช้สำหรับกดเพื่อควบคุมการเข้า Boot โหมด หรือ สิ่งทำงานต่าง ๆ โดยต่อไว้กับ

GPIO 16

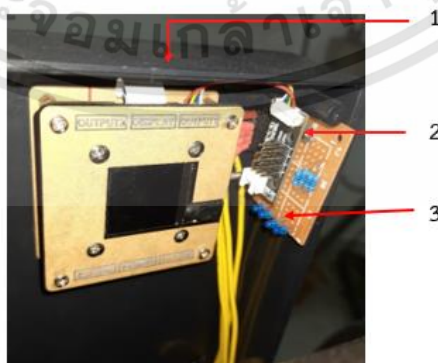
6.ช่องเสียบ Micro-SD ใช้สำหรับการโหลดไฟล์หรือบันทึกไฟล์ต่างๆ

7.หลอดไฟ RGB สำหรับแสดงสถานะต่างๆ



รูปที่ 3.4 รายละเอียดภายในของบอร์ด Corgitude

### 3.4.1 การวางตำแหน่งอุปกรณ์ภายในกล่องควบคุม



รูปที่ 3.4.1 การวางตำแหน่งของอุปกรณ์ภายในกล่อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

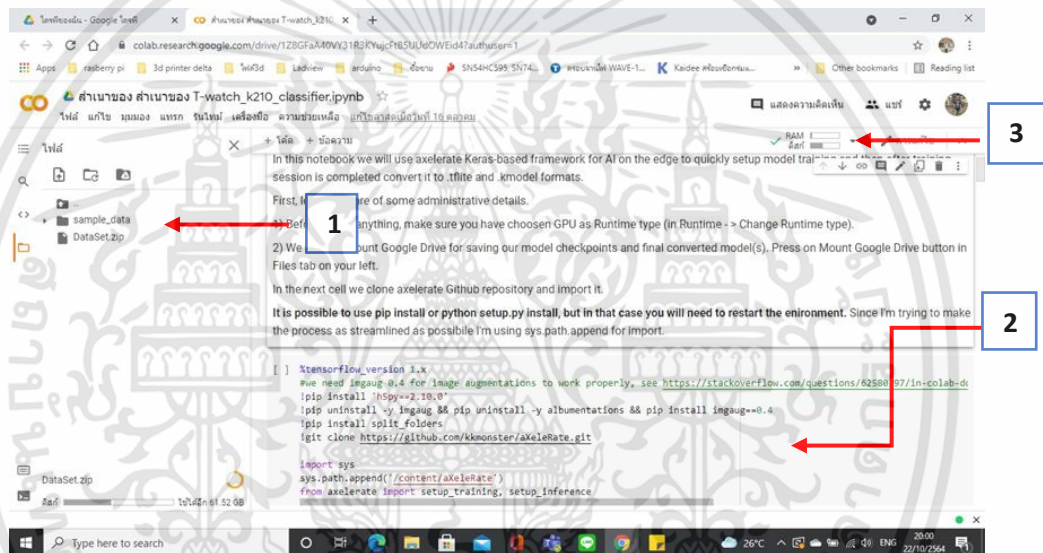
ในการออกแบบระบบกล้องควบคุมสำหรับใช้งานในรถแทรกเตอร์ไฟฟ้าเป็นการออกแบบอุปกรณ์ที่ติดตั้งเสริมไปในตัวรถแทรกเตอร์ไฟฟ้าทำให้ระบบที่ใช้ในการนำทางมีประสิทธิภาพเพิ่มมากขึ้นโดยมีการวางตำแหน่งอุปกรณ์ภายในกล้องควบคุมมีดังนี้

หมายเลข 1 คือ ตัวบอร์ด CorgiDude ประกอบไปตัวกล้องและหน้าจอแสดงผล

หมายเลข 2 คือ ตัวขยายพอร์ตสำหรับต่อใช้งานกับรถแทรกเตอร์ไฟฟ้า

หมายเลข 3 คือ ไฟแสดงสถานะ

### 3.5 โปรแกรม Colaboratory



รูปที่ 3.5 หน้าต่างของโปรแกรม Colaboratory

ในส่วนของการเขียนโปรแกรมเพื่อให้ตัวบอร์ด CorgiDude เพื่อให้ตัวบอร์ดได้จำค่าที่ได้กำหนดไว้ในเงื่อนไขของโปรแกรมหน้าต่างของโปรแกรมประกอบด้วยส่วนต่างๆดังรูป

หมายเลข 1 คือ เป็นการนำเข้าไฟล์จากคอมพิวเตอร์

หมายเลข 2 คือ พื้นที่สำหรับเขียนโปรแกรมและเงื่อนไขต่างๆ

หมายเลข 3 คือ สถานะของหน่วยความจำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 4

### ผลการทดลอง

การทดลองระบบช่วยการขับเคลื่อนยานยนต์ไฟฟ้าโดยใช้การประมวลผลภาพ และการแสดงผลแบบเวลาจริงบนโปรแกรม Colaboratory ประกอบไปด้วยการตรวจจับวัตถุแต่ละประเภทบนถนน แฉง้เตือนเมื่อมีวัตถุผ่านหรือขวางหน้ารถ

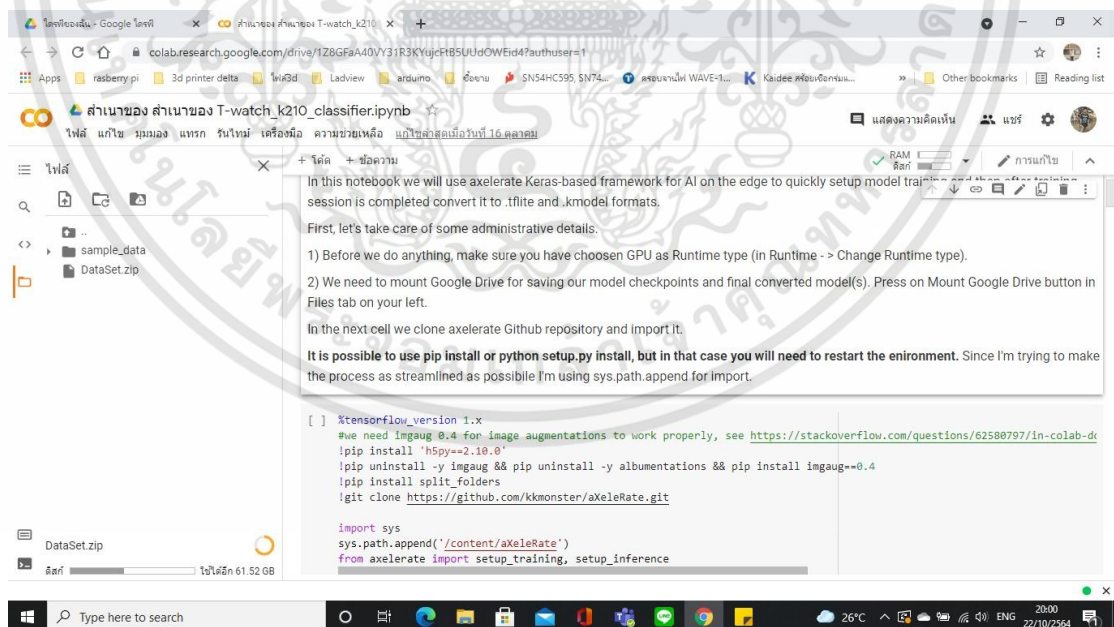
#### 4.1 ขั้นตอนการทดลอง

การทดลองมี 2 ส่วน

- 1) การเทรนนิ่งบอร์ด CorgiDude โดยใช้โปรแกรม Colaboratory
- 2) ติดตั้งอุปกรณ์กับตัวรถแทรกเตอร์ไฟฟ้าและทดสอบกับวัตถุจริง

##### 4.1.1 ขั้นตอนการเทรนนิ่งโดยใช้โปรแกรม Colaboratory

ขั้นตอนที่ 1 : ขั้นตอนการเทรนนิ่ง AI กด Run ตรงที่วงกลมเพื่อที่จะทำการติดตั้งไลบรารี tensorflow Version.1.x และ ติดตั้ง h5py Version 2.10.0 และ ติดตั้ง split\_folders และทำการ Clone-Module : <https://github.com/kkmonster/axeleRate.git>



รูปที่ 4.1.1 หน้าต่างเริ่มต้นของโปรแกรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

!tensorflow_version 1.x
#we need imgaug 0.4 for image augmentations to work properly, see https://stackoverflow.com/questions/62580797/in-colab-dt
!pip install 'h5py==2.10.0'
!pip uninstall -y imgaug && pip uninstall -y albumentations && pip install imgaug==0.4
!pip install split_folders
!git clone https://github.com/kkmonster/axelerate.git

import sys
sys.path.append('/content/axelerate')
from axelerate import setup_training, setup_inference

```

### รูปที่ 4.1.2 การติดตั้ง Library

ขั้นตอนที่ 2 : ทำการแตกไฟล์ รูปภาพที่ทำการอัปโหลด จากการที่ไปเก็บข้อมูลบนท้องถนน แยกไปด้วยกัน 4 Class :

- Class1 หรือ C1 = รถจักรยานยนต์หรือมอเตอร์ไซค์
- Class2 หรือ C2 = รถยนต์
- Class3 หรือ C3 = รถจักรยาน
- Class4 หรือ C4 = บุคคลบนทางถนน

```

!wget --id 1055P9M5t3vB1M-Da9kgh2WZ9ux7t11vF
Downloading...
From: https://drive.google.com/uc?id=1055P9M5t3vB1M-Da9kgh2WZ9ux7t11vF
To: /content/datasets.zip
12.9MB [00:01, 12.1MB/s]

Unzip datasets

!unzip -qq /content/Dataset.zip

split datasets for training and validating

import splitfolders

# The path to the directory where the original dataset was uncompressed
input_folder = '/content/Dataset'
# The directory where we will store our smaller dataset
output_folder = '/content/data_for_training'

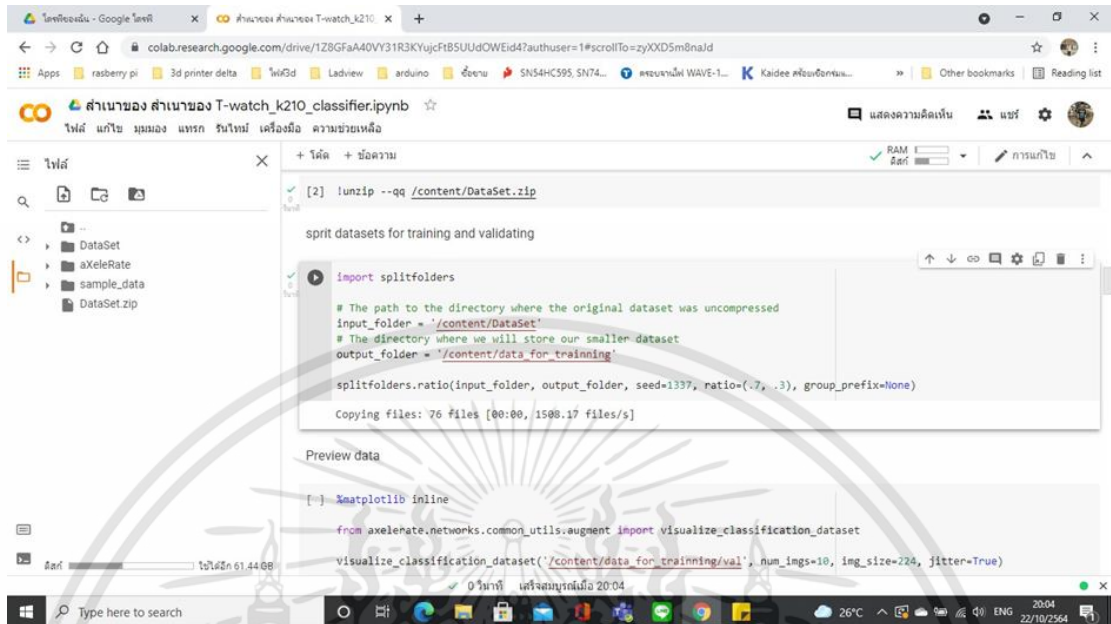
splitfolders.ratio(input_folder, output_folder, seed=1337, ratio=(.7, .3), group_prefix=None)

```

### รูปที่ 4.1.3 การแตกไฟล์รูปภาพเข้าโปรแกรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### ขั้นตอนที่ 3 : ทำการแบ่งไฟล์รูปภาพบนท้องถนนเพื่อไปทำการเทรนนิ่ง AI



```

[2] !unzip --qq /content/DataSet.zip

split datasets for training and validating

import splitfolders

# The path to the directory where the original dataset was uncompressed
input_folder = "/content/DataSet"
# The directory where we will store our smaller dataset
output_folder = "/content/data_for_training"

splitfolders.ratio(input_folder, output_folder, seed=1337, ratio=(.7, .3), group_prefix=None)

Copying files: 76 files [00:00, 1500.17 files/s]

Preview data

!matplotlib inline

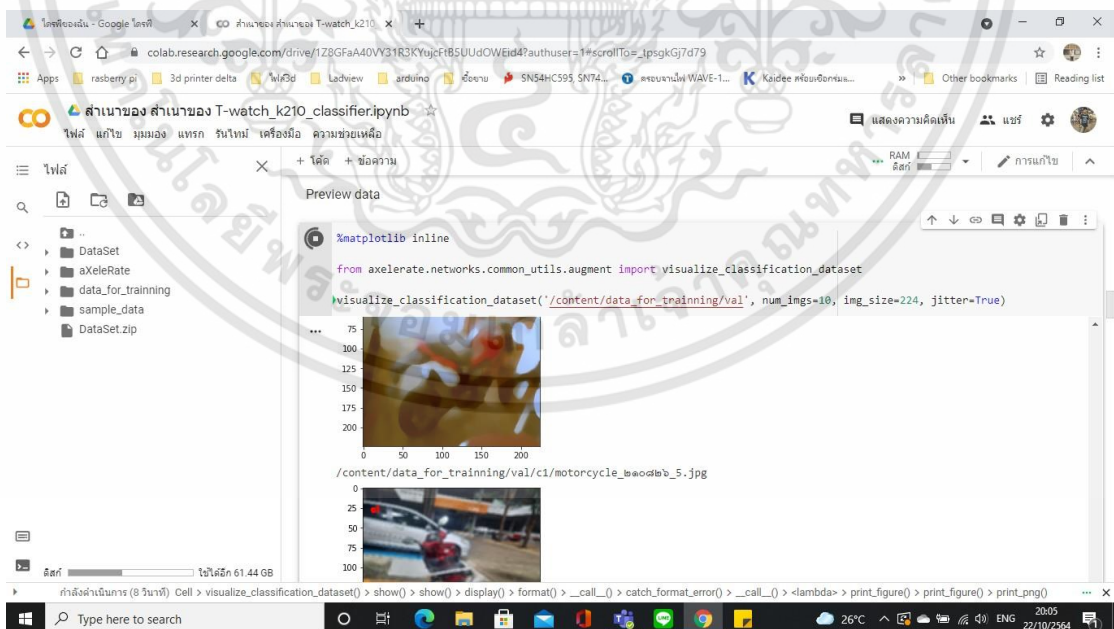
from axelerate.networks.common_utils.augment import visualize_classification_dataset

visualize_classification_dataset('/content/data_for_training/val', num_imgs=10, img_size=224, jitter=True)

```

รูปที่ 4.1.4 การแบ่งไฟล์รูปภาพบนท้องถนนเพื่อไปทำการเทรนนิ่ง AI

ขั้นตอนที่ 4 : ทำการเทรนรูปภาพ โดยการใช้คำสั่ง matplotlib เพื่อที่จะได้ไฟล์ .kmodel ออกมาเป็น OUTPUT เพื่อนำไปใช้ในางานของโครงการนี้ต่อไป



```

!matplotlib inline

from axelerate.networks.common_utils.augment import visualize_classification_dataset

visualize_classification_dataset('/content/data_for_training/val', num_imgs=10, img_size=224, jitter=True)

...

/c/content/data_for_training/val/c1/motorcycle_๒๑๐๘๒๖_5.jpg

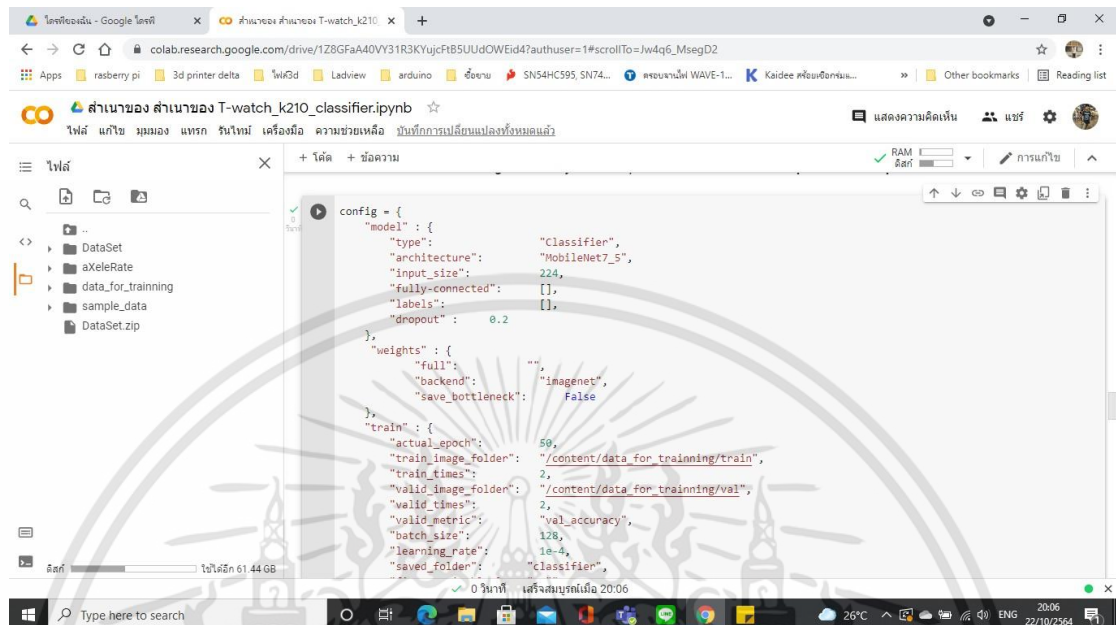
/c/content/data_for_training/val/c1/car_๒๑๐๘๒๖_5.jpg

```

รูปที่ 4.1.5 ทำการเทรนนิ่งรูปภาพเพื่อให้ได้ out put

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนที่ 5 : Resize ตัวข้อมูลในไฟล์ .kmodel เพื่อที่จะให้ความคมชัดใกล้เคียงกับ ตัวกล้อง OV2640 ที่นำมาเป็นโมดูลที่รับข้อมูลจากท้องถนนจากไปยัง บอร์ด Corgidue



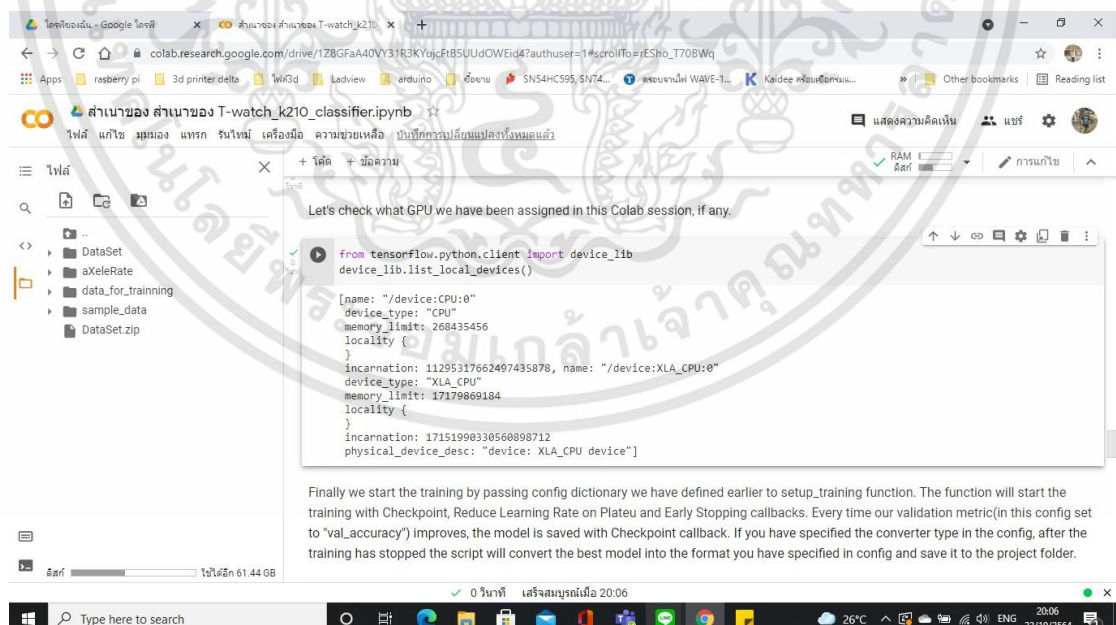
```

config = {
  "model": {
    "type": "Classifier",
    "architecture": "MobileNetV2",
    "input_size": 224,
    "fully-connected": [],
    "labels": [],
    "dropout": 0.2
  },
  "weights": {
    "full": "",
    "backend": "imagenet",
    "save_bottleneck": False
  },
  "train": {
    "actual_epoch": 50,
    "train_image_folder": "/content/data for training/train",
    "train_times": 2,
    "valid_image_folder": "/content/data for training/val",
    "valid_times": 2,
    "valid_metric": "val_accuracy",
    "batch_size": 128,
    "learning_rate": 1e-4,
    "saved_folder": "classifier"
  }
}

```

รูปที่ 4.1.6 การ Resize ตัวข้อมูลในไฟล์ .kmodel

ขั้นตอนที่ 6 : นำข้อมูลทั้งหมดที่ได้เทรนใน Colab นำออกมาเป็นข้อมูลไฟล์.Kmodel



```

from tensorflow.python.client import device_lib
device_lib.list_local_devices()

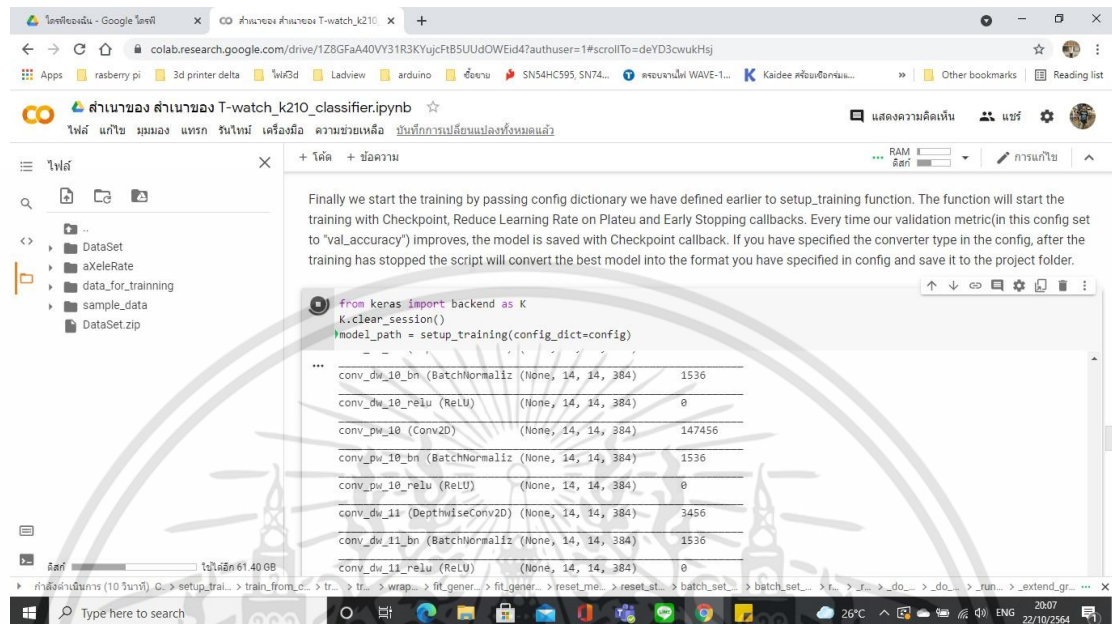
[('name': '/device:CPU:0',
  'device_type': 'CPU',
  'memory_limit': 268435456,
  'locality': {}),
 ('incarnation': 11295317662497435878, 'name': '/device:XLA_CPU:0',
  'device_type': 'XLA_CPU',
  'memory_limit': 17179869184,
  'locality': {}),
 ('incarnation': 17151998330560898712,
  'physical_device_desc': "device: XLA_CPU device")]

```

รูปที่ 4.1.7 การนำข้อมูลที่เทรนใน Colab ออกมาเป็น ไฟล์ Kmodel

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

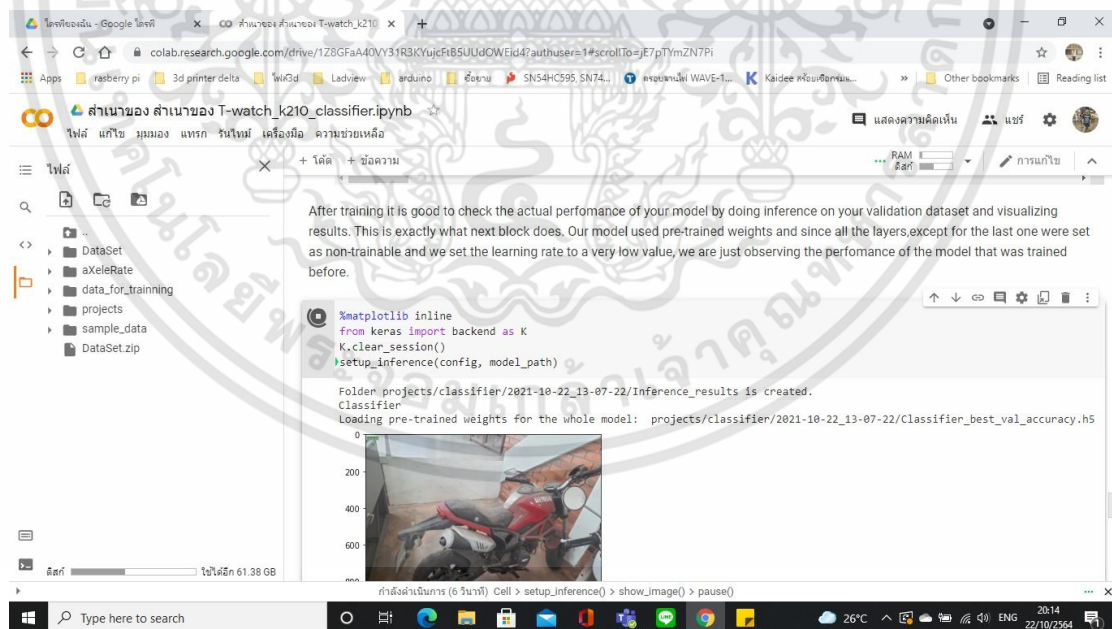
ขั้นตอนสุดท้าย : นำไฟล์ที่เทรนและได้ Output ออกมา ทำการตรวจสอบข้อมูลในไฟล์ และทำการสุ่มตัวอย่างข้อมูลในไฟล์มาตรวจสอบค่าความแม่นยำ



Finally we start the training by passing config dictionary we have defined earlier to `setup_training` function. The function will start the training with Checkpoint, Reduce Learning Rate on Plateau and Early Stopping callbacks. Every time our validation metric (in this config set to "val\_accuracy") improves, the model is saved with Checkpoint callback. If you have specified the converter type in the config, after the training has stopped the script will convert the best model into the format you have specified in config and save it to the project folder.

```
from keras import backend as K
K.clear_session()
model_path = setup_training(config_dict=config)
...
conv_dw_10_bn (BatchNormaliz (None, 14, 14, 384) 1536
conv_dw_10_relu (ReLU (None, 14, 14, 384) 0
conv_dw_10 (Conv2D (None, 14, 14, 384) 147456
conv_dw_10_bn (BatchNormaliz (None, 14, 14, 384) 1536
conv_dw_10_relu (ReLU (None, 14, 14, 384) 0
conv_dw_11 (DepthwiseConv2D (None, 14, 14, 384) 3456
conv_dw_11_bn (BatchNormaliz (None, 14, 14, 384) 1536
conv_dw_11_relu (ReLU (None, 14, 14, 384) 0
```

รูปที่ 4.1.8 นำไฟล์ที่เทรนและได้ Output ออกมา ทำการตรวจสอบข้อมูลในไฟล์



After training it is good to check the actual performance of your model by doing inference on your validation dataset and visualizing results. This is exactly what next block does. Our model used pre-trained weights and since all the layers, except for the last one were set as non-trainable and we set the learning rate to a very low value, we are just observing the performance of the model that was trained before.

```
%matplotlib inline
from keras import backend as K
K.clear_session()
setup_inference(config, model_path)

Folder projects/classifier/2021-10-22-13-07-22/Inference_results is created.
Classifier
Loading pre-trained weights for the whole model: projects/classifier/2021-10-22-13-07-22/Classifier_best_val_accuracy.h5
```

รูปที่ 4.1.9 การสุ่มตัวอย่างข้อมูลในไฟล์มาตรวจสอบค่าความแม่นยำ

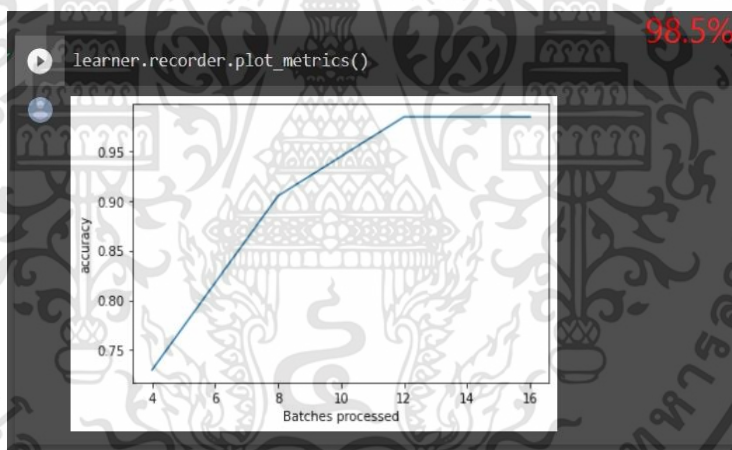
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 4.2 ผลการทดสอบโมเดลของเทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน

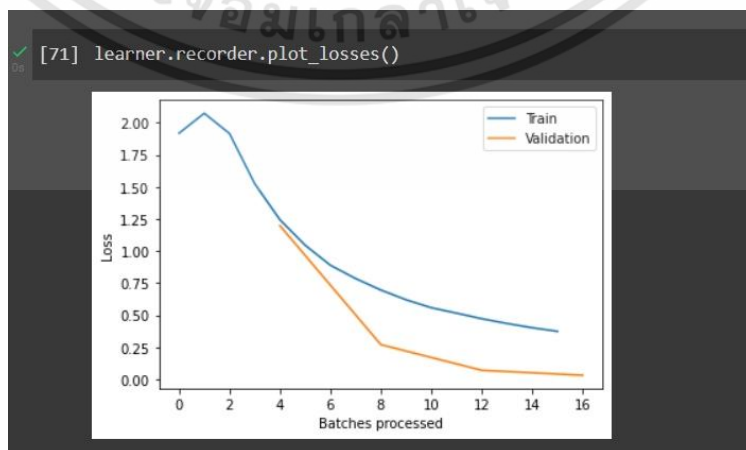
ผลการทดสอบโมเดลของเทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน ได้แก่ ผลการทดสอบประสิทธิภาพโมเดล ผลการทดสอบหาค่าเทรซโฮลด์ และผลการทดสอบโมเดลจากตารางคอนฟิวชันเมทริกซ์

### 4.2.1 ผลการทดสอบประสิทธิภาพ

ผลการทดสอบประสิทธิภาพจากกราฟอนุमान (Frozen inference graph) จากชุดข้อมูลสอนจำนวน 1000 ภาพ และชุด ข้อมูลตรวจสอบจำนวน 300 ภาพ โดยใช้เวลา 1 ชั่วโมง 30 นาที จากการสอน 20,000 รอบ ของ โมเดลที่ได้รับการสอนล่วงหน้า Faster R-CNN Resnet50 COCO พบว่าค่าแสดงผลจากTensorboard มีความแม่นยำเฉลี่ย (mean Average Precision : mAP) เท่ากับ 0.4310, ความ แม่นยำเฉลี่ยที่ไอโอยู 50 (mAP@.50IOU (Intersection Over Union)) เท่ากับ 0.9344, ความ แม่นยำเฉลี่ยที่ไอโอยู 75 mAP@.75JOU) เท่ากับ 0.4039 และการสูญเสียทั้งหมด (Total Loss) เท่ากับ 0.1464



รูปที่ 4.2ก กราฟแสดงความแม่นยำเฉลี่ย (mAP)

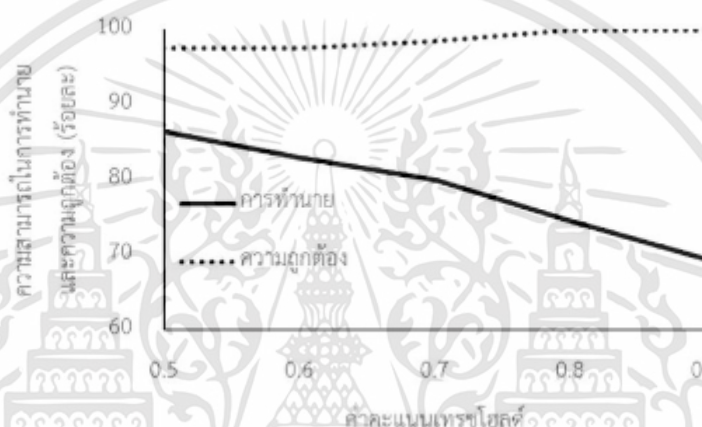


รูปที่ 4.2ข กราฟแสดงการสูญเสียทั้งหมด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

#### 4.2.2 ผลการทดสอบหาค่าคะแนนเทรซโฮลด์

ในการทำนายผลลัพธ์สำหรับตรวจจับความสมบูรณ์บริเวณปิดผนึก ต้องมีการกำหนดค่าคะแนนเทรซโฮลด์ สำหรับเป็นเกณฑ์ขั้นต่ำในการระบุประเภทของวัตถุว่าเป็นประเภทไหน ถ้าไม่ถึงเกณฑ์ที่กำหนดก็ไม่สามารถระบุประเภทวัตถุนั้นได้ แต่ถ้าถึงเกณฑ์ที่กำหนดจะมีการทำนายของภาพป้อนเข้า โดยผลลัพธ์มีทั้งถูกและผิดจากการทำนาย ผู้ทำวิจัยจึงสร้างกราฟความสัมพันธ์ระหว่างค่าร้อยละความสามารถในการทำนายและความถูกต้อง กับค่าเทรซโฮลด์ที่ 0.5-0.9 ดังรูปที่ 4.7 จากชุดข้อมูลตรวจสอบจำนวน 300 ภาพ



รูปที่ 4.2.2 กราฟความสัมพันธ์ระหว่างค่าร้อยละความสามารถในการทำนายและความถูกต้อง กับค่าเทรซโฮลด์ที่ 0.5-0.9

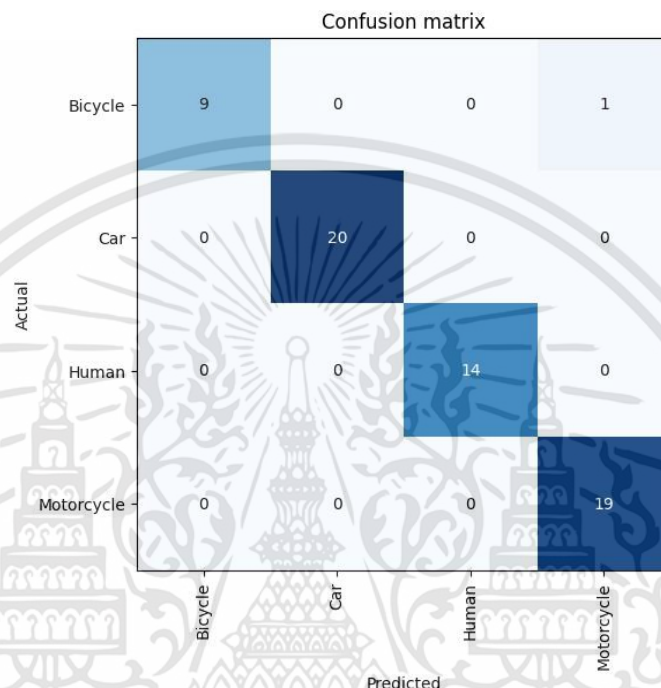
การทดสอบกราฟความสัมพันธ์ระหว่างค่าร้อยละความสามารถในการทำนายและความถูกต้อง กับค่าเทรซโฮลด์ที่ 0.5-0.9 (ดังรูปที่ 4.7) พบว่าค่าความถูกต้องของแต่ละค่าเทรซโฮลด์ มีค่าไม่แตกต่างกันมากนัก จึงเลือกค่าคะแนนเทรซโฮลด์ที่ 0.6 ที่มีมีค่าความสามารถการทำนายและความถูกต้องที่ยอมรับได้ รวมทั้งป้องกันการทำนายที่ผิดพลาดจากสิ่งอื่นที่อาจมารบกวน

#### 4.2.3 ผลการทดสอบตารางคอนฟิวชันเมทริกซ์ ของเทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน

หลังจากเสร็จสิ้นการสอนโมเดล และการกำหนดค่าค่าเทรซโฮลด์ ทำการทดสอบประสิทธิภาพ ของโมเดลสำหรับตรวจความสมบูรณ์ของบรรจุภัณฑ์ จากคอนฟิวชันเมทริกซ์ ด้วยชุดข้อมูลทดสอบ สำหรับทดสอบประสิทธิภาพโปรแกรมจำนวน 500 ภาพ และภาพถ่ายความรื้อนรูปแบบต่างๆที่ใช้ สำหรับเป็นตัวแทนของประเภทไม่สามารถจำแนกได้จำนวน 100 ภาพ ซึ่งใช้เวลาในการประมวลผล 5 วินาทีต่อภาพ สำหรับในการประมวลผล โดยแสดงผลลัพธ์ในรูปแบบของคอนฟิว

ชั้นเมทริกซ์ ดัง ตารางที่ 4.2a และค่าของความแม่นยำ, ความครบถ้วน, ค่าประสิทธิภาพโดยรวม และ ความถูกต้อง ดังตารางที่ 4.2b

ตารางที่ 4.2a คอนฟิวชันเมทริกซ์ของโมเดลตรวจสอบความสมบูรณ์ของบรรจุภัณฑ์ ด้วยเทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน จำนวนภาพและสัดส่วนการทำนาย)



ตารางที่ 4.2b ผลลัพธ์ค่าความแม่นยำ, ความครบถ้วน, ค่าประสิทธิภาพโดยรวม และ ความถูกต้อง ของเทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน (สัดส่วนการทำนาย)

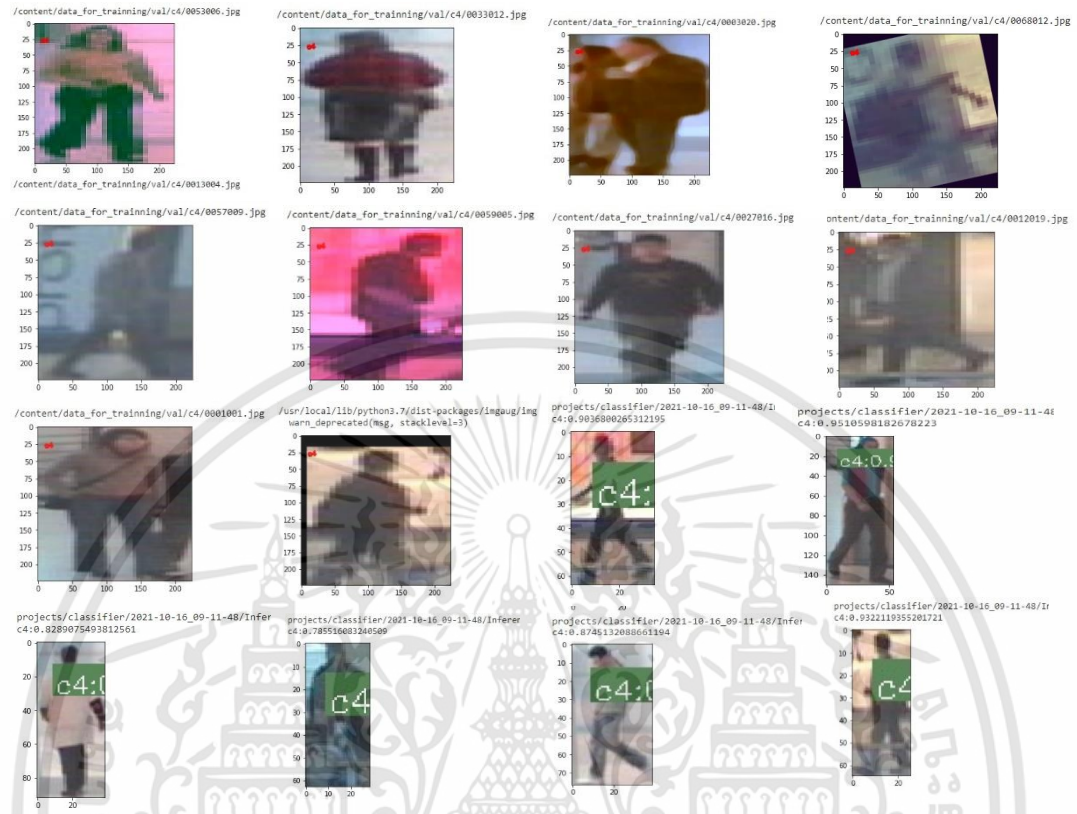
Results

TP: 62  
Overall Accuracy: 98.41%

Class	n (truth) ①	n (classified) ②	Accuracy	Precision	Recall	F1 Score
1	9	10	98.41%	0.90	1.0	0.95
2	20	20	100%	1.0	1.0	1.0
3	14	14	100%	1.0	1.0	1.0
4	20	19	98.41%	1.0	0.95	0.97

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 4.3 ผลการเทรนนิ่งตัวกลิ้งผ่านรูปภาพแต่ละประเภท



รูปที่ 4.3.ก ภาพคนที่น่ามาเทรนนิ่ง



รูปที่ 4.3.ข ภาพรถที่น่ามาเทรนนิ่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.3.ค ภาพรถจักรยานยนต์ที่นำมาเทรนนิ่ง



รูปที่ 4.3.ง ภาพรถยนต์ที่นำมาเทรนนิ่ง

#### 4.4 ทดสอบโดยนำอุปกรณ์มาติดตั้งที่รถแทรกเตอร์ไฟฟ้า

การทดสอบโดยใช้วัตถุจริงมาทดสอบขวงหน้ารถแทรกเตอร์ไฟฟ้าเพื่อดูผลลัพธ์ของตัวกล้องว่ากล้องสามารถตรวจจับวัตถุและสามารถควบคุมรถแทรกเตอร์ไฟฟ้าไม่ให้ชนกับวัตถุที่กีดขวางหน้ารถโดยมีการทดสอบดังรูป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.4.1 ทดสอบโดยนักรถจักรยานมาขวางหน้ารถ



รูปที่ 4.4.2 ทดสอบโดยให้มีคนมาขี่มาขวางหน้ารถ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.4.3 ทดสอบโดยนำรถยนต์มาขวางหน้ารถ



รูปที่ 4.4.4 ทดสอบโดยนำรถจักรยานยนต์มาขวางหน้ารถ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

#### 4.5 เปรียบเทียบผลการทดลอง

โดยผลการทดลองการวัดระยะความห่างที่รถหยุดนิ่งจากวัตถุกับรถโดยการวัดระยะและนำค่าทั้งหมดมาแสดงในตารางแสดงผลการทดลอง

ตารางที่ 4.5.1 แสดงผลการเปรียบเทียบระยะการหยุดนิ่งของรถระหว่างคนกับรถจักรยาน

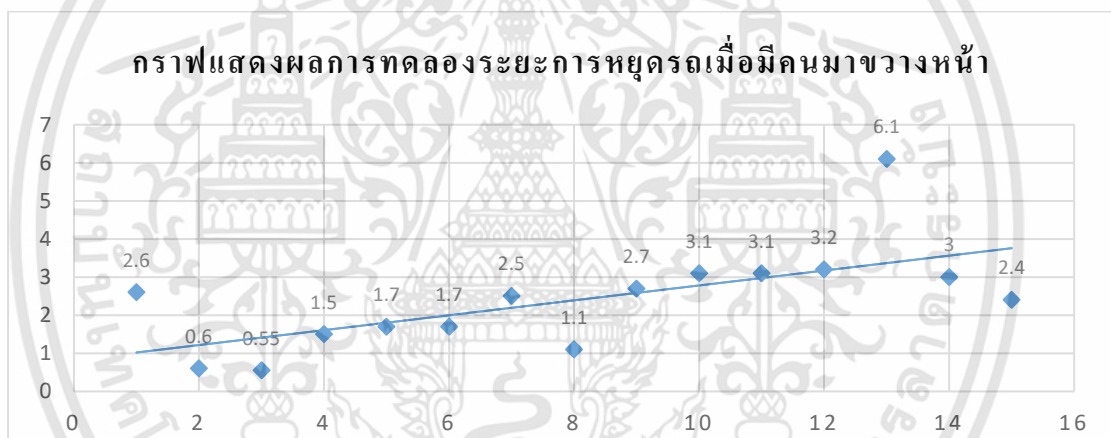
ตารางผลการทดลอง	คน	จักรยาน
ครั้งที่	ระยะการหยุดรถ (เมตร)	ระยะการหยุดรถ (เมตร)
1	2.6	1.3
2	0.6	1.2
3	0.55	1.4
4	1.5	1.5
5	1.7	1.3
6	1.7	4
7	2.5	2.8
8	1.1	3.1
9	2.7	4
10	3.1	2.6
11	3.1	3.5
12	3.2	1.8
13	6.1	2
14	3	2.9
15	2.4	2.6
ค่าเฉลี่ย	2.39	2.4

ตารางที่ 4.5.2 แสดงผลการเปรียบเทียบระยะการหยุดนิ่งของรถระหว่างรถจักรยานยนต์กับรถยนต์

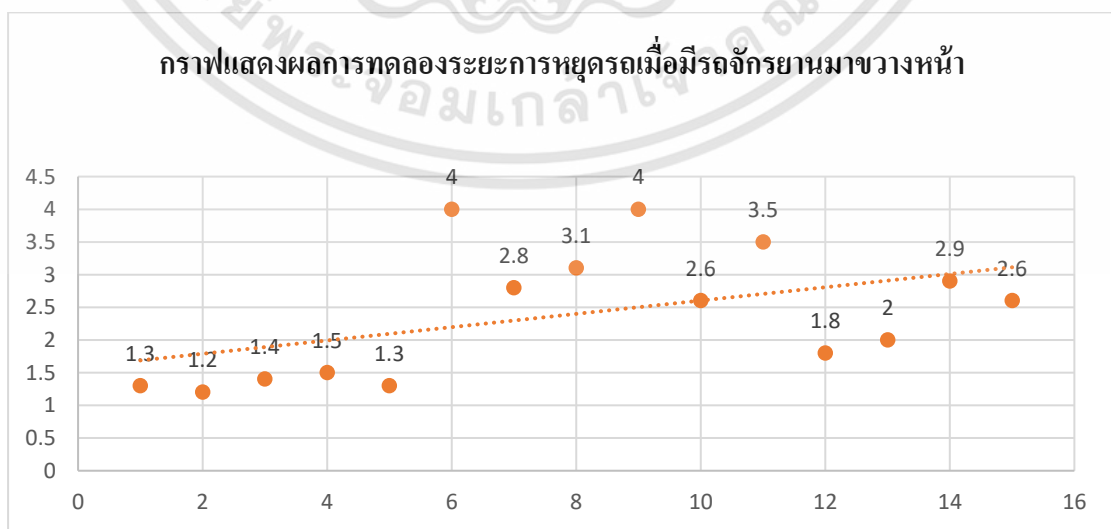
ตารางผลการทดลอง	จักรยานยนต์	รถยนต์
ครั้งที่	ระยะการหยุดรถ (เมตร)	ระยะการหยุดรถ (เมตร)
1	3.5	3.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2	5	3.2
3	2.4	2.6
4	5	1.6
5	3.7	2.8
6	3	6.5
7	6.5	4.5
8	2.7	3
9	3	1.6
10	4.1	3.5
ค่าเฉลี่ย	2.593333333	2.166666667

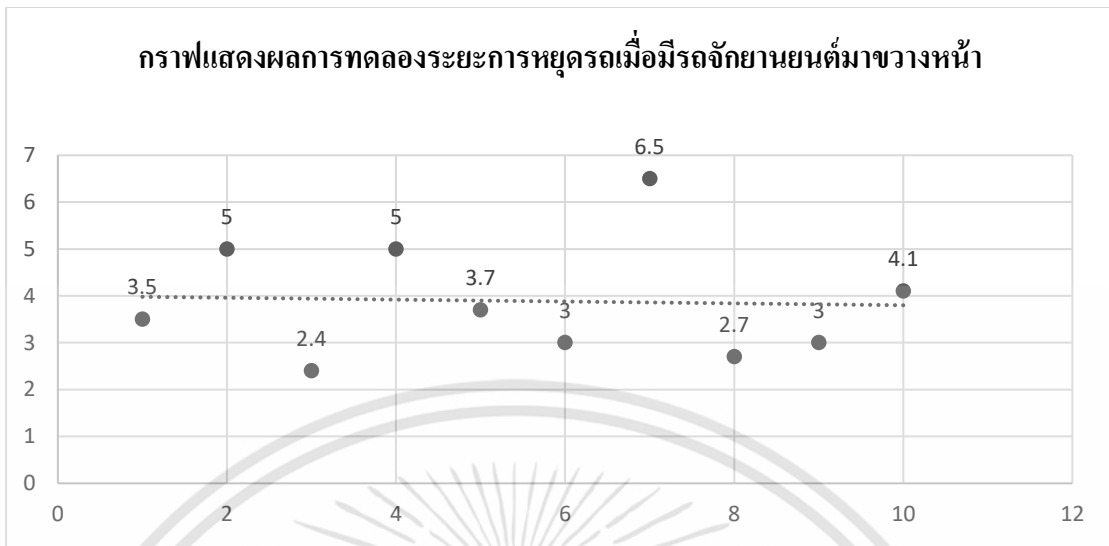


รูปที่ 4.5.ก กราฟแสดงผลระยะที่รถหยุดนิ่งเมื่อมีคนมาขวางหน้ารถ

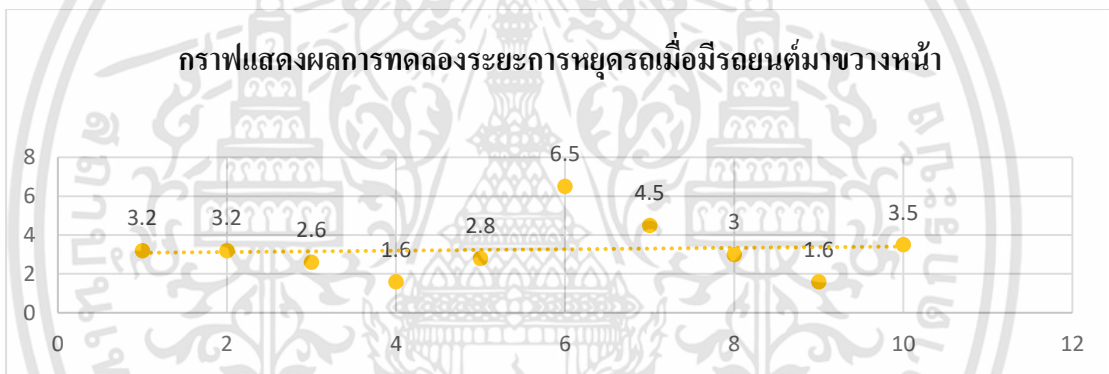


รูปที่ 4.5.ข กราฟแสดงผลระยะที่รถหยุดนิ่งเมื่อมีรถจักรยานมาขวางหน้ารถ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.5.ค กราฟแสดงผลระยะที่รถหยุดนิ่งเมื่อมีรถจักรยานยนต์มาขวางหน้ารถ



รูปที่ 4.5.ง กราฟแสดงผลระยะที่รถหยุดนิ่งเมื่อมีรถยนต์มาขวางหน้ารถ

ตารางที่ 4.5.3 แสดงผลการเปรียบเทียบระยะการหยุดนิ่งของรถระหว่างรถจักรยานยนต์กับรถจักรยานเมื่อมีแสงมากและเมื่อมีแสงน้อย

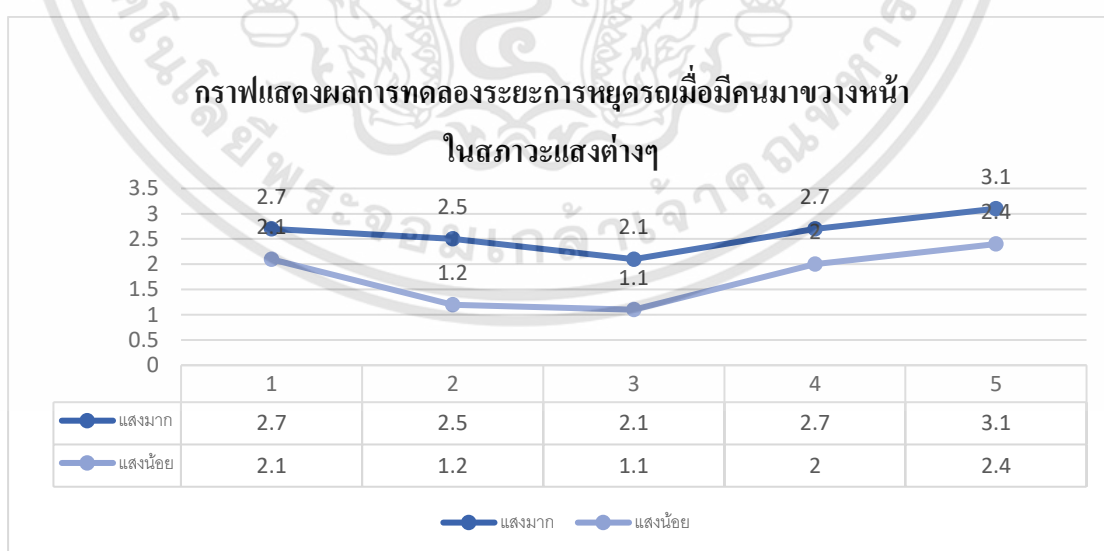
ตารางผลการทดลอง	จักรยาน		จักรยานยนต์	
	ระยะการหยุดรถเมื่อแสงมาก (เมตร)	ระยะการหยุดรถเมื่อแสงน้อย (เมตร)	ระยะการหยุดรถเมื่อแสงมาก (เมตร)	ระยะการหยุดรถเมื่อแสงน้อย (เมตร)
ครั้งที่ 1	4	3.5	3.5	3
ครั้งที่ 2	2.8	1.8	2.5	2.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3	3.1	2	2.4	3.1
4	4	2.9	5	3
5	2.6	2.6	3.7	5
เฉลี่ย	3.3	2.56	3.42	3.26

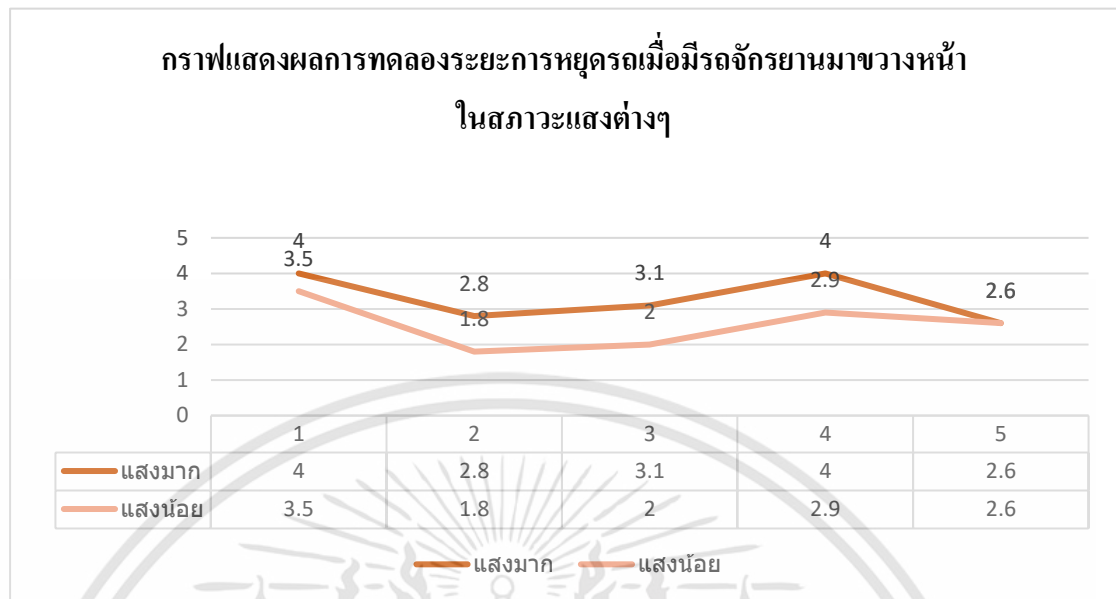
ตารางที่ 4.5.4 แสดงผลการเปรียบเทียบระยะการหยุดนิ่งของรถระหว่างคนกับรถยนต์เมื่อมีแสงมาก และเมื่อมีแสงน้อย

ตารางผล การทดลอง	คน		รถยนต์	
	ระยะการหยุด เมื่อแสงมาก (เมตร)	ระยะการหยุด เมื่อแสงน้อย (เมตร)	ระยะการหยุด เมื่อแสงมาก (เมตร)	ระยะการหยุด เมื่อแสงน้อย (เมตร)
ครั้งที่ 1	2.7	2.1	4	2.3
2	2.5	1.2	3.2	2.3
3	2.1	1.1	2.6	1.5
4	2.7	2	2.7	1.2
5	3.1	2.4	2.8	1.5
เฉลี่ย	2.62	1.76	3.06	1.76

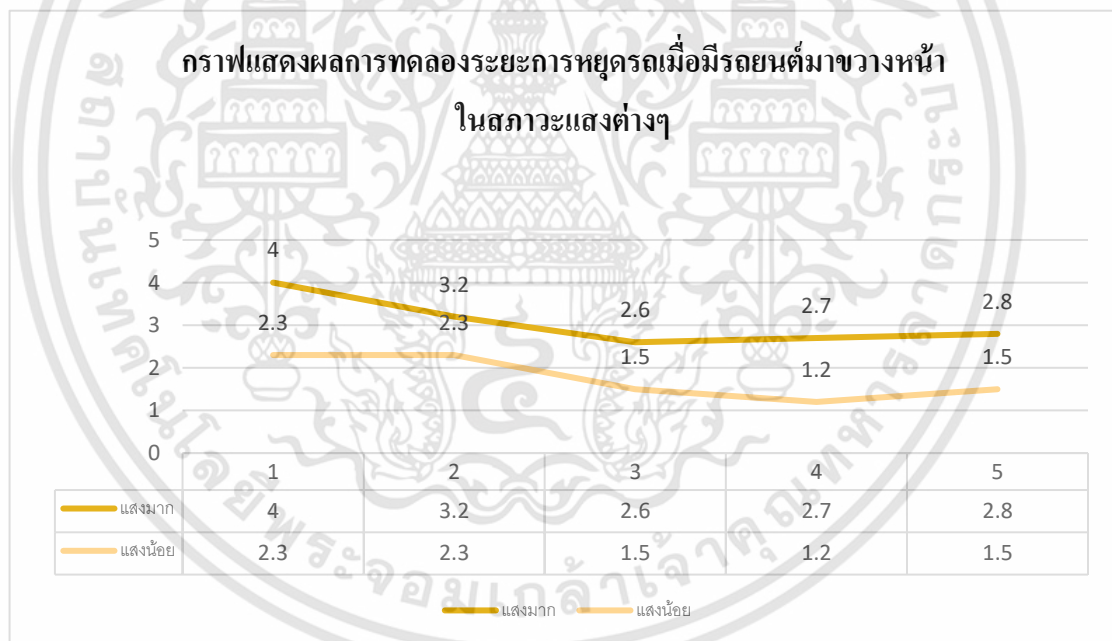


รูปที่ 4.5.a กราฟแสดงผลระยะที่รถหยุดนิ่งเมื่อมีคนมาขวางหน้ารถในสภาวะแสงต่างๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

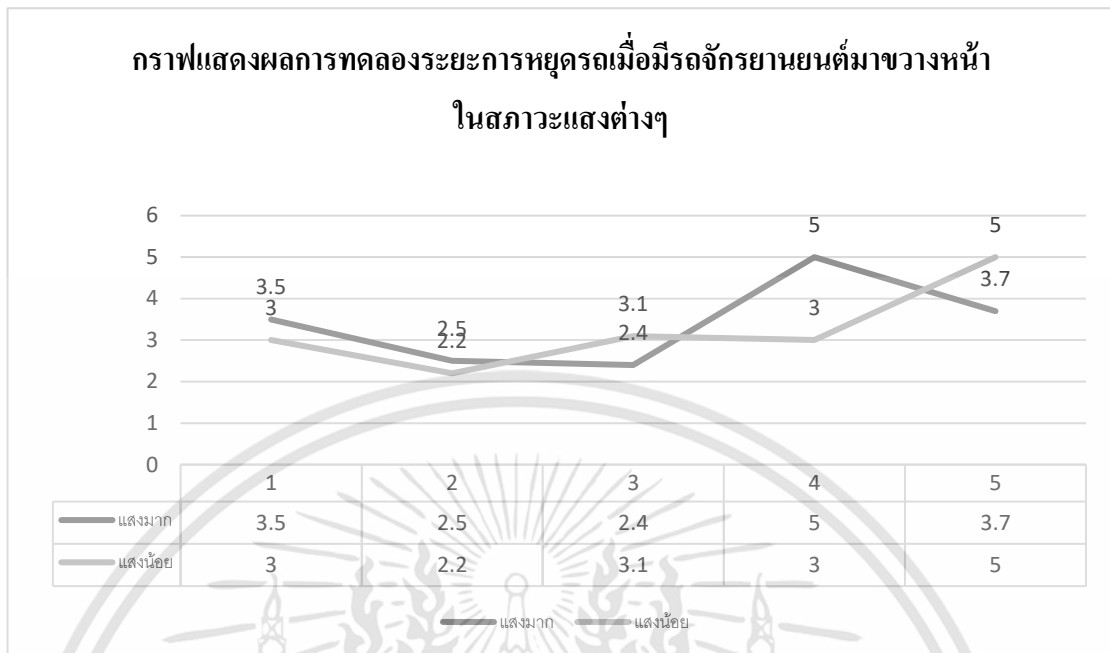


รูปที่ 4.5.b กราฟแสดงผลระยะที่รถหยุดนิ่งเมื่อมีรถจักรยานมาขวางหน้ารถในสภาวะแสงต่างๆ



รูปที่ 4.5.c กราฟแสดงผลระยะที่รถหยุดนิ่งเมื่อมีรถยนต์มาขวางหน้ารถในสภาวะแสงต่างๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.5.d กราฟแสดงผลระยะเวลาที่รถหยุดนิ่งเมื่อมีรถจักรยานยนต์มาขวางหน้ารถในสถานะแสงต่างๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 5

### สรุปผลการทดลอง

#### 5.1 สรุปผลการดำเนินงาน

จากการดำเนินงานโครงการ เรื่องระบบช่วยการขับขี่ยานยนต์ไฟฟ้าโดยใช้การประมวลผลภาพด้วยปัญญาประดิษฐ์ ทางผู้จัดทำสามารถสรุปผลการดำเนินงานได้ดังนี้

ผลจากการศึกษาข้อมูล การเก็บรวบรวมข้อมูล และการวิเคราะห์ตัวแปรที่มีต่อการควบคุมยานยนต์ไฟฟ้าโดยใช้กล้อง OV2640 camera และควบคุมด้วยบอร์ด AI CorgiDude ซึ่งผลการทดลองความแตกต่างระหว่างการตรวจจับวัตถุของกล้องในสภาวะแสงน้อยและแสงมาก โดยให้รถวิ่งอยู่ในเลนตามปกติแล้วให้วัตถุเคลื่อนที่ตัดหน้ารถในระยะที่เท่าๆกัน

จากการทดลองการตรวจจับของกล้องในสภาวะแสงน้อยและแสงมาก แล้วทำการวัดระยะการหยุดของรถเมื่อวัตถุเคลื่อนที่ตัดหน้ารถในขณะที่วิ่งในความเร็วคงที่ พบว่าค่ามีค่าเฉลี่ยที่แตกต่างกันไป โดยในสภาวะแสงน้อยกล้องระยะการตรวจจับจะสั้นกว่าในสภาวะแสงมาก

#### ตารางที่ 5.1 ตารางเปรียบเทียบระยะการหยุดในสภาวะต่างๆ

ตารางแสดงการเปรียบเทียบระยะการหยุดนิ่งของรถในสภาวะต่างๆ		
	แสงน้อย	แสงมาก
คน	1.76 m	2.63 m
รถจักรยาน	2.56 m	3.3 m
รถจักรยานยนต์	3.26 m	3.42 m
รถยนต์	1.76 m	3.06 m
ค่าเฉลี่ย	2.335	3.1025

จากตารางที่ 5.1 แสงนั้นเป็นปัจจัยหลักในการตรวจจับวัตถุ ทำให้ระยะในการตรวจจับในแต่ละสภาวะแสงจะไม่เท่ากัน โดยที่สภาวะแสงน้อยจะตรวจจับวัตถุได้สั้นกว่าสภาวะแสงมากเฉลี่ยอยู่ที่ 2.3 เมตร ขณะที่ในสภาวะแสงมากจะเฉลี่ยอยู่ที่ 3.1 เมตร ดังนั้นเมื่อแสงน้อยจะทำให้ประสิทธิภาพของกล้องลดลงแต่ระยะในการหยุดรถนั้นยังอยู่ในเกณฑ์ที่ยอมรับได้

#### 5.2 ปัญหาที่พบและแนวทางในการแก้ปัญหา

##### 5.2.1 ปัญหาที่พบ

1. เนื่องจากตัวกล้องมีมุมมองความกว้างน้อยทำให้การตรวจจับวัตถุอาจจะคลาดเคลื่อน เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. แรงเฉื่อยที่เกิดขึ้นจากการหยุดรถอาจทำให้ระยะในการหยุดรถคลาดเคลื่อน
3. กล้อง OV2640 มีความละเอียดที่ 2 ล้านพิกเซล ทำให้การตรวจจับวัตถุในสถานะแสงน้อยไม่ดีเท่าที่ควร
4. เนื่องจากรถที่ใช้ติดตั้งระบบมีพวงมาลัยแบบ manual ทำให้ไม่สามารถที่จะควบคุมการหักเลี้ยวโดยอัตโนมัติได้

### 5.2.2 แนวทางการแก้ไขปัญหา

1. ติดตั้งเซนเซอร์เพิ่มเติมรอบตัวรถทำให้ช่วยเรื่องการตรวจจับวัตถุดีขึ้น
2. ทำการติดตั้งเบรกมือเสริมเพื่อป้องกันกรณีที่รถมีแรงเฉื่อยมากจนการเบรกด้วยมอเตอร์ไม่หยุดในทันที
3. ติดตั้งไฟหน้ารถเพื่อเสริมประสิทธิภาพของกล้องในสถานะที่แสงน้อย

### 5.3 ข้อเสนอแนะ

1. ควรใช้กล้องที่มีความละเอียดที่สูงขึ้นเพื่อประสิทธิภาพการตรวจจับในสถานะแสงน้อย
2. เนื่องจากตอนนี้การควบคุมอัตโนมัติของรถอยู่ในเลเวล 1 หากเปลี่ยนเป็นพวงมาลัยไฟฟ้าจะสามารถควบคุมการหักเลี้ยวของรถได้ที่ทำให้การควบคุมอัตโนมัติเพิ่มเป็นเลเวล 2 – 3 ได้โดยง่าย
3. หากต่อยอดสามารถทำเป็นโมดูลเพื่อนำไปเสริมในรถยนต์ทั่วไปตามท้องตลาดเพื่อเพิ่มประสิทธิภาพของรถยนต์ได้

## บรรณานุกรม

- [1] Walsh J., Campbell S., Mahony N. and Carvalho A. "Deep Learning vs. Traditional Computer Vision", **Conference: Computer Vision Conference (CVC) 2019.**
- [2] Baldi P. "Gradient descent learning algorithm overview: a general dynamical systems perspective" **IEEE Transactions on Neural Networks**, vol. 6, no.1, 1995. pp. 182-195.
- [3] Vasilev I., Roelants P., Zocca V., Slater D. and Spacagna G. **Python Deep Learning - Second Edition**, Packt, 2019.
- [4] Chollet F. **Deep Learning with Python**, Manning, 2017.
- [5] Huang J., Rathod V., Sun C., Zhu M. Korattikara A., Fathi A., Frscher I., Wojna Z., Song Y. and Guadarrama S. "Speed/Accuracy Trade-Offs for Modern Convolutional Object Detectors", **IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**, 2017.
- [6] Girshick R., Donahue J., Darrell T. and Malik J. "Region-Based Convolutional Networks for Accurate Object Detection and Segmentation", **IEEE Transactions on Pattern Analysis and Machine Intelligence**, vol.38, 2016.
- [7] KDnuggets. 2020. **Deep Learning for Object Detection: A Comprehensive Review**. [Online]. Available : <https://www.kdnuggets.com/2017/10/deep-learning-object-detection-comprehensive-review.html>
- [8] Girshick R., "Fast R-CNN", **IEEE International Conference on Computer Vision (ICCV)**, 2015.
- [9] Ren S., He K., Girshick R. and Sun J., "Faster R-CNN: Towards Real-Time Objects Detection with Region Proposal Networks", **IEEE Transactions on Pattern**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

*Analysis and Machine Intelligence*, vol.39 , 2017, pp. 1137-1149.

[10] He K., Zhang X., Ren S. and Sun J., "Deep Residual Learning for Image Recognition", *IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*, 2016.

[11] Frans Coenen. 2020. *On The Use of Confusion Matrixes*. [Online]. Available: [https://cgi.csc.liv.ac.uk/~frans/Notes/confusion Matrixes2012-5-12.pdf](https://cgi.csc.liv.ac.uk/~frans/Notes/confusion%20Matrixes2012-5-12.pdf)



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## ภาคผนวก ก

### คอร์สโค้ด

#### ก.1 คู่มือการใช้งานบอร์ดคอร์สโค้ด CorgiDude

CorgiDude คือบอร์ดพัฒนาไมโครคอนโทรลเลอร์ประสิทธิภาพสูงที่มีความสามารถในการนำโมเดลทางคณิตศาสตร์ วิทยาการคำนวณ หรือแม้กระทั่งโมเดลปัญญาประดิษฐ์ Machine Learning อย่างการจำแนกภาพ ค้นหาวัตถุ ค้นหาใบหน้า มาทำการประมวลผลที่บอร์ดได้อย่างรวดเร็ว สามารถเชื่อมต่อ กล้อง จอภาพแสดงผล หรือเซนเซอร์อื่นมาต่อและเขียนโปรแกรมเพื่อควบคุมและสั่งการได้อย่างมีประสิทธิภาพ สามารถเขียนโปรแกรมหลักด้วยภาษา Python (micropython) ที่ง่ายต่อการเรียนรู้ หรือเลือกพัฒนาด้วยภาษา C/C++ เพื่อเข้าถึงประสิทธิภาพสูงสุดก็ได้เช่นกัน โดยสามารถเลือกพัฒนาได้ผลงานได้อย่างหลากหลาย เช่น

คุณสมบัติของ CorgiDude

- โมเดลจำแนกรูปภาพรูปภาพสิ่งของต่าง ๆ ได้ถึง 1000 ชนิด
- ทำเครื่องคัดแยกขยะ แยกกระป๋อง แก้วน้ำ
- ตรวจสอบวัตถุแปลกปลอมในสายพานลำเลียง
- ระบบคัดแยกพืชอัตโนมัติ
- โมเดลค้นหาวัตถุในภาพ
- ตรวจสอบมอเตอร์ไซค์บนทางเท้า
- เครื่องไล่สุนัขหน้าบ้าน
- เครื่องนับรถการจราจรบนท้องถนน
- โมเดลค้นหา / รู้จำใบหน้า
- ระบบเช็คชื่อ เข้างาน เข้าเรียนด้วยใบหน้า
- ระบบตรวจสอบใบหน้าก่อนเปิดประตู
- ระบบรู้จำลูกค้าโดยใช้ใบหน้า
- พัฒนาโมเดล AI เอง
- รถวิ่งอัตโนมัติ หลบสิ่งกีดขวาง
- ระบบพยากรณ์ล่วงหน้าจากข้อมูลเซนเซอร์
- ระบบรู้จำเสียง สั่งงานด้วยเสียง
- ค้นหาบาร์โค้ด QR Code
- ตรวจสอบสัตว์ตก ระบบประมวลผลภาพ
- เชื่อมต่อ WiFi ส่งการแจ้งเตือน ทำระบบ IoT ต่าง ๆ (เฉพาะในรุ่น WiFi)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# ภาคผนวก ข

## โปรแกรม Colab

### ข.1 การเริ่มต้นใช้งานโปรแกรม

Colab คืออะไร

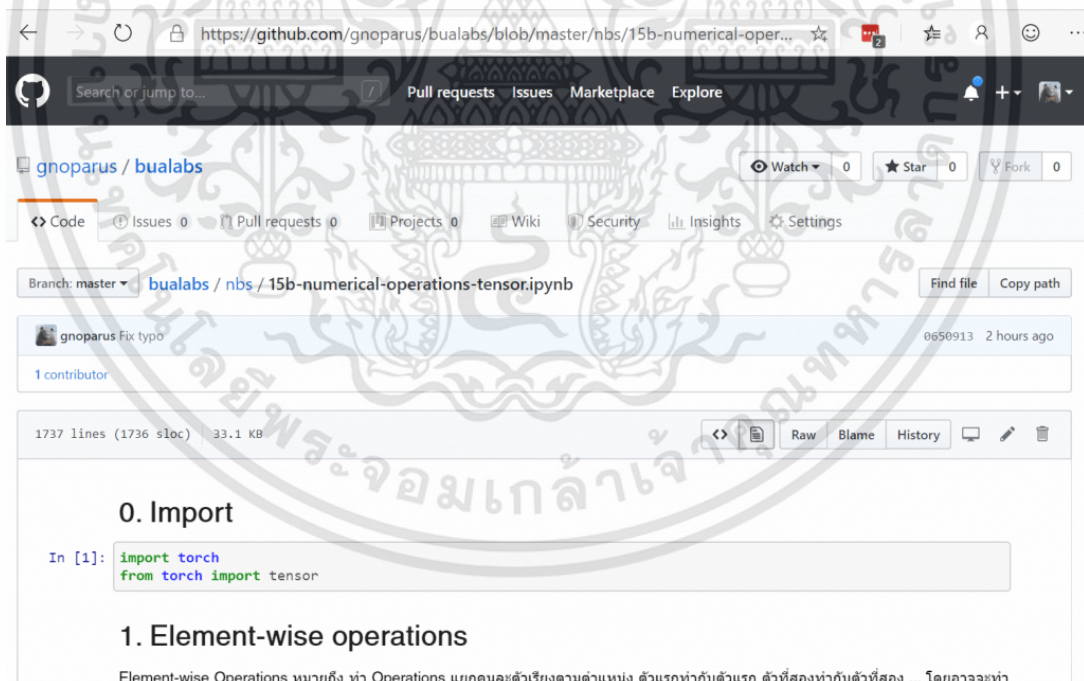
Google Colab ชื่อเต็มคือ Google Colaboratory เป็นบริการ Software as a Service (SaaS) โฮสต์โปรแกรม Jupyter Notebook บน Cloud จาก Google

Google Colab ใช้ยังไง เราสามารถใช้ Google Colab สร้าง Notebook เขียนโปรแกรมภาษา Python

#### 1. เข้าไปดู URL ของ Notebook ที่ GitHub

เข้าไปใน Repository ของเรา Browse ไปยัง Notebook ที่เราต้องการ แล้ว Copy URL มาเก็บไว้ก่อน สังเกตที่ตัวหนา หลัง github.com/ ตัวอย่างเช่น

<https://github.com/gnoparus/bualabs/blob/master/nbs/15b-numerical-operations-tensor.ipynb>



#### 2. นำ URL มาใส่ใน โค้ด HTML

เราจะนำ URL ด้านบน ตั้งแต่ <https://github.com/> เป็นต้นไป สังเกตตัวหนา ไปใส่ไว้ในลิงค์ หลัง <https://colab.research.google.com/github/> แล้วใส่ในช่อง Web Address ได้เลย เช่น โค้ดตัวอย่าง ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ลิงค์ Link

<https://colab.research.google.com/github/gnoparus/bualabs/blob/master/nbs/15b-numerical-operations-tensor.ipynb>

ปุ่ม HTML Button

```
<a
href="https://colab.research.google.com/github/gnoparus/bualabs/blob/master/nbs/15b-numerical-operations-tensor.ipynb">

</a>
```

ปุ่ม Markdown

```
[[OpenInColab](https://colab.research.google.com/assets/colabbadge.svg)](https://colab.research.google.com/github/gnoparus/bualabs/blob/master/nbs/15b-numerical-operations-tensor.ipynb)
```

### 3. Save และ Publish

Save และ Publish ขึ้น CMS ที่เราใช้จัดการเว็บไซต์อยู่ จะได้ปุ่มตั้งด้านล่าง

### 4. กดปุ่ม Open In Colab

จะเปิดหน้า Google Colab ขึ้นมา ถ้าเป็น Private Repository ก็จะให้ Login และทำการ Link Google Account และ GitHub Account ถ้าเป็น Public Repository ก็ไม่ต้องทำอะไรเลย

The screenshot shows a Google Colab notebook interface. The browser address bar displays the URL: `https://colab.research.google.com/github/gnoparus/bualabs/blob/master/n...`. The notebook title is `15b-numerical-operations-tensor.ipynb`. The menu bar includes `File Edit View Insert Runtime Tools Help`. The toolbar shows `+ Code + Text Copy to Drive` and a `Connecting` status. The notebook content is as follows:

```
[ ] import torch
    from torch import tensor
```

**0. Import**

**1. Element-wise operations**

Element-wise Operations หมายถึง ทำ Operations แยกคนละตัวเรียงตามตำแหน่ง ตัวแรกทำกับตัวแรก ตัวที่สองทำกับตัวที่สอง ... โดยอาจจะทำไปพร้อม ๆ กัน ทำมานานกันไปก็ได้

**1.1 Basic operations**

```
[ ] a = tensor([1., 2., 3., 4.])
```

บวก ลบ คูณ หาร ยกกำลัง กับ ตัวเลขธรรมดา ใส่ จุด หลังตัวเลข หมายถึง ให้เป็นเลขแบบ Float จุดทศนิยม รองรับการคำนวณคณิตศาสตร์ เช่น 1. เหมือนกับ 1.0

Google Collaps gnoparus/bualabs/nbs/15b จูปีเตอร์โน้ตบุ๊กภาพหน้าจอ

อาจจะมีข้อความเตือนว่า Notebook นี้ไม่ได้พัฒนาโดย Google เป็น Notebook ของผู้อื่น โปรดใช้ความระมัดระวังก่อน Run โปรแกรมใด ๆ

## 5. Save Notebook ไปยัง Google Drive หรือ GitHub

เมื่อเราแก้ไข ใช้งาน [Jupyter Notebook](#) เรียบร้อยแล้ว ถ้าเราต้องการจะ Save กลับไปยัง GitHub หรือ Google Drive ให้ไปที่เมนู

File → Save a copy in Drive หรือ File → Save a copy to GitHub