รายงานสหกิจศึกษาฉบับสมบูรณ์

ระบบอ่านป้ายทะเบียนอัตโนมัติ
Automated License Plate Recognition

นายณัฐดนัย บูรณะหิรัญวงษ์

หลักสูตรวิศวกรรมระบบควบคุม
ภาควิชาวิศวกรรมการวัดและควบคุม
คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2561

รายงานสหกิจศึกษาฉบับสมบูรณ์

ระบบอ่านป้ายทะเบียนอัตโนมัติ

Automated License Plate Recognition

นายณัฐดนัย บูรณะหิรัญวงษ์

หลักสูตรวิศวกรรมระบบควบคุม

ภาควิชาวิศวกรรมการวัดและควบคุม

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2561

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชื่อโครงงานสหกิจศึกษา	ระบบอ่านป้ายทะเบียนอัตโนมัติ
ชื่อ - สกุล นักศึกษา	นายณัฐดนัย บูรณะหิรัญวงษ์
คณะ	วิศวกรรมศาสตร์
สาขาวิชา	วิศวกรรมการวัดและควบคุม
ชื่อ - สกุล อาจารย์นิเทศ	ดร.พิชชา ประสิทธิ์มีบุญ
ชื่อ - สกุล ผู้นิเทศงาน	คุณชูศักดิ์ ชูแสงสุนทร
ชื่อสถานประกอบการ	บริษัท ซีเกท เทคโนโลยี (ประเทศไทย) จำกัด

บทคัดย่อ

โครงงานสหกิจศึกษานี้นำเสนอระบบอ่านป้ายทะเบียนอัตโนมัติ ซึ่งเป็นระบบที่ใช้ในการตรวจสอบรถที่เข้ามาจอดภายในบริษัท ซีเกท เทคโนโลยี (ประเทศไทย) จำกัด เพื่อเพิ่มความปลอดภัยในการจัดการกับรถที่เข้ามา ซึ่งทำงานโดยการติดตั้งกล้องบริเวณด้านหน้าบริษัท กล้องจะทำงานร่วมกับคอมพิวเตอร์ที่จะมี Software ทำงานในส่วนของการอ่านป้ายทะเบียนโดยใช้เทคนิคด้าน Image Processing เป็นโครงสร้างหลักของ Software และนำเอาเทคนิคด้าน Machine Learning เข้ามาประยุกต์ใช้ในบางส่วนของระบบเพื่อเพิ่มประสิทธิภาพให้กับ Software และแสดงผลการอ่านป้ายทะเบียนกับผู้ใช้งานผ่านหน้าจอ User Interface

คำสำคัญ: Image Processing , Machine Learning , ALPR

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Co-operative Title:	Automated License Plate Recognition
Student intern name:	Mr. Natdanai Buranahiranwong
Faculty:	Engineering
Department:	Instrumentation and Control Engineering
Advisor name:	Mr. Chousak Chousangsuntorn
Mentor name:	Dr. Pitcha Prasitmeeboon
Company:	Seagate Technology (Thailand) co. ltd

ABSTRACT

This project presents an automated license plate recognition system which is a system used to inspect vehicles that come to park inside Seagate Technology (Thailand) Co. Ltd. to increase the safety of handling incoming vehicles. This project needs to install the camera in front of company's gateway. The camera work with computer that running ALPR software to read license plate by using image processing technique as the main structure of this software and applying machine learning technique to apply in some parts of Software to increase efficiency. Then display result to user via user interface

keyword: Image Processing, Machine Learning, ALPR

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

การจัดทำโครงงานสหกิจฉบับสมบูรณ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความกรุณาของบริษัท ซีเกท เทคโนโลยี (ประเทศไทย) จำกัด ได้ให้โอกาสเข้าร่วมโครงการสหกิจศึกษาและทำให้ข้าพเจ้าได้เป็นส่วนหนึ่งในการทำงานกับบริษัท และผู้นิเทศงานคุณชูศักดิ์ ชูแสงสุนทรที่คอยให้ความช่วยเหลือตลอดเวลาที่เข้าร่วมโครงการ รวมถึงพี่พนักงานในแผนกทุกท่านที่ให้คอยสนับสนุนตลอดระยะเวลาหกเดือนของโครงการ

ขอขอบคุณ ผศ.ดร.ทัตยา ปุคคละนันทน์ อาจารย์ที่ปรึกษาที่ได้ให้คำแนะนำและเป็นที่ปรึกษาให้กับข้าพเจ้าตลอดเวลาของโครงการสหกิจศึกษา และขอขอบคุณอาจารย์ภาควิชาวิศวกรรมระบบควบคุมทุกท่านที่ได้ให้ความรู้และคำแนะนำอันเป็นประโยชน์ต่อการทำโครงงานสหกิจให้เสร็จสมบูรณ์

ผู้จัดทำ

ณัฐดนัย บูรณะหิรัญวงษ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้า
บทคัดย่อ	I
ABSTRACT	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญรูป	VII
สารบัญตาราง	IX
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญ	1
1.2 วัตถุประสงค์ของโครงงาน	1
1.3 ขอบเขตของโครงงาน	1
1.4 ขั้นตอนการศึกษา	2
1.5 ประโยชน์ที่คาดว่าจะได้รับ	3
บทที่ 2 ทฤษฎีและโครงงานที่เกี่ยวข้อง	4
2.1 โครงงานที่เกี่ยวข้อง	4
2.1.1 ระบบจำแนกป้ายทะเบียนอัตโนมัติ 2560	4
2.1.2 ระบบตรวจสอบทะเบียนและข้อมูลรถยนต์ภายใน 2555	4
2.2 Machine Learning (ML)	4
2.2.1 Supervised Machine Learning	5
2.2.2 Unsupervised Machine Learning	5
2.2.3 Reinforcement Machine Learning	6
2.3 Deep Learning	6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
2.4 Supported Vector Machine	7
2.5 การประมวลผลภาพ	8
2.6 Gray Scale	8
2.7 พิกเซล (Pixel)	9
2.8 Binary Image	11
2.9 Threshold Thresholding	12
2.10 ระบบสี	13
2.10.1 ระบบสี RGB	13
2.10.2 ระบบสีRYB	14
2.10.3 ระบบสีCMYK	14
2.10.4 ระบบสีHSV และ HSL	14
2.11 Connected Component Labeling	15
2.11 Histogram of Orientation Gradient (HOG)	20
บทที่ 3 ขั้นตอนการดำเนินงาน	26
3.1 ศึกษาและออกแบบ Software ที่ใช้ในการทำโปรแกรมอ่านป้ายทะเบียน	26
3.2.1 License Plate Detection	27
3.2.2 Character Segmentation	27
3.2.3 Character Recognition	27
3.2 ศึกษาและเลือกกล้องที่เหมาะสมกับการนำมาทำ Software ประมวลผลภาพ	28
3.3.1 Sensor	28
3.3.2 Resolution	28

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
3.3.3 Angle of View	28
3.3.4 Aperture	29
3.3.5 Focal Length	29
3.3.6 Framerate	29
3.3.7 S/N ratio	29
3.3 ทำ Software เพื่อตรวจจับป้ายทะเบียนจากภาพที่ได้จากกล้อง	31
3.4 ทำ Software เพื่อแยกตัวอักษรที่ติดกัน	36
3.5 ทำ Software เพื่อใช้ในการรู้จำตัวอักษร	37
3.6 ทำหน้าต่าง User Interface สำหรับใช้งาน	40
3.7 ติดตั้งกล้องและทดสอบโปรแกรม	42
บทที่ 4 ผลการดำเนินงาน	44
4.1 ผลจากการทำ Software เพื่อตรวจจับป้ายทะเบียนจากภาพที่ได้จากกล้อง	44
4.2 ผลการทดสอบ Software จำแนกตัวอักษรที่ติดกัน	46
4.3 ผลจากการทำ Software รู้จำตัวอักษร	50
4.4 ผลการทดลองกับวิดีโอจากกล้องวงจรปิด	56
บทที่ 5 สรุปผลการดำเนินงานและข้อเสนอแนะ	56
5.1 สรุปผลการดำเนินงาน	56
5.2 ปัญหาที่พบและแนวทางแก้ไข	56
5.3 ข้อเสนอแนะ	58
เอกสารอ้างอิง	59

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่	หน้า
2.1 แสดงการเปรียบเทียบความสว่างของแต่ละเฉดสี	9
2.2 แสดงให้เห็นถึงพิกเซลในภาพ	10
2.3 ภาพ Binary ต้นฉบับ	15
2.4 เริ่มอัลกอริที่มจากพิกเซลแรก	16
2.5 เริ่มอัลกอริที่มไป 3 พิกเซล	16
2.6 รูปแสดงเมื่อพิกเซลที่เจอไม่ติดกับ Label ก่อนหน้า	17
2.7 รูปแสดงเมื่อเจอพิกเซลที่รอบข้างมี 2 Label	18
2.8 เมื่อตรวจสอบครบทั้งภาพ	18
2.9 รูปเมื่ออัลกอริที่มทำงานสมบูรณ์	19
2.10 Preprocessing ภาพต้นฉบับ	20
2.11 Sobel Operator	21
2.12 แบ่งภาพเป็นเซลล์ 8 x 16	22
2.13 คำนวณ Gradient ในเซลล์	23
2.14 histogram ของเซลล์ ๆ หนึ่ง	23
2.15 กำหนดบล็อคที่ประกอบไปด้วยเซลล์ 2 x 2	25
2.16 ภาพเมื่อทำ HOG สมบูรณ์	25
3.1 ลำดับการทำงานของ Software	26
3.2 Hikvision 4 MP IR Fixed Bullet Network Camera	30
3.3 Dataset ป้ายทะเบียน	32
3.4 ตัวอย่างโปรแกรมจาก Github ของ AlexeyAB	33
3.5 การตั้งค่าไฟล์ cfg	34
3.6 dataset ตัวอักษรจากเว็บ NECTEC	38
3.7 ตัวอย่างการทำ HOG	38
3.8 ไฟล์ csv ที่ประกอบไปด้วยเวกเตอร์คุณลักษณะ	39

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่	หน้า
3.9 โปรแกรม Qt Creator	40
3.10 ตัวอย่างการเขียนโปรแกรมด้วย PyQt5	41
3.11 ภาพตัวอย่างการติดตั้งกล้อง	42
3.12 ภาพตัวอย่างการติดตั้งกล้อง	42
4.1 ภาพตัวอย่างการสอน YOLO เมื่อครบ 13301 รอบ	44
4.2 ภาพเมื่อทดลองกับคลิปวิดีโอถ่ายรถที่จอดอยู่	45
4.3 ภาพป้ายทะเบียนที่ตัดออกมาจากการตรวจเจอป้ายทะเบียน	45
4.4 ภาพป้ายทะเบียนเมื่อผ่าน Denoising	46
4.5 ภาพป้ายทะเบียนเมื่อผ่าน Gaussian Blur	46
4.6 ภาพป้ายทะเบียนต้นฉบับ	46
4.7 ภาพป้ายทะเบียนเมื่อทำการ Gray Scale	47
4.8 ภาพป้ายทะเบียนเมื่อทำ Threshold	47
4.9 ภาพเมื่อผ่าน Connected Component Labeling	47
4.10 ภาพเมื่อทำการหากลุ่มพิกเซลเยอะสุด 10 Label	48
4.11 ภาพเมื่อตัดตัวอักษรออกมาแต่ละตัว	48
4.12 ทำ HOG ตัวอักษร	50
4.13 หน้าต่าง User Interface	56

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่	หน้า
1.1 ระยะเวลาการทำงาน	2
2.1 ตารางเปรียบเทียบค่า bit per pixel	11
2.2 ตารางรายชื่อสีและค่าที่แทนค่าสีต่าง ๆ	13
4.1 แสดงให้เห็นถึงผลการแยกตัวอักษรที่ติดกันกับป้ายทะเบียนตัวอย่างทั้งหมด 25 รูป	49
4.2 แสดงผลตัวอย่างของการรู้จำตัวอักษรจากภาพป้ายทะเบียนตัวอย่าง 65 ภาพที่ได้มาจากขั้นตอนการจำแนกตัวอักษรที่ติดกัน	51
4.3 แสดงสาเหตุของการรู้จำที่ผิดพลาด	53

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญ

ปัจจุบันบริษัทซีเกทมีการตรวจสอบและจัดการกับรถที่เข้า - ออกบริษัทโดยการติดสติ๊กเกอร์ไว้ที่กระจกด้านหน้าของรถ และให้พนักงานรักษาความปลอดภัย (Security) ที่ประจำอยู่ป้อมที่ประตูหน้าบริษัทเป็นคนคอยตรวจสอบรถที่มีสติ๊กเกอร์กับรถที่ไม่มีสติ๊กเกอร์ รถที่มีสติ๊กเกอร์จะสามารถผ่านเข้าไปภายในบริษัทได้ ส่วนรถที่ไม่มีสติ๊กเกอร์จะต้องหยุดด้านหน้าประตูและติดต่อกับพนักงานรักษาความปลอดภัย ซึ่งเมื่อใช้ระบบนี้ในการตรวจสอบนี้ อาจก่อให้เกิดความผิดพลาดที่เกิดจากคนได้ (Human Error) ที่อาจจะเกิดจากความเหนื่อยล้าของพนักงาน หรือการที่รถเข้ามาเยอะจนไม่สามารถตรวจสอบได้อย่างละเอียด และส่งผลต่อความปลอดภัยของบริษัทในการควบคุมรถแต่ละคันที่เข้า - ออก

โครงงานนี้จึงนำเสนอวิธีการตรวจสอบรถที่เข้า – ออกบริษัทที่เป็นระบบอัตโนมัติด้วยคอมพิวเตอร์ที่ทำงานคู่กับกล้องวงจรปิดเพื่ออ่านป้ายทะเบียนของรถที่เข้ามาถึงด้านหน้าประตูของบริษัทและตรวจสอบสิทธิในการเข้าไปในบริษัทโดยอัตโนมัติ ซึ่งจะช่วยอำนวยความสะดวกในการตรวจสอบรถให้กับพนักงานรักษาความปลอดภัยมากขึ้น และเพิ่มความปลอดภัยให้กับบริษัทซีเกท

1.2 วัตถุประสงค์ของโครงงาน

1. จัดการควบคุมรถที่เข้า - ออกภายในบริษัทของพนักงานและบุคคลที่ไม่ใช่พนักงาน
2. อำนวยความสะดวกในการตรวจสอบรถที่เข้า – ออกบริษัทให้กับพนักงานรักษาความปลอดภัย
3. เพิ่มความปลอดภัยให้กับบริษัทซีเกท

1.3 ขอบเขตของโครงงาน

1. สามารถระบุป้ายทะเบียนของรถที่เข้ามาด้านหน้าบริษัทได้โดยมีความแม่นยำ 80% ขึ้นไป
2. สามารถบันทึกเวลาเข้าออกของรถที่เข้าออกแต่ละคันได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. สามารถแสดงผลของป้ายทะเบียนที่ตรวจสอบมาขึ้นบนจอแสดงผลได้

1.4 ขั้นตอนการศึกษา

1. ศึกษาและหา Tool ที่ใช้ในแต่ละส่วนของโปรแกรมที่เขียน
2. ศึกษาและเลือกอุปกรณ์ที่จะนำมาใช้ในระบบ
3. ศึกษาการเขียนโปรแกรม
4. เขียนโปรแกรมส่วนต่าง ๆของระบบ
5. ทดลองติดตั้งกล้องในมุมต่าง ๆ
6. ทดสอบและแก้ไขโปรแกรม
7. จัดทำเอกสารวิธีการใช้โปรแกรม

ตารางที่ 1.1 ระยะเวลาการทำงาน

แผนการดำเนินงาน	เดือน	มิถุนายน				กรกฎาคม				สิงหาคม			
	สัปดาห์ที่	1	2	3	4	5	6	7	8	9	10	11	12
หา tool สำหรับเขียนโปรแกรม													
ศึกษาและเลือกอุปกรณ์													
ศึกษาการเขียนโปรแกรม													
เขียนโปรแกรมส่วนต่าง ๆ													

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 1.1 ระยะเวลาการทำงาน (ต่อ)

แผนการดำเนินงาน	เดือน	กันยายน				ตุลาคม				พฤศจิกายน			
	สัปดาห์ที่	13	14	15	16	17	18	19	20	21	22	23	24
เขียนโปรแกรมส่วนต่าง ๆ													
ทดลองติดตั้งอุปกรณ์													
ทดสอบและแก้ไขโปรแกรม													
ทำเอกสารวิธีการใช้โปรแกรม													

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. สามารถนำโครงงานไปพัฒนาต่อไปให้เป็นระบบที่ดีขึ้นได้
2. นำความรู้ด้าน Image Processing และ Machine Learning ไปประยุกต์ใช้ต่อด้านอื่น ๆ
3. ได้รับความรู้และประสบการณ์ทำงานในสถานประกอบการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ทฤษฎีและโครงงานที่เกี่ยวข้อง

2.1 โครงงานที่เกี่ยวข้อง

2.1.1 ระบบจำแนกป้ายทะเบียนอัตโนมัติ 2560

นาย ชัยธวัช มินเดินเรือ ได้จัดทำระบบจำแนกป้ายทะเบียนรถอัตโนมัติ สามารถนำมาใช้ร่วมกับลานจอดรถ ในตำมหมู่บ้านเพื่อ เก็บบันทึก รถยนต์ของลูกบ้านที่เข้ามาจอดในลานจอดรถ ว่ามีการเข้าและออก ในเวลาใด เพื่อง่ายต่อการ สืบค้นข้อมูลรถต่าง ๆ ในกรณีที่เกิดการโจรกรรมรถยนต์โดยมีการตรวจสอบที่จุด CheckPoint และถ้าเกิดสิ่งผิดปกติจะมีการแจ้งเตือนหรือ Alert ไปที่เจ้าของบ้าน โดยใช้การทำ Image Processing บน ไลบรารี่ OpenCV

2.1.2 ระบบตรวจสอบทะเบียนและข้อมูลรถยนต์ภายใน 2555

โครงงานนี้มีการใช้ความรู้ในด้านการประมวลผลภาพ ระบบฐานข้อมูลเพื่อเพิ่มประสิทธิภาพในการตรวจสอบรถยนต์เข้าออกภายในมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรีโดยจะมีการบันทึกเวลาเข้าออกของรถแต่ละคันที่เข้าออกภายในมหาวิทยาลัยและสามารถแสดงข้อมูลของรถยนต์พร้อมทั้งบุคคลที่เกี่ยวข้องมีการนำข้อมูลรถจากกองพัฒนาอาคารและสถานที่

2.2 Machine Learning (ML)

เป็นขั้นตอนวิธี (Algorithm) ที่นำทฤษฎีสถิติศาสตร์มาประยุกต์ใช้เพื่อทำให้ระบบคอมพิวเตอร์มีการพัฒนาประสิทธิภาพในหน้าที่ๆหนึ่งด้วยตัวเครื่องเอง ซึ่ง Machine Learning จะมีการสร้างโมเดลทางคณิตศาสตร์ของข้อมูลตัวอย่าง (Sample Data) เพื่อใช้เป็นข้อมูลในการฝึกสอนระบบคอมพิวเตอร์ (Training Data) Machine Learning แบ่งออกเป็น 3 ประเภท ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.1 Supervised Machine Learning

เป็นการเรียนรู้ในลักษณะที่ผู้เขียนโปรแกรมมีชุดข้อมูลอินพุตและเอาท์พุตที่จับคู่เชื่อมต่อกันไว้ (Mapping) อยู่ เพื่อให้ระบบคอมพิวเตอร์เรียนรู้ เมื่อคอมพิวเตอร์รู้เอาท์พุตของข้อมูลแล้ว จะเริ่มหาความสัมพันธ์ของอินพุตเพื่อโยงไปหาเอาท์พุตให้ได้ว่าอินพุตแบบไหนจะให้เอาท์พุตออกมาแบบไหน ยกตัวอย่างเช่น ต้องการให้คอมพิวเตอร์แยกแอปเปิลกับส้ม ข้อมูลอินพุตที่เตรียมไว้คือลักษณะต่าง ๆ ของแอปเปิลและส้ม เช่น แอปเปิลมีผลสีแดง ส้มมีผลสีส้ม และคอมพิวเตอร์จะทำการเรียนรู้ความแตกต่างของส้มกับแอปเปิลนี้ ระบบคอมพิวเตอร์จะทำการสร้างโมเดลที่มาจากการเรียนรู้จากชุดข้อมูลที่เตรียมไว้ แบ่งออกเป็น 2 ชนิด

1. Classification จัดการกับปัญหาที่มีค่าเชิงเส้น (Discrete Value) เป็นการจัดหมวดหมู่ของข้อมูล ซึ่งเอาท์พุตที่ได้จะมีลักษณะเป็นชนิด (Class) ของข้อมูล เช่น การทำนายว่าอีเมลที่เข้ามาใหม่เป็นอีเมลสแปมหรือไม่ การทำนายว่าหุ้นจะราคาขึ้นหรือลง

2. Regression จะใช้กับปัญหาที่เราต้องการเอาท์พุตเป็นตัวเลข หรือเป็นเมื่อต้องการที่จะทำนายค่าออกมาเป็นตัวเลขจำนวนจริงหรือการทำนายข้อมูลที่มีค่าต่อเนื่อง (Continuous Value) เช่น เมื่อปัญหาของผู้ใช้คืออยากรู้ว่ามีงบราคา X บาท จะสามารถซื้อบ้านได้ขนาดใหญ่แค่ไหน

2.2.2 Unsupervised Machine Learning

เป็นการเรียนรู้ในลักษณะแบบไม่มีผู้สอน ผู้เขียนโปรแกรมจะมีชุดข้อมูลของอินพุต แต่ไม่ได้บอกว่าเอาท์พุตของข้อมูลเหล่านั้นคืออะไร ซึ่งคอมพิวเตอร์จะทำหน้าที่เพียงแค่จัดหมวดหมู่ของข้อมูลอินพุตที่มีลักษณะ (Features) คล้ายกันให้อยู่ในกลุ่มเดียวกัน ซึ่งก็จะแยกชนิดข้อมูลออกจากกันเป็นอย่างน้อย 2 กลุ่ม ตัวอย่างเช่น แอปเปิลกับส้ม Unsupervised ML จะทำหน้าที่แยกลักษณะของสองสิ่งนี้เพียงแค่ว่าชนิดหนึ่งมีผลสีแดง เป็นผลไม้ชนิดที่ 1 อีกชนิดหนึ่งมีผลสีส้ม เป็นผลไม้ชนิดที่ 2 แต่จะไม่เข้าใจว่าแต่ละอย่างนั้นเรียกว่าอะไร เมื่อเราส่งข้อมูลที่ต้องการทดสอบ(Test Data) เข้าไปเป็นแอปเปิล Unsupervised ML ก็จะทำนายออกมาว่าเป็นผลไม้ชนิดที่ 1 ที่คล้ายกับที่ได้แบ่งกลุ่มไว้ Unsupervised Machine Learning แบ่งออกเป็น 2 ชนิด

1. Cluster Analysis คือการนำข้อมูลมาแบ่งกลุ่ม จัดหมวดหมู่ตามลักษณะ (Features) ที่คล้ายกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. Association Analysis คือการหาความสัมพันธ์ของข้อมูลจำนวน 2 กลุ่มขึ้นไป เพื่อหาสิ่งที่ข้อมูลมีลักษณะเหมือนกัน ใช้กันอย่างมากในการทำเหมืองข้อมูล (Data Mining)

2.2.3 Reinforcement Machine Learning

เป็นการเรียนรู้แบบเสริมกำลัง ผู้เขียนโปรแกรมจะมีการสร้างเป้าหมายที่อยากให้คอมพิวเตอร์ทำตามไว้ โดยใช้รางวัลหรือแรงจูงใจมาช่วย (Reward) เมื่อทำสิ่งที่ตรงตามกับเป้าหมาย จะได้รับ Reward ไป ซึ่งอาจจะเป็นการได้รับคะแนนเพื่อบอกกับคอมพิวเตอร์ว่าทำถูกต้อง เมื่อทำผิดพลาด ก็จะถูกหักคะแนนออกเพื่อให้คอมพิวเตอร์เลือกที่จะไม่ทำตาม โดยจะไม่มี Supervisor มาบอกว่าควรทำแบบไหน ถูกนำมาใช้บ่อยกับ Robot Navigation และการเล่นเกม ยกตัวอย่างเช่น การเขียนโปรแกรมเพื่อเล่นเกมมาริโอ เมื่อเล่นแล้วมาริโอตายจะถูกหักคะแนน เมื่อสามารถเก็บเห็ดหรือเหรียญได้จะได้รับคะแนน และคอมพิวเตอร์จะเรียนรู้ด้วยตัวเองไปเรื่อยๆจนสามารถเล่นจนจบเกม

2.3 Deep Learning

การเรียนรู้เชิงลึก (อังกฤษ: Deep learning) เป็นสาขาของการเรียนรู้ของเครื่อง พื้นฐานของการเรียนรู้เชิงลึกคือ อัลกอริทึมที่พยายามจะสร้างแบบจำลองเพื่อแทนความหมายของข้อมูลในระดับสูงโดยการสร้างสถาปัตยกรรมข้อมูลขึ้นมาที่ประกอบไปด้วยโครงสร้างย่อย ๆหลายอัน และแต่ละอันนั้นได้มาจากการแปลงที่ไม่เป็นเชิงเส้น การเรียนรู้เชิงลึก อาจมองได้ว่าเป็นวิธีการหนึ่งของการเรียนรู้ของเครื่องที่พยายามเรียนรู้วิธีการแทนข้อมูลอย่างมีประสิทธิภาพ ตัวอย่างเช่น รูปภาพภาพหนึ่ง สามารถแทนได้เป็นเวกเตอร์ของความสว่างต่อจุดพิกเซล หรือมองในระดับสูงขึ้นเป็นเซ็ตของขอบของวัตถุต่าง ๆ หรือมองว่าเป็นพื้นที่ของรูปร่างใด ๆก็ได้ การแทนความหมายดังกล่าวจะทำให้การเรียนรู้ที่จะทำงานต่าง ๆทำได้ง่ายขึ้น ไม่ว่าจะเป็นการรู้จำใบหน้าหรือการรู้จำการแสดงออกทางสีหน้า การเรียนรู้เชิงลึกถือว่าเป็นวิธีการที่มีศักยภาพสูงในการจัดการกับฟีเจอร์สำหรับการเรียนรู้แบบไม่มีผู้สอนหรือการเรียนรู้แบบกึ่งมีผู้สอน นักวิจัยในสาขานี้พยายามจะหาวิธีการที่ดีขึ้นในการแทนข้อมูลแล้วสร้างแบบจำลองเพื่อเรียนรู้จากตัวแทนของข้อมูลเหล่านี้ในระดับใหญ่ บางวิธีการก็ได้แรงบันดาลใจมาจากสาขาประสาทวิทยาขั้นสูง โดยเฉพาะเรื่องกระบวนการตีความหมายในกระบวนการประมวลผลข้อมูลในสมอง ตัวอย่างของกระบวนการที่การเรียนรู้เชิงลึกนำไปใช้ได้แก่ การเข้ารหัสประสาท อันเป็นกระบวนการหาความสัมพันธ์ระหว่างตัวกระตุ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กับการตอบสนองของเซลล์ประสาทในสมอง นักวิจัยด้านการเรียนรู้ของเครื่องได้เสนอสถาปัตยกรรมการเรียนรู้หลายแบบบนหลักการของการเรียนรู้เชิงลึกนี้ ได้แก่ โครงข่ายประสาทเทียมแบบลึก (Deep Artificial Neural Networks) โครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Networks) โครงข่ายความเชื่อแบบลึก (Deep Belief Networks) และโครงข่ายประสาทเทียมแบบวนซ้ำ (Recurrent Neural Network) ซึ่งมีการนำมาใช้งานอย่างแพร่หลายในทางคอมพิวเตอร์วิทัศน์ การรู้จำเสียงพูด การประมวลผลภาษาธรรมชาติ การรู้จำเสียง และชีวสารสนเทศศาสตร์

2.4 Supported Vector Machine

เป็นอัลกอริทึมที่สามารถนำมาช่วยแก้ปัญหาการจำแนกข้อมูล ใช้ในการวิเคราะห์ข้อมูลและจำแนกข้อมูล โดยอาศัยหลักการของการหาสัมประสิทธิ์ของสมการเพื่อสร้างเส้นแบ่งแยกกลุ่มข้อมูลที่ถูกป้อนเข้าสู่กระบวนการสอนให้ระบบเรียนรู้ โดยเน้นไปยังเส้นแบ่งแยกแยะกลุ่มข้อมูลได้ดีที่สุด เกิดจากการที่นำค่าของกลุ่มข้อมูลมาวางลงในฟีเจอร์สเปซ (Feature Space) จากนั้นจึงหาเส้นที่ใช้แบ่งข้อมูลทั้งสองออกจากกันโดยจะสร้างเส้นแบ่ง (Hyperplane) ที่เป็นเส้นตรงขึ้นมา และเพื่อให้ทราบว่าเส้นตรงที่แบ่งสองกลุ่มออกจากกันนั้น เส้นตรงใดเป็นเส้นที่ดีที่สุด สำหรับรากฐานเดิมของ Support Vector Machine ถูกนำมาใช้กับข้อมูลที่เป็นเชิงเส้น แต่ในความเป็นจริงแล้วข้อมูลที่นำมาใช้ในระบบการสอนให้ระบบเรียนรู้ส่วนใหญ่มักเป็นข้อมูลแบบไม่เป็นเชิงเส้น ซึ่งสามารถแก้ปัญหาดังกล่าวด้วยการนำ Kernal Function มาใช้ การจำแนกข้อมูลบนระนาบหลายมิติ จะใช้ส่วนการเลือกที่มีความเหมาะสมที่สุดเรียกว่าโครงสร้างในการคัดเลือก (feature selection) ซึ่งโครงสร้างในการคัดเลือกมาจากข้อมูลที่สอนให้ระบบเรียนรู้ จำนวนเซตของโครงสร้างที่ใช้อธิบายในกรณีหนึ่ง เรียกว่า เวกเตอร์ (vector) ดังนั้นจุดมุ่งหมายของตัวแบบ SVM คือ แบ่งแยกกลุ่มของเวกเตอร์ในกรณีนี้ด้วยหนึ่งกลุ่มของตัวแปรเป้าหมายที่อยู่ข้างหนึ่งของระนาบ และกรณีของกลุ่มอื่นที่อยู่ทางระนาบต่างกัน ซึ่งเวกเตอร์ที่อยู่ข้างระนาบหลายมิติทั้งหมดเรียกว่า ซัพพอร์ตเวกเตอร์(Support Vectors) สมมติว่าเราต้องการจำแนกข้อมูลออกเป็น 2 กลุ่ม โดยใช้เส้นแบ่งที่เป็นเส้นตรง จะเห็นว่ามีเส้นตรงจำนวนมากที่สามารถจำแนกได้ แต่เส้นตรงเส้นไหนที่ดีที่สุด เราจะนิยาม Margin เป็นผลรวมระยะห่างของเส้นตรงที่เป็นเส้นแบ่ง ถึงเส้นตรงที่ผ่านข้อมูลที่ใกล้ที่สุดและขนานกับเส้นแบ่งของทั้งสองกลุ่ม จะเห็นว่า H1 แม้จะสามารถแบ่งข้อมูลทั้งสองกลุ่มออกได้เช่นกัน แต่ระยะในการแบ่งจากเส้นแบ่งไปถึงข้อมูลที่ใกล้ที่สุดนั้นมีขนาดน้อย แต่จากเส้น H2 จะเป็นเส้นที่แบ่งกลุ่ม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ที่กว้างมากที่สุดของทั้งสองกลุ่มคือให้ค่า maximum margin เราเรียกข้อมูลที่อยู่บน margin นี้ว่า Support Vector

2.5 การประมวลผลภาพ

Image Processing เรียกอีกอย่างหนึ่งว่า Image processing หมายถึง การเรียกใช้ขั้นตอนหรือกรรมวิธีใด ๆ มากระทำกับภาพ โดยมีวัตถุประสงค์เพื่อปรับปรุงคุณภาพ ให้ได้ภาพใหม่ที่มี คุณสมบัติตามต้องการ โดยทั่วไปแล้ววัตถุประสงค์ของ Image processing คือ

1. Image processing (Image in image out) เป็นขั้นตอนหรือวิธีที่มีการใช้กระบวนการทำการประมวลผลภาพเพื่อทำให้ได้ภาพออกมาอย่างสมบูรณ์เช่นการตกแต่งภาพด้วยโปรแกรม Photoshop เป็นต้น

2. Image analysis (Image in measurements out) เป็นขั้นตอนหรือวิธีที่มีการใช้กระบวนการทางการประมวลภาพ เพื่อทำให้ได้ค่าการวัดออกมา เช่นการวัดขนาดในงานอุตสาหกรรมเป็นต้น

3. Image understanding (Image in high-level description out) เป็นขั้นตอนหรือวิธีที่มีการใช้กระบวนการทางการประมวลผลภาพเพื่อให้ได้ ผลลัพธ์ออกมาเป็นความหมาย และความหมายนั้นสามารถที่จะทำให้เกิดประโยชน์สามารถนำมา ประยุกต์ใช้ต่อได้ตัวอย่างเช่นการรู้จำหรือการจดจำตัวอักษร (Optical Character Recognition หรือ OCR) เป็นต้น รูปแบบในการประมวลผลภาพสามารถแบ่งออกได้เป็น 2 ระดับคือการประมวลผลภาพใน ระดับต่ำ (Low-level image processing) และการประมวลผลภาพในระดับสูง (High-level image processing)

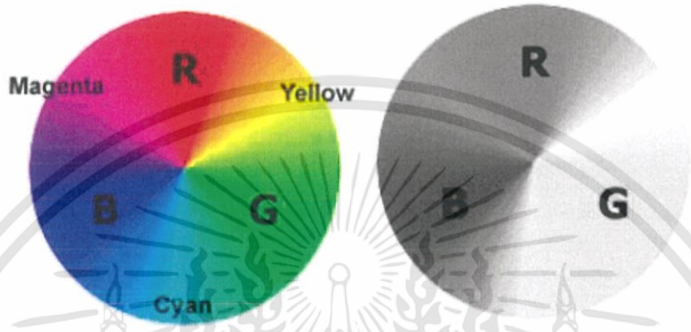
2.6 Gray Scale

สีในโทนสีเทา (Gray scale) จะมีค่าแม่สีของแสงสีแดง เขียว และน้ำเงินเป็นค่าเดียวกัน เช่น สีเทาอ่อน ซึ่งมีค่าแบบ 24 บิต (True color) เป็น 0xc3c3c3 จะเห็นว่าเกิดจากการรวมกันของแม่สีที่ค่า 0xc3 ซึ่งเป็นค่าสีเดียวกันทุกแม่สี ปัญหาคือ จะทำอย่างไรที่จะทำให้เกิดการรวมตัวกันของ แม่สีที่มีค่าต่าง ๆซึ่งไม่จำเป็นต้องเท่ากันให้กลายเป็นค่าที่เท่ากัน ซึ่งอาจจะทำได้โดยการนำค่าของแสงสีแต่ละสีมารวมกันแล้วหารสาม หรือการนำค่าของแสงสีแต่ละสีมาหาค่าเฉลี่ย ซึ่งหากทำโดยวิธีนี้อาจทำให้เกิดปัญหาได้คือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ถ้าค่าของแม่สีใดมีค่ามากว่าสีอื่นมาก ๆ จะทำให้ค่าเฉลี่ยของสีหนักไปทางค่านั้น ซึ่งจะทำให้สีเทาที่ได้ดูอ่อนเกินไป มีผลทำให้ภาพทั้งภาพมีโทนสีที่ผิดเพี้ยนไปได้

ดังนั้นเราต้องคำนึงถึงความสว่าง (Brightness) ของแม่สีซึ่งก็คือแม่สีแต่ละสีเมื่อวัด ค่าของแม่สีที่เท่ากันจะมีความสว่างไม่เท่ากัน



รูปที่ 2.1 แสดงการเปรียบเทียบความสว่างของแต่ละเฉดสี

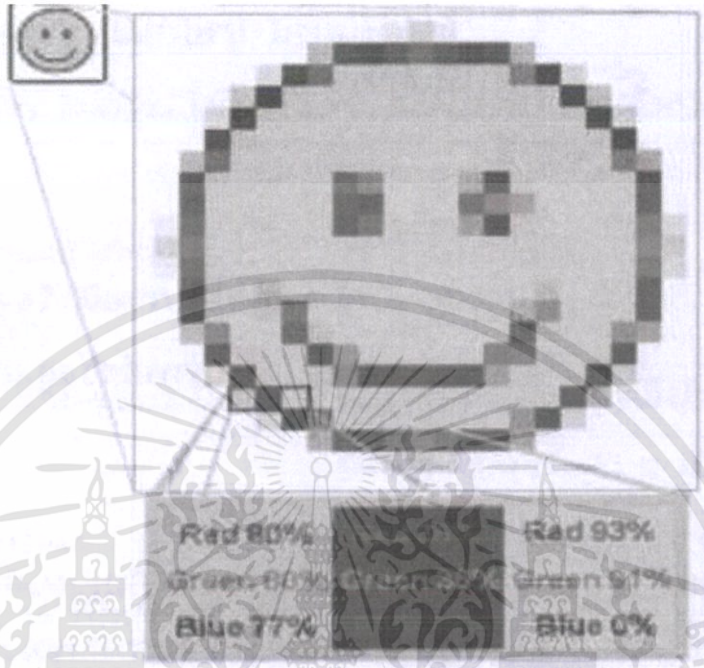
แสงสีแดงจะมีความสว่างน้อยกว่าแสงสีเขียว แต่มากกว่าแสงสีน้ำเงิน แสงสีเขียวมีความ สว่างมากที่สุด สำหรับแสงสีน้ำเงินมีความสว่างน้อยที่สุด เพื่อให้ภาพที่มีสีที่ใกล้เคียงกับต้นฉบับมากที่สุด การนำค่าของแม่สีแต่ละสีมารวมกันหรือผสมกันให้เป็นสีเทา เราจะต้องดึงค่ามาตามความ สว่างของแต่ละแม่สี เช่น สีเขียวก็จะดึงความสว่างมามากกว่าสีอื่น สีน้ำเงินจะดึงความสว่างมาน้อย กว่าสีอื่น เพราะสีน้ำเงินมีความสว่างน้อยที่สุดส่วนสีแดงจะดึงความสว่างมาพอประมาณ เป็นต้น แล้วจึงนำมารวมกันจะได้ค่าของสีเทาที่ต้องการ มาตรฐานของเปอร์เซ็นต์การดึงค่าสีตามความสว่าง ของแม่สีที่นิยมใช้กันนั้นจะใช้วิธีดึงค่าของสีแดง 29.9% สีเขียว 58.7% และสีน้ำเงิน 11.4%2.3 ทฤษฎีของ Gray scale

2.7 พิกเซล (Pixel)

โดยคำว่า Pixel มาจากการผสมคำระหว่าง Picture กับ Element ทำให้ได้เป็นความหมายว่าองค์ประกอบของภาพหรือหน่วยของภาพโดยความหนาแน่นของพิกเซลในภาพนั้นจะเป็นตัวบ่งบอกความละเอียดของภาพนั้น เช่น ภาพที่มีขนาด Resolution เท่ากับ 2272 x 1704 จะมีค่าเท่ากับ 3,871,488 พิกเซล หรือ ประมาณ 4MP โดย พิกเซลเป็นหน่วยที่เล็กที่สุด ของภาพ และแต่ละพิกเซลมีที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อยู่หรือตำแหน่งที่อ้างอิงจาก x , y หรือ กว้าง , ยาวตาม resolution ของภาพนั้น แต่ละพิกเซลจะเก็บค่าของสี และเมื่อพิกเซลแต่ละพิกเซลรวมตัวกันมากพอ มันจะทำให้เกิดเป็นภาพได้สังเกตุได้จากรูปที่ 2



รูปที่ 2.2 แสดงให้เห็นถึงพิกเซลในภาพ

บิตต่อพิกเซล เป็นการนิยามให้ในแต่ละพิกเซล แทนค่าสีได้ตามจำนวนของ เช่น จากสมการเปรียบเทียบในตารางด้านล่าง บ่งบอกถึงปริมาณสีที่ความสามารถแทนค่าได้ในแต่ละพิกเซล โดยยิ่งมีจำนวนบิตเยอะก็ยิ่งสามารถแทนค่าสีได้เยอะตามไปด้วย แล้วจะทำให้ภาพมีความสมจริงยิ่งขึ้น จะเห็นได้จากตารางที่ 3 ที่เปรียบเทียบค่าในขนาด bit per pixel ในขนาดที่ต่างกันออกไป 7 pixel

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2.1 ตารางเปรียบเทียบค่า bit per pixel

Bit per pixel	จำนวนสี
1 bpp	2^1=2 สี (ขาวดำ)
2 bpp	2^2=4 สี
3 bpp	2^3=8 สี
8 bpp	2^8=256 สี
16 bpp	2^{16}=65,536 สี
24 bpp	2^{24}=16,800,000 สี

ไม่ว่าขนาดของ resolution หรือ ค่าบิตในแต่ละพิกเซล ถ้ายิ่งมากเท่าไร ผลดีกับระบบ จำแนก ป้าย ทะเบียนรถอัตโนมัติ คือ จะทำให้วิเคราะห์ภาพได้ละเอียดมากยิ่งขึ้นแยกแยะช่องว่างกับตัวอักษรได้แม่นยำขึ้น ส่วนผลเสีย คือระยะเวลาในการประมวลก็จะช้าตามไปด้วย

2.8 Binary Image

รูปภาพขาว-ดำ คือ รูปภาพที่ประกอบด้วยสีเฉพาะสีขาวและสีดำ ใน Digital images ภาพจะอยู่ในรูปของแผ่นตาราง โดยแต่ละช่องจะเป็นส่วนหนึ่งของภาพหรืออักษร เรียกแต่ละจุดหรือช่อง นั้นว่า พิกเซลแต่ละพิกเซลจะถูกกำหนดให้มีระดับของความเข้ม (สีดำและสีขาว) ซึ่งแสดงให้อยู่ในรูปของรหัส Binary (0 และ 1) แต่ละพิกเซลก็จะแทนด้วย Binary digital (bits) จะถูกเก็บเป็นลำดับ ในคอมพิวเตอร์ โดยส่วนใหญ่แล้วคอมพิวเตอร์สามารถวิเคราะห์ข้อมูลจากรูปภาพขาวดำได้ง่ายกว่า รูปภาพแบบอื่น ๆ เนื่องจากรูปภาพขาวดำมีค่าข้อมูลของสีเพียง 2 ค่าซึ่งแทนด้วยค่า 0, 1 หรือ 0, 255 โดยที่ 0 แทนสีดำ และ 1 หรือ 255 แทนสีขาว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.9 Threshold Thresholding

เป็นเทคนิคที่ถูกใช้หลากหลายในหลายกระบวนการ Image processing ซึ่งเป็น กระบวนการแปลงภาพสีให้มีการแสดงผลได้แค่ 2 ระดับ คือ ขาว และดำซึ่งจะแปลงข้อมูลภาพให้ เป็นภาพ Binary (Binary image) โดยมีกระบวนการแปลงภาพที่มีความเข้มหลายระดับ (Multilevel image) ให้เป็นภาพที่มีความเข้มเพียง2ระดับหรือ1 บิต(bit) คือ0และ1โดยที่0แทนจุดที่มี ภาพสีดำ และ 1 แทนด้วยจุดที่มีภาพสีขาว โดยจะกำหนดให้พิกเซลที่มีค่าต่ำกว่าค่า Threshold เป็น ค่าๆหนึ่ง และมีค่าสูงกว่าเป็นค่าอีกค่าหนึ่งซึ่ง Image threshold ถือเป็น Segmentation อย่างหนึ่ง Thresholding technique คือการพิจารณาจุด pixel ในภาพว่าจุดใดควรจะเป็นจุดขาว หรือจุด ใดควรจะเป็นจุดที่มีค่าเท่ากับ 1 โดยจะทำการเปรียบเทียบค่าความเข้มของแสงของแต่ละพิกเซล (f(x,y)) กับค่าคงที่ที่เรียกว่า Threshold (Threshold value) เทคนิคนี้นิยมใช้กันมากในกรณีที่ความ แตกต่างระหว่างวัตถุ (Object) และพื้นหลัง (Background) ซึ่งเป็นลักษณะเดียวกับสีขาวของป้าย ทะเบียน และสีดำของตัวอักษรบนป้ายทะเบียนรถยนต์ ค่าพิกเซลในภาพที่มีค่าน้อยกว่าค่า Threshold จะถูกกำหนดเป็น 0 (จุดดำ) และถ้าค่าของพิกเซลใด ๆ ในภาพมีค่ามากกว่าหรือเท่ากับค่า Threshold จะถูกกำหนดให้เป็น 1 (จุดขาว)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.10 ระบบสี

2.10.1 ระบบสี RGB

จะประกอบไปด้วยแม่สี ทั้งหมด 3 สี ได้แก่ สีแดง สีเขียว และสีน้ำเงิน โดยการ ผสมสีในระบบนี้เป็นลักษณะของแสง ซึ่งแสงมีลักษณะเป็นคลื่นเมื่อใดที่แสงมาซ้อนทับ กันจะทำให้เกิดการรวมตัวของความยาวคลื่นดังนั้นจะทำให้เกิดแสงสีต่าง ๆ ใช้ในการ แสดงผล ซึ่งอุปกรณ์ที่ใช้ในกำรแสดงผลหรือ output ที่ได้ทำงจะแสดงบนอุปกรณ์ที่เป็น โทรทัศน์ คอมพิวเตอร์หรือ smart phone ส่วน input ของการรับค่าสีในระบบ RGB เช่น กล้องวิดิโอ หรือ webcam และ สแกนเนอร์ เป็นต้น ซึ่งระบบสี RGB จะใช้ในการวิเคราะห์ในระบบจำแนกป้ายทะเบียนรถยนต์อัตโนมัติโดยจะเห็นได้จากตาราง

ตารางที่ 2.2 ตารางรายชื่อสีและค่าที่แทนค่าสีต่าง ๆ

Color	Name	Hex Code #RRGGBB	Decimal Code (R,G,B)
	Black	#000000	(0,0,0)
	White	#FFFFFF	(255,255,255)
	Red	#FF0000	(255,0,0)
	Lime	#00FF00	(0,255,0)
	Blue	#0000FF	(0,0,255)
	Yellow	#FFFF00	(255,255,0)
	Cyan / Aqua	#00FFFF	(0,255,255)
	Magenta / Fuchsia	#FF00FF	(255,0,255)
	Silver	#C0C0C0	(192,192,192)
	Gray	#808080	(128,128,128)
	Maroon	#800000	(128,0,0)
	Olive	#808000	(128,128,0)
	Green	#008000	(0,128,0)
	Purple	#800080	(128,0,128)
	Teal	#008080	(0,128,128)
	Navy	#000080	(0,0,128)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จะสังเกตได้ว่าค่าสีขาวเป็นค่า 255 ในทุกแม่สี และสีดำจะมีค่า 0 ในทุกแม่สี ดังนั้นจะเป็นการง่ายในการวิเคราะห์องค์ประกอบที่สนใจในภาพได้ง่ายโดยองค์ประกอบที่สนใจนั้นก็คือ ป้ายทะเบียน และตัวอักษรภายในป้าย ซึ่งจะใช้คุณสมบัตินี้ในการแบ่งสีของภาพให้เป็น สีขาว และดำ

2.10.2 ระบบสีRYB

ระบบสีนี้ประกอบไปด้วยแม่สีทั้ง 3 ได้แก่ สีแดง สีเหลือง และสีน้ำเงิน เป็น มาตรฐานที่ใช้ในการผสมสีในงานศิลปะ เช่นการวาดรูป

2.10.3 ระบบสีCMYK

ระบบสีนี้คือการผสมสี 4 สี ได้แก่ สีน้ำเงินอมเขียว สีแดงอมม่วง สีเหลือง สีดำ ซึ่งใช้ในงานพิมพ์สี ซึ่งจะสังเกตุได้จาก หมึกพิมพ์ของเครื่องปริ้น ink jet ทั่วไปจะมี 4 สี

2.10.4 ระบบสีHSV และ HSL

เป็นระบบสีที่ปรับปรุงระบบสี RGB เพื่อให้ได้การแสดงผลสีของรูปทรงต่างๆได้ อย่างมีคุณภาพใช้งานด้าน คอมพิวเตอร์กราฟฟิก เพื่อความสมจริงของแสงและเงาจะสังเกตุได้ว่าค่าสีขาวเป็นค่า 255 ในทุกแม่สี และสีดำจะมีค่า 0 ในทุกแม่สี ดังนั้นจะเป็นการง่ายในการวิเคราะห์ องค์ประกอบที่เราสนใจในภาพได้ง่าย โดย องค์ประกอบที่เราสนใจ นั้นก็คือ ป้ายทะเบียน และตัวอักษรภายในป้าย ซึ่งเราจะใช้คุณสมบัตินี้ในการแบ่งสีของภาพให้เป็น สีขาว และดำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.11 Connected Component Labeling

ใช้ในการทำ Computer vision เพื่อตรวจหากลุ่มของพิกเซลที่เป็นส่วนเดียวกันในภาพ Binary Image ซึ่งหาโดยวิธีการหาพิกเซลที่ติดกัน ซึ่งเมื่อกำหนดกลุ่มของพิกเซลแต่ละกลุ่มแล้ว แต่ละกลุ่มจะถูกกำหนด Label ขึ้นมาเพื่อนำไปใช้งานต่อไป

หลักการทำงาน เริ่มจากการตรวจพิกเซลแต่ละพิกเซลของภาพ โดยเริ่มไล่จากพิกเซลแรกสุดของภาพ เมื่อพิกเซลที่เราสนใจคือส่วนที่เป็นสีดำในภาพ เพราะฉะนั้นจึงข้ามพิกเซลส่วนที่เป็นสีขาวไปและคำนวณเพียงแค่สีดำเท่านั้น

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

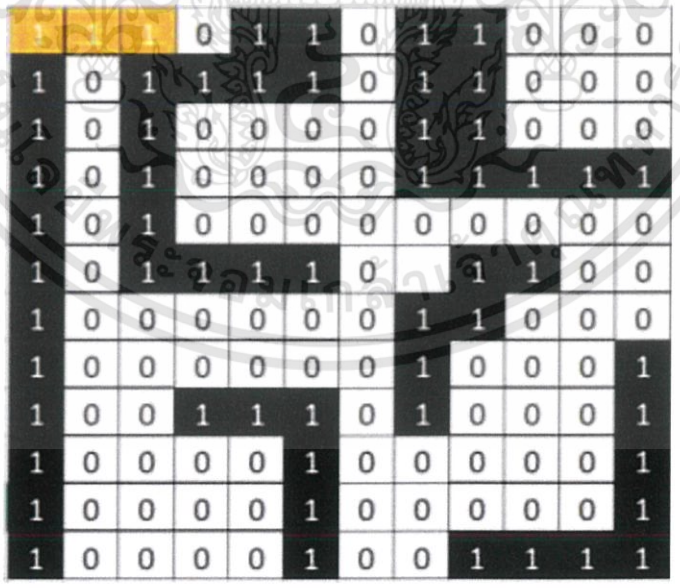
รูปที่ 2.3 ภาพ Binary ต้นฉบับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อเราเจอสีดำที่พิกเซลแรกกำหนด Label ของพิกเซลนั้นให้เป็นค่าค่าหนึ่ง สมมติว่าค่านั้นคือ 1

รูปที่ 2.4 เริ่มอัลกอริทึ่มจากพิกเซลแรก

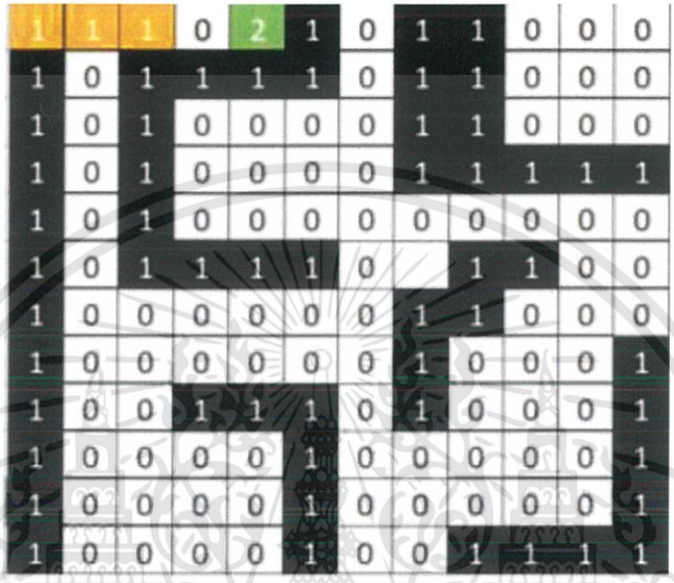
เมื่อขยับไปที่พิกเซลต่อไปจะทำการตรวจสอบพิกเซลรอบข้างว่ามีพิกเซลรอบข้างที่ถูกกำหนดเป็น Label ไว้ก่อนแล้วหรือไม่ ซึ่งในพิกเซลที่ 2 จากรูปจะเห็นว่าพิกเซลทางด้านซ้ายถูกกำหนดเป็น Label 1 ไปแล้ว ดังนั้น พิกเซลที่ 2 จึงถูกกำหนดเป็น Label 1 เหมือนกัน



รูปที่ 2.5 เริ่มอัลกอริทึ่มไป 3 พิกเซล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

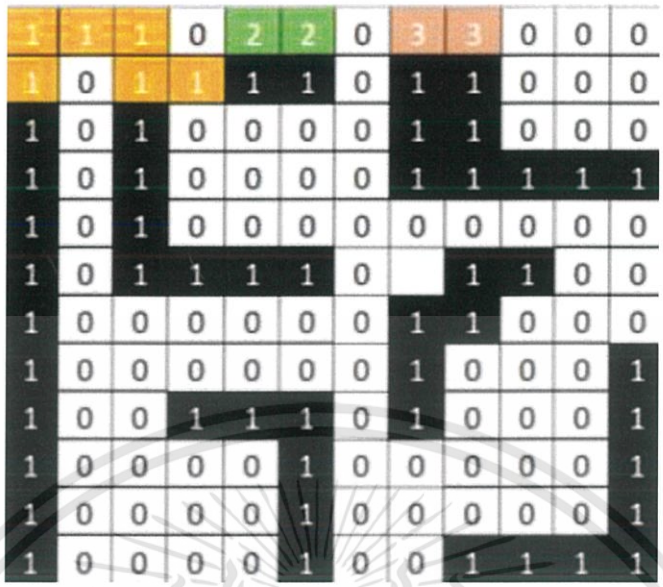
เมื่อมาถึงพิกเซลที่ 5 จะเห็นได้ว่าเป็นพิกเซลสีดำซึ่งเรารู้ว่าเป็นกลุ่มพิกเซลเดียวกับ Label 1 แต่ว่าเมื่อพิกเซลนี้ตรวจดูเงื่อนไขรอบข้างจะไม่พบพิกเซลรอบข้างที่ถูกกำหนด Label ไว้เลย ดังนั้นพิกเซลนี้จะถูกกำหนดเป็น Label 2 แทน



รูปที่ 2.6 รูปแสดงเมื่อพิกเซลที่เจอไม่ติดกับ Label ก่อนหน้า

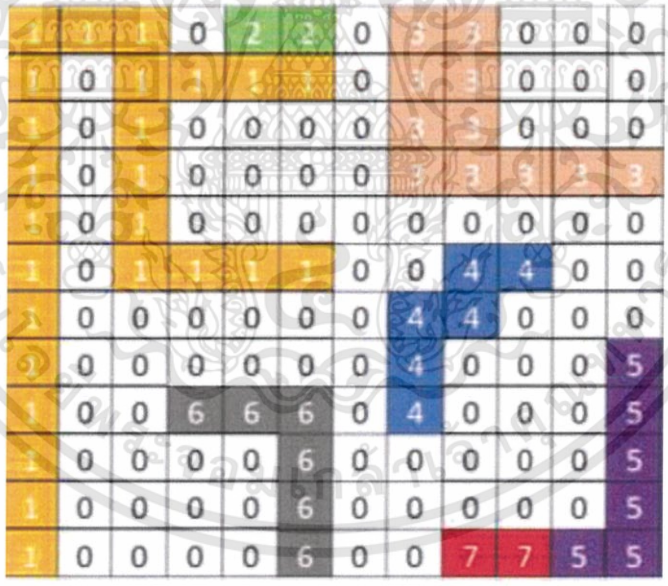
จุดสำคัญอยู่ที่พิกเซลที่ 4 ในแถวที่ 2 เพราะจุดนี้เมื่อตรวจดูเงื่อนไขรอบข้างนั้น จะเจอทั้ง Label 1 และ Label 2 ซึ่งในกรณีแบบนี้จะถูกกำหนดให้พิกเซลนั้นเป็นพิกเซลที่มี Label ที่มีค่าน้อยกว่าก็คือ Label 1 และในขณะเดียวกัน Label 2 ก็จะถูกกำหนดเก็บค่าไว้ว่าเป็นกลุ่มพิกเซลเดียวกันกับ Label 1 ด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.7 รูปแสดงเมื่อเจอพิกเซลที่รอบข้างมี 2 Label

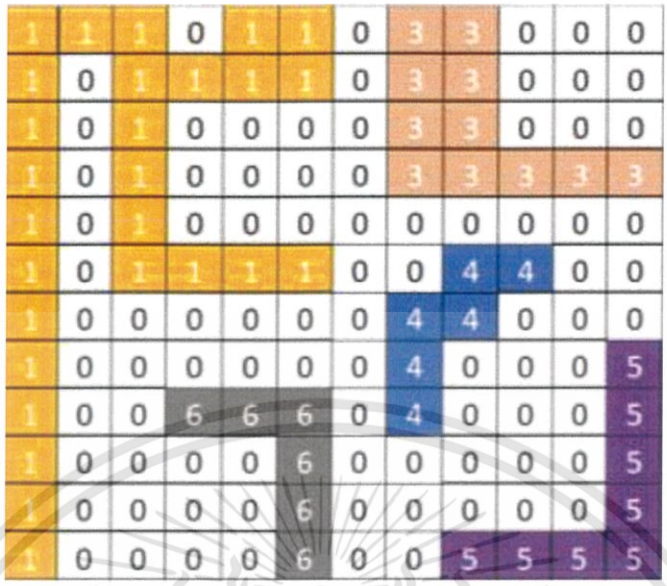
เมื่อทำครบทั้งภาพจะได้ Label ออกมาเป็นเหมือนกับภาพดังนี้ซึ่งยังไม่เสร็จสมบูรณ์



รูปที่ 2.8 เมื่อตรวจสอบครบทั้งภาพ

ขั้นตอนสุดท้ายคือการจัดการกับกลุ่มพิกเซลที่มีหลาย Label ซ้อนกันอยู่ซึ่งต้องแก้ให้กลุ่มพิกเซลที่ Label มากกว่าเปลี่ยนเป็น Label ที่น้อยกว่า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



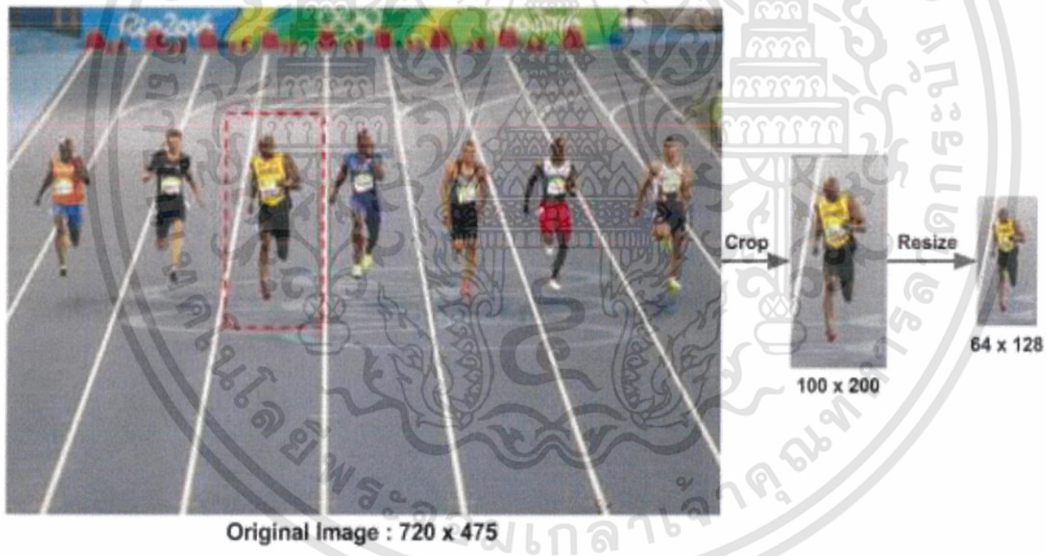
รูปที่ 2.9 รูปเมื่ออัลกอริทึมทำงานสมบูรณ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.11 Histogram of Orientation Gradient (HOG)

เป็นตัวแยกคุณลักษณะ (Feature Extraction) ตัวหนึ่ง มีหน้าที่ในการแปลงรูปภาพที่เรามองในลักษณะของภาพ RGB ให้เป็น Blocks of Edges เพื่อให้คอมพิวเตอร์สามารถมองรูปภาพนี้ได้ โดย Gradient ของ HOG หมายถึงการดึงขอบของวัตถุในภาพออกมา และ Histogram คือการนับสถิติว่าในบริเวณที่สนใจมีขอบในทิศทางไหนมากกว่ากัน ซึ่ง HOG นั้นจะแปลงจากขนาดของภาพ กว้าง x ยาว x 3 (channels) ให้อยู่ในรูปของเวกเตอร์คุณลักษณะ (Feature vector) ซึ่งจะต้องมีการปรับค่าให้ได้เวกเตอร์คุณลักษณะที่เหมาะสม โดยมีหลักการคำนวณดังนี้

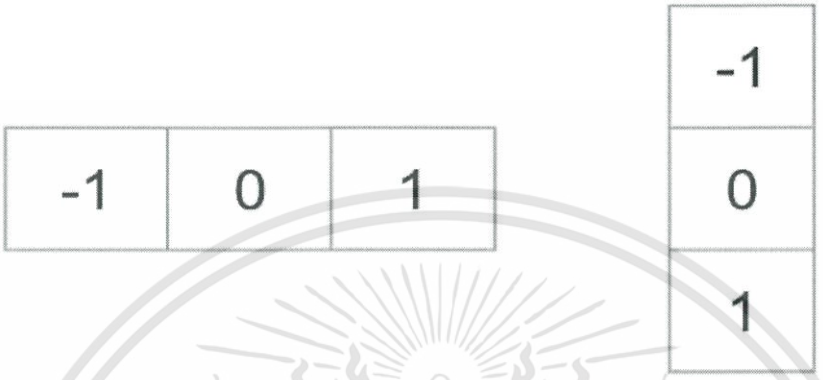
สมมติว่าภาพที่เราสนใจคือภาพนักวิ่งคนหนึ่ง สิ่งแรกที่ต้องทำคือการปรับขนาดของภาพให้มีอัตราส่วนของภาพเป็น 1: 2 ซึ่งเราจะปรับขนาดภาพให้เป็น 64 x 128 โดยเดิมภาพแก้วน้ำนี้อาจจะมีขนาดอยู่ที่ 720 x 475 ให้ทำการตัดภาพและปรับขนาดภาพตามที่ต้องการ



รูปที่ 2.10 Preprocessing ภาพต้นฉบับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อเราได้ภาพที่พร้อมจะทำ HOG แล้วก็จะเริ่มคำนวณหาเวกเตอร์คุณลักษณะต่อไปโดยเราต้องหาค่า Gradient ของภาพในแต่ละพิกเซลทั้งในแนวตั้งและแนวนอนโดยการใช้ Sobel operator ในไลบรารี่ OpenCV



รูปที่ 2.11 Sobel Operator

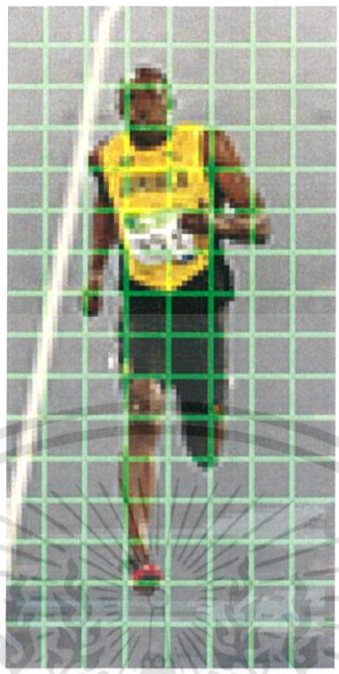
เมื่อทำทุกพิกเซลจะทำให้เราได้ค่า Gradient ของแต่ละพิกเซลในแนวตั้ง (g_y) และแนวนอน (g_x) แล้วเราจะสามารถหาขนาด (Magnitude) และทิศทาง (Direction) ของ Gradient ได้จากสูตร

$$g = \sqrt{g_x^2 + g_y^2} \tag{2.1}$$

$$\theta = \arctan\frac{g_y}{g_x} \tag{2.2}$$

หลังจากนั้นเราจะแบ่งภาพให้อยู่ในหน่วยเซลล์แทนเพื่อแบ่งส่วนของภาพหาผลรวมของขนาดและทิศทางของ Gradient ซึ่งเราจะแบ่งภาพของคนวิ่งนี้ออกเป็น 8 x 16 cells

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



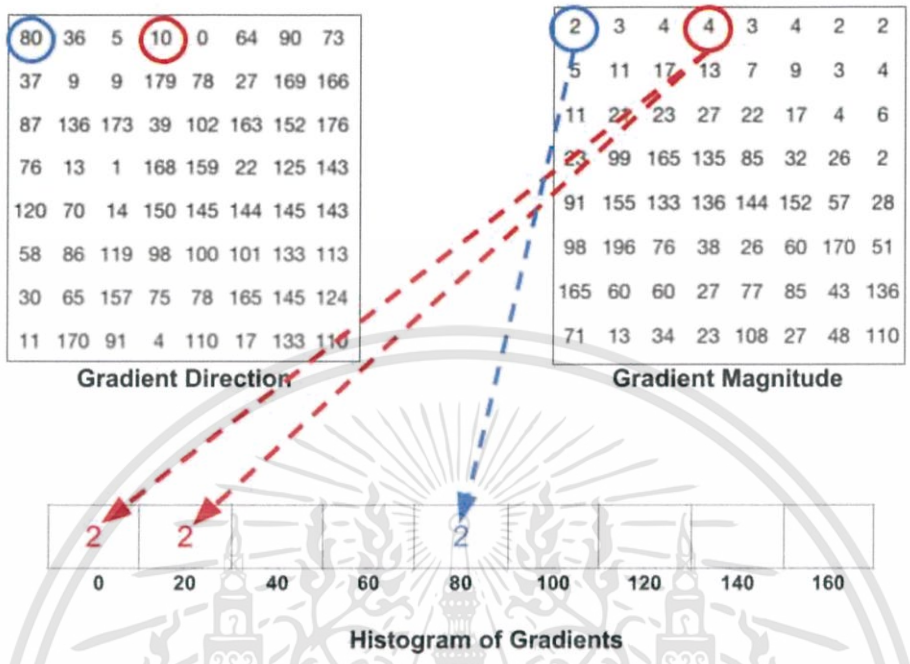
รูปที่ 2.12 แบ่งภาพเป็นเซลล์ 8 x 16

เมื่อแบ่งเซลล์เป็น 8 x 16 เซลล์นั้นทำให้ใน 1 เซลล์มีพิกเซลอยู่ 8 x 8 พิกเซลหรือมีทั้งหมด 64 พิกเซลซึ่งแต่ละพิกเซลมีค่าของ Gradient อยู่ 2 ค่าคือขนาดและทิศทาง หมายความว่าเราจะมีค่าของ Gradient อยู่ทั้งหมด 128 ค่าต่อหนึ่งเซลล์ซึ่งค่านี้จะไปถูกแบ่งแสดงผลบน 9-bin histogram ซึ่งจะถูกเก็บเป็นเวกเตอร์ 9 ค่า

โดย 9-bin histogram จะเป็นเวกเตอร์ที่แสดงถึงตัวเลข 9 ตัวที่สัมพันธ์กับค่ามุม 0,20,40,60 ... 160 องศา ตามตัวอย่างดังนี้

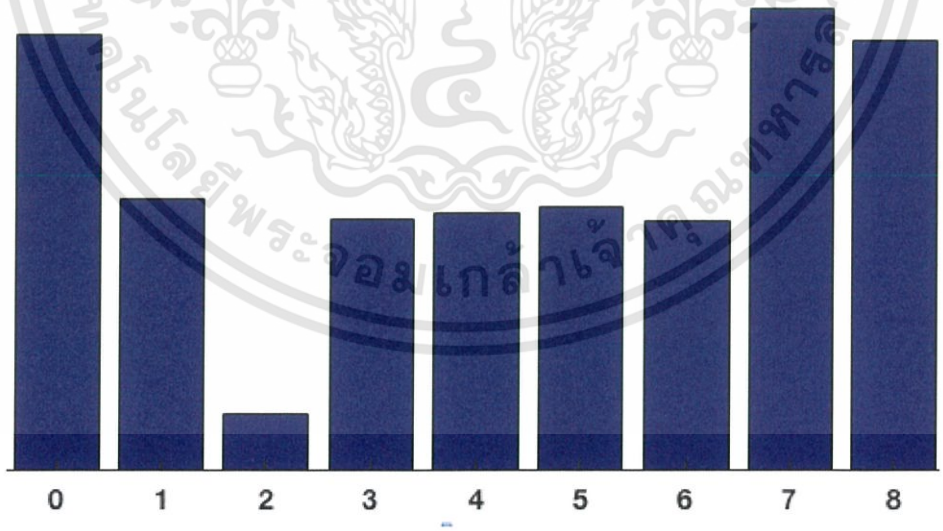
รูปนี้แสดงให้เห็นถึงค่าของขนาดและทิศทางของ Gradient ในแต่ละพิกเซลในเซลล์ ๆ หนึ่งซึ่งค่านี้จะถูกแทนลงใน Histogram โดยเมื่อดูที่พิกเซลแรก มีขนาดเท่ากับ 2 และมุม 80 องศา ดังนั้นจึงนำค่า 2 ไปเพิ่มใน histogram ที่ 80 องศา หรืออีกค่าหนึ่งคือพิกเซลที่ 4 มีขนาดเท่ากับ 4 และมีมุม 10 องศา ซึ่งอยู่กึ่งกลางระหว่าง 0 กับ 20 องศาดังนั้นจึงต้องมีการแบ่งไปตามอัตราส่วนซึ่งจะได้ว่าแบ่งไปอยู่ที่ 0 องศาและ 20 องศาอย่างละ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.13 คำนวณ Gradient ในเซลล์

ทำอย่างนี้ไปเรื่อย ๆ จนครบทุกพิกเซลใน 1 เซลล์ท้ายที่สุดจะทำให้เราได้ค่า Histogram ของเซลล์นี้มา



รูปที่ 2.14 histogram ของเซลล์ ๆ หนึ่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนต่อมาคือการทำ Normalization เพื่อจัดการกับแสงของภาพที่อาจจะส่งผลต่อสีภาพที่เรานำมาคำนวณ โดยสมมติว่าที่พิกเซลหนึ่งมีค่าสี RGB อยู่ที่ [128, 64, 32] เราสามารถ Normalization ได้โดยการคำนวณดังนี้

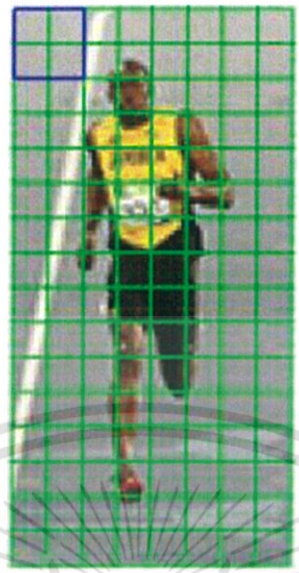
$$\text{normalized vector} = \frac{\sqrt{128^2+64^2+32^2}}{[128,\quad 64,\quad 32]} \tag{2.3}$$

ซึ่งจะได้เวกเตอร์ที่ normalized แล้วเป็นค่า [0.87, 0.43, 0.22] ต่อมาเมื่อพิกเซลเดียวกันนี้มีค่าแสงที่เพิ่มขึ้นเป็น 2 เท่าทำให้ค่า RGB กลายเป็น [256, 128, 64] เมื่อเราลองทำการ normalization อีกครั้งกับพิกเซลนี้

$$\text{normalized vector} = \frac{\sqrt{256^2+128^2+64^2}}{[256,\quad 128,\quad 64]} \tag{2.4}$$

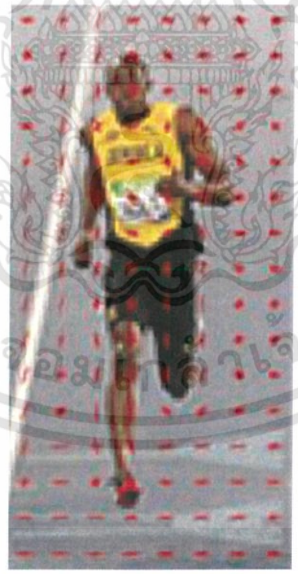
จะได้ค่าเวกเตอร์ [0.87, 0.87, 0.22] ซึ่งมีค่าเท่ากัน แสดงให้เห็นว่าการ normalization จะช่วยจัดการปัญหาเรื่องแสงให้ได้ซึ่งเราจะกำหนดให้บล็อค ๆ หนึ่งประกอบไปด้วยเซลล์ 2 x 2 เซลล์ซึ่งจะทำให้เรามีค่า 9-bin orientation อยู่ทั้งหมด 4 เวกเตอร์ทำให้เรามีค่าเวกเตอร์แต่ละบล็อกเป็น 36 x 1 หรือ 36 ค่าและเราก็ทำการ normalization เวกเตอร์ 36 x 1 นี้เหมือนกับที่เราทำกับเวกเตอร์ 3 x 1 ข้างบน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.15 กำหนดบล็อคที่ประกอบไปด้วยเซลล์ 2 x 2

ท้ายที่สุดคือเมื่อเราแบ่งบล็อกเป็น 2 x 2 เซลล์ต่อ 1 บล็อกจะทำให้เรามีบล็อกอยู่ทั้งหมด 7 x 15 = 105 บล็อก โดยแต่ละบล็อกมีเวกเตอร์อยู่ทั้งหมด 36 ค่า ทำให้เรามีเวกเตอร์คุณลักษณะทั้งหมด 105 x 36 = 3780 ค่า ซึ่งได้เป็นเวกเตอร์คุณลักษณะ 3780 x 1



รูปที่ 2.16 ภาพเมื่อทำ HOG สมบูรณ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

ขั้นตอนการดำเนินงาน

จากบทที่ 2 จึงได้นำความรู้ที่ศึกษาค้นคว้ามาประยุกต์ใช้กับโครงงานและเริ่มวางแผนการทำงาน โดยขั้นตอนการทำงาน มีดังต่อไปนี้

3.1 ศึกษาและออกแบบ Software ที่ใช้ในการทำโปรแกรมอ่านป้ายทะเบียน

3.2 ศึกษาและเลือกกล้องที่เหมาะสมกับการนำมาทำ Software ประมวลผลภาพ

3.3 ทำ Software เพื่อตรวจจับป้ายทะเบียนจากภาพที่ได้จากกล้อง

3.4 ทำ Software เพื่อแยกตัวอักษรที่ติดกัน

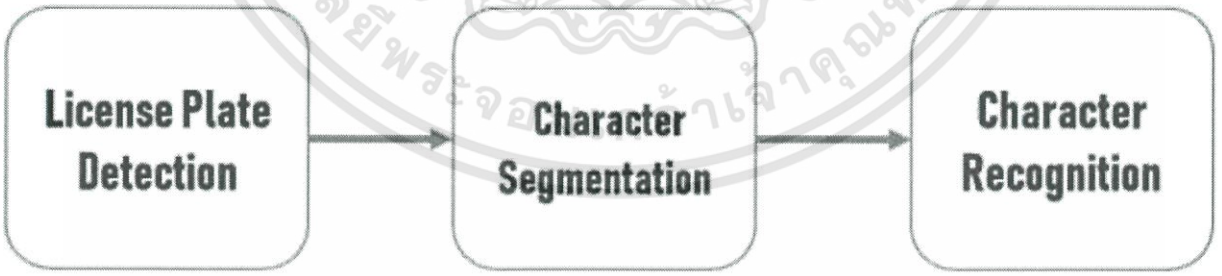
3.5 ทำ Software เพื่อใช้ในการรู้จำตัวอักษร

3.6 ทำหน้าต่าง User Interface สำหรับใช้งาน

3.7 ติดตั้งกล้องและทดสอบโปรแกรม

3.1 ศึกษาและออกแบบ Software ที่ใช้ในการทำโปรแกรมอ่านป้ายทะเบียน

จากการศึกษาและออกแบบโปรแกรมนั้น ทำให้ทราบว่าโครงสร้างของ Software สำหรับอ่านป้ายทะเบียนนั้น แบ่งออกเป็น 3 ส่วนหลัก ๆ จากไดอะแกรมรูปที่ 3.1



รูปที่ 3.1 ลำดับการทำงานของ Software

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2.1 License Plate Detection

คือ Software ส่วนแรกสุดเพื่อใช้ในการตรวจจับป้ายทะเบียนจากภาพที่ได้จากกล้อง ซึ่ง Input ของส่วนนี้ก็คือภาพในเฟรม ๆ หนึ่งซึ่งได้จากวิดีโอของกล้อง และ output ที่ได้คือภาพที่ถูกตัดออกมาเฉพาะแค่ส่วนของป้ายทะเบียนจากภาพรวมทั้งหมด ซึ่งสาเหตุของการเขียน Software ส่วนนี้ขึ้นมานั้นเนื่องจากถ้าเราไม่ตัดภาพเฉพาะส่วนป้ายทะเบียนมาประมวลผลแต่ประมวลผลจากภาพทั้งหมดในทุก ๆ pixel นั้นจะสิ้นเปลืองเวลาและทรัพยากรในการประมวลผลแต่ละเฟรมมากโดยที่ไม่จำเป็น การประมวลผลภาพที่เหลือแค่ส่วนป้ายทะเบียนจะลดเวลาประมวลผลไปอย่างมาก

3.2.2 Character Segmentation

คือ Software เพื่อแยกตัวอักษรที่ติดกันซึ่งเป็นพื้นฐานสำคัญในการทำการรู้จำตัวอักษรที่ใช้กันอย่างแพร่หลาย ประสิทธิภาพในการแยกตัวอักษรที่ติดกันนั้นส่งผลอย่างมากกับประสิทธิภาพและความแม่นยำในการรู้จำตัวอักษร Input ของส่วนนี้คือภาพป้ายทะเบียนที่ถูกตัดออกมาในขั้นตอนที่แล้ว และ output ที่ได้คือภาพของตัวอักษรแต่ละตัวบนป้ายทะเบียนที่ถูกแยกออกจากกันเป็นภาพอักษรเดี่ยวเรียบร้อยแล้ว

3.2.3 Character Recognition

คือส่วนที่สำคัญที่สุดของโครงงานนี้ การรู้จำตัวอักษรคือการเปลี่ยนภาพที่มีตัวอักษรอยู่นั้นให้ออกมาในรูปแบบตัวแปร String หรือเปลี่ยนให้เป็น text file ซึ่งขั้นตอนนี้จะทำให้เราได้รู้ว่ารถที่เป็น Input เข้ามาในขั้นตอนแรกนั้นมีป้ายทะเบียนอะไร โดย Input ของส่วนนี้ก็คือรูปตัวอักษรเดี่ยวที่ได้มาจากขั้นตอนที่แล้ว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 ศึกษาและเลือกกล้องที่เหมาะสมกับการนำมาทำ Software ประมวลผลภาพ

กล้องเป็นอุปกรณ์ที่มีความสำคัญต่อโครงงานนี้อย่างมากที่สุด เพราะทุกส่วนของ Software นั้นเป็นการใช้หลักการประมวลผลภาพทั้งหมด เพราะฉะนั้นกล้องที่เลือกมาจะต้องมีความคมชัด ระบุตัวอักษรในภาพได้ชัดเจนแม้จะต้องเจอกับสภาพแวดล้อมที่เปลี่ยนไป เช่น ฟ้าครึ้ม ฝนตก แดดแรง แสงสะท้อน เป็นต้น ดังนั้นจึงต้องมีการกำหนดคุณสมบัติของกล้องที่จะนำมาทดสอบตามหัวข้อดังนี้

3.3.1 Sensor

ในส่วนของ Sensor นั้นมีหลายส่วนที่ต้องพิจารณา ซึ่งจะแบ่งออกเป็นดังนี้

3.3.1.1 ขนาดของ Sensor ควรจะมีขนาด 1/3” หรือมากกว่า เนื่องจากยิ่งขนาดของพิกเซลมีขนาดใหญ่มากเท่าไร ก็จะส่งผลให้สามารถทำระบบการรู้จำได้ดีขึ้นในสภาพที่แสงไม่เป็นไปตามต้องการ

3.3.1.2 Minimum illumination ควรมีค่าอย่างน้อย 0.1-0.01 lux เมื่อใช้กับ focal length ที่ f/1.0,f/1.2,f/1.3 หรือ f/1.4

3.3.1.3 ขนาดของพิกเซล (Pixel size) ควรมีขนาด 3.75 ไมโครเมตร หรือสูงกว่า เนื่องจากยิ่งขนาดของของพิกเซลยิ่งสูงเท่าไร ค่าของ sensitivity ยิ่งมากเท่านั้น

3.3.1.4 Dynamic Range แนะนำให้ใช้กล้องที่มีโหมด Wide dynamic range (WDR) ซึ่งจะช่วยให้ภาพที่ได้จากกล้องนั้น ถูกปรับค่าแสงให้สามารถประมวลผลได้ง่ายขึ้นเมื่อนำเข้าสู่ Software

3.3.2 Resolution

ควรจะเป็น Full HD 1920 x 1080 หรือจำนวนพิกเซล 2 MP ขึ้นไป เพื่อความคมชัดในการประมวลผล เนื่องจากภาพที่จะนำมาประมวลผลจะถูกตัดและย่อลงไปอีก ภาพต้นฉบับควรจะละเอียดพอสมควร

3.3.3 Angle of View

ควรจะมีความกว้างของ Angle of View ที่มากกว่า 80 องศาเพื่อให้สามารถเก็บภาพได้กว้าง แต่น้อยกว่า 100 องศา สาเหตุเนื่องจากว่าเมื่อมี Angle of View ที่สูงมาก หรือเป็นกล้องชนิดเลนส์ Wide จะทำให้ภาพที่ได้นั้นบิดเบี้ยว โค้ง ทำให้นำไปประมวลผลต่อได้ยาก เนื่องจากตัวอักษรก็จะบิดเบี้ยวไปด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.4 Aperture

รูรับแสงอาจจะไม่ค่อยส่งผลกับภาพในการเอามาประมวลผลต่อ แต่แนะนำให้เลือกเลนส์ที่มีความกว้างมาก เพื่อทำให้สามารถถ่ายได้ดีในกรณีแสงน้อย คือ f/1.0 - f/1.8

3.3.5 Focal Length

ทางยาวโฟกัส ส่งผลต่อค่าของ Angle of View ดังนั้นจึงมีความสำคัญ เพื่อที่จะให้ได้ทางยาวโฟกัสที่เหมาะสม ควรจะมีการกำหนดระยะความกว้างและระยะห่างจากกล้องถึงวัตถุที่จะถ่ายซึ่งในโครงงานนี้ก็คือรถที่เข้ามาในบริษัท เมื่อกำหนดได้แล้วจึงนำมาคำนวณหาค่าทางยาวโฟกัส ซึ่งเราได้กำหนดค่าความกว้างของกล้องที่ควรจะถ่ายอยู่ที่ ..... เมตร และระยะห่างจากกล้องถึงรถอยู่ที่ 5 เมตร ทางยาวโฟกัสกล้องที่เหมาะสมจึงควรจะเป็น 2.8 – 3.6 mm

3.3.6 Framerate

เนื่องจากว่ากล้องที่สามารถถ่ายวิดีโอได้ 15 fps ขึ้นไปก็สามารถนำมาประมวลผลได้ดีแล้วสำหรับโครงงานนี้ และโดยทั่วไปกล้องก็มี Framerate อยู่ที่ 25-30 fps เป็นอย่างต่ำอยู่แล้ว ทำให้ส่วนนี้ไม่มีปัญหามากนัก

3.3.7 S/N ratio

ยิ่งมากเท่าไร noise ของภาพก็จะยิ่งน้อยเท่านั้น แต่ไม่ใช่ตัวเลือกสำคัญที่จะต้องนำมาพิจารณาเป็นหลัก

หลังจากที่กำหนดคุณสมบัติของกล้องมาเรียบร้อยแล้วจึงดำเนินการเลือกกล้องที่จะนำมาทดสอบเป็นอย่างแรกในช่วงหลังจากที่ Software นั้นเขียนเสร็จ กล้องที่ได้มาทดสอบก็คือ Hikvision 4 MP IR Fixed Bullet Network Camera ซึ่งเป็นกล้อง CCTV ที่มีคุณสมบัติดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Sensor: 1/3” Progressive Scan CMOS

Resolution: 2688 x 1520

Framerate: 30 fps

Focal Length: 2.8 fixed lens

Minimum illumination: 0.01 Lux

Dynamic range: 120 dB WDR

Aperture: F1.6

Angle of View: Horizontal 103 degree

S/N ratio: 50 dB

รูปที่ 3.2 Hikvision 4 MP IR Fixed Bullet Network Camera

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3 ทำ Software เพื่อตรวจจับป้ายทะเบียนจากภาพที่ได้จากกล้อง

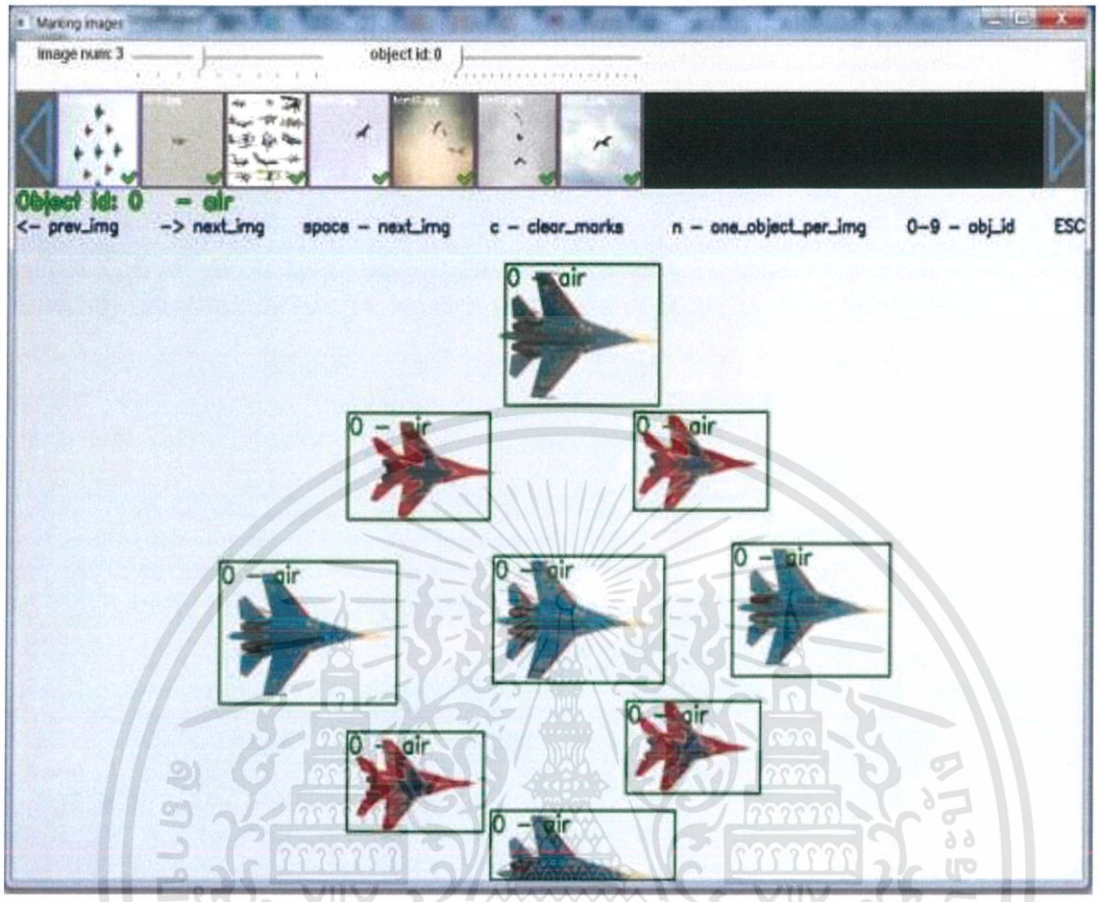
เป็นส่วนแรกของ Software เนื่องจากสิ่งข้อมูลที่เราต้องการมากที่สุดคือเลขที่ทะเบียนรถยนต์ ดังนั้นจึงต้องหาป้ายทะเบียนที่เข้ามาในเฟรมของกล้องให้ได้ก่อน เพื่อที่จะได้ตัดเฉพาะส่วนป้ายทะเบียนออกมาประมวลผลในขั้นตอนต่อไป ซึ่งอัลกอริทึมที่เราเลือกใช้คือ You Only Look Once (YOLO) ซึ่งเป็น Deep learning ที่มีผู้พัฒนา pjreddie พัฒนาจาก Framework ที่ชื่อว่า Darknet ขึ้นมาเป็นระบบ Object Detection ที่ทำหน้าที่ตรวจหาวัตถุที่เราสนใจในภาพหรือวิดีโอที่เป็น input ซึ่งวัตถุที่เราสนใจในโครงงานนี้ก็คือป้ายทะเบียน ลักษณะการทำงานโดยทั่วไปของ Deep learning นั้นต้องมีการสอนคอมพิวเตอร์เพื่อให้คอมพิวเตอร์รู้ว่า ลักษณะของป้ายทะเบียนเป็นอย่างไร ดังนั้นเราจึงต้องมีการเตรียมชุดข้อมูลของป้ายทะเบียนเพื่อสอนคอมพิวเตอร์ก่อน ซึ่งเราได้เตรียมรูปภาพสำหรับการสอน Deep learning ไว้ทั้งหมด 789 รูป รูปที่เตรียมจะเป็นรูปป้ายทะเบียนเต็มภาพหรือเป็นภาพของรถที่มองเห็นป้ายทะเบียนก็ได้ โดยการจะใช้ YOLO นั้นควรจะต้องมีไลบรารี่ของ Nvidia ที่ชื่อว่า CUDA และ CUDnn ทั้ง 2 ไลบรารี่เนื่องจากว่าถ้าเกิดไม่ลงไลบรารี่ 2 ตัวนี้ YOLO จะประมวลผลบน CPU ของคอมพิวเตอร์ซึ่งแทบจะไม่สามารถใช้งานได้เลยเนื่องจาก YOLO ใช้ทรัพยากรสูงมากในการประมวลผลภาพ การลงไลบรารี่นี้จะทำให้ YOLO ประมวลผลบน GPU ทำให้การประมวลผลเร็วขึ้นซึ่งการ์ดจอนั้นควรจะมีประสิทธิภาพที่สูงพอสมควร ซึ่งการ์ดจอที่เราใช้คือ GTX 1050 Ti ทำให้สามารถประมวลผลได้ค่อนข้างเร็วและได้เฟรมเรทในระดับที่พอใช้งานได้ และต่อมาคือ YOLO ถูกพัฒนาขึ้นมาด้วยภาษา C ดังนั้นจึงต้องใช้โปรแกรม Visual Studio 2015 (ต้องเป็นเวอร์ชั่น 2015 เท่านั้นเนื่องจากไลบรารี่ CUDA และ CUDnn ไม่รองรับเวอร์ชั่นอื่น) ในการเขียนโปรแกรมพัฒนา YOLO ต่อให้เข้ากับโครงงานของเรา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 3.3 Dataset ป้ายทะเบียน

เมื่อเตรียมรูปเพื่อสอน Deep learning เรียบร้อยแล้ว เราจะต้องทำการกำหนด Label ให้กับแต่ละรูปที่เลือกมา เพื่อบอกว่าส่วนไหนของรูปที่เป็นป้ายทะเบียน ซึ่ง YOLO จะนำส่วนที่เราทำ Label ไว้ไปเรียนรู้ ซึ่งเราสามารถใช้โปรแกรม YOLO Mark ที่ AlexeyAB ได้พัฒนาขึ้นมาเพื่อสร้าง Label สำหรับใช้กับ YOLO โดยเฉพาะขึ้นมาไว้แล้ว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.4 ตัวอย่างโปรแกรมจาก Github ของ AlexeyAB

ลักษณะการใช้โปรแกรมคือการสร้างเส้นกรอบล้อมรอบ (Bounding Box) วัตถุที่เราต้องการจะ Label ในภาพสิ่งที่ต้องทำในโปรแกรมนี้คือการตั้งค่า Class ที่ต้องการในไฟล์ obj.name ซึ่งเราได้ตั้ง Class ไว้ว่า “License Plate” และกำหนดจำนวน Class ที่ต้องการทำ Label ในไฟล์ obj.data ซึ่งในที่นี้ก็คือมีแค่ 1 Class หลังจากนั้นก็เริ่มทำการสร้าง Bounding Box กับรูปป้ายทะเบียนทั้งหมด 789 รูป หลังจากสร้าง Bounding Box เสร็จแล้วออกจากโปรแกรม YOLO Mark จะทำให้เราได้ไฟล์ text มาโดยแต่ละรูปก็จะมีไฟล์ text ของตัวเอง ซึ่งในไฟล์นั้นจะบอกถึงตำแหน่งของวัตถุที่เราสนใจภายในภาพ ซึ่งเราจะต้องใช้ไฟล์ที่บอกตำแหน่งนี้คู่กับรูปป้ายทะเบียนในการนำไปสอน YOLO ในขั้นตอนถัดไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อเราได้ทำ Label กับรูปทุกรูปที่จะนำมาใช้สอน YOLO แล้วนั้น ขั้นตอนต่อไปคือการตั้งค่าตัวแปรต่างๆสำหรับการนำไปสอนให้เหมาะสมที่สุด ขั้นตอนแรกคือการสร้างไฟล์เพื่อตั้งค่า ไฟล์แรกที่ต้องสร้างขึ้นมาคือไฟล์สกุล cfg ซึ่งเป็นไฟล์สำหรับตั้งค่าโมเดลโครงสร้าง Convolution Neural Network

```
[net]
# Testing
batch=64
subdivisions=32
# Training
# batch=64
# subdivisions=16
width=512
height=512
channels=3
momentum=0.9
decay=0.0005
angle=0
saturation = 1.5
exposure = 1.5
hue=.1

learning_rate=0.0001
burn_in=1000
max_batches = 500200
policy=steps
steps=400000,450000
scales=.1,.1

[convolutional]
batch_normalize=1
filters=32
size=3
stride=1
pad=1
activation=leaky

# Downsample

[convolutional]
batch_normalize=1
filters=64
size=3
stride=2
```

รูปที่ 3.5 การตั้งค่าไฟล์ cfg

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ซึ่งเราสามารถตั้งค่าให้เหมาะสมได้ตามคู่มือการใช้ YOLO ใน github ของ AlexeyAB โดยไฟล์นี้เราทำการ Copy ไฟล์ของ yolo-voc.cfg ซึ่งเป็นโมเดลมาตรฐานทั่วไปของการใช้ YOLO ที่ถูกตั้งค่ามาเหมาะสมในระดับหนึ่งแล้ว และเราจะนำมาปรับค่าเพื่อให้เข้ากับโครงงานนี้ โดยค่าแรกที่ต้องปรับค่า Class ที่ถูกตั้งเป็น 20 ไว้ ให้เปลี่ยนเป็นค่าของ Class ที่เราใช้ นั่นคือ 1 เพราะมี Class แค่ป้ายทะเบียนเพียงอย่างเดียว ส่วนค่าต่อไปคือ filters ที่บรรทัด 224 ซึ่งถูกตั้งค่าไว้ที่ 125 ให้เปลี่ยนเป็นค่าที่คำนวณจาก (Class + 5)*5 ทำให้ค่า filters ในโครงงานนี้เปลี่ยนเป็น 30 เท่านี้ก็เป็นอันเสร็จสำหรับไฟล์ cfg

ไฟล์ต่อมาที่จำเป็นต้องมีคือไฟล์ที่บอกถึงที่อยู่ของรูปที่จะใช้ในการสอน และไฟล์ที่บอกถึงที่อยู่ของรูปที่ใช้ในการทดสอบผลลัพธ์การสอน ซึ่งก็จะสร้างเป็นไฟล์ txt ขึ้นมา เนื้อหาข้างในคือที่อยู่ของไฟล์รูปแต่ละรูป

2 ไฟล์สุดท้ายที่ต้องสร้างขึ้นมาก็คือไฟล์ obj.names ซึ่งในไฟล์จะมีแค่ชื่อของ Class ที่เราต้องการสร้าง ซึ่งในที่นี้ก็จะมีแค่ Class ชื่อ License Plate เพียงอันเดียวเท่านั้น และไฟล์ obj.data ซึ่งในไฟล์จะประกอบไปด้วยจำนวน Class ที่ต้องการสอน ชื่อไฟล์ที่บอกถึงที่อยู่ของรูปที่ใช้ในการสอน ชื่อไฟล์ที่บอกถึงที่อยู่ของรูปที่ใช้ในการทดสอบผลลัพธ์ ตำแหน่งไฟล์ obj.names และโฟลเดอร์ที่ใช้เก็บไฟล์ weights ของ YOLO

ต่อมาก่อนที่จะเริ่มการสอน จะต้องดาวน์โหลดไฟล์ Pre-trained ซึ่งเป็นไฟล์ค่า weights เริ่มต้นสำหรับการสอนโดยทั่วไปที่ pjreddie ได้สร้างเอาไว้ หลังจากนั้นจึงเริ่มทำการสอน YOLO โดยใช้ไฟล์ทั้งหมดที่สร้างมา หลังจากนั้น YOLO จะค่อย ๆ บันทึกไฟล์ weights ทุก ๆ การสอนครบ 100 ครั้ง เมื่อได้ผลที่พอใจแล้วจึงนำไฟล์ weights ที่บันทึกมาไปใช้ในการทดสอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.4 ทำ Software เพื่อแยกตัวอักษรที่ติดกัน

ก่อนที่เราจะเริ่มทำการแยกตัวอักษรที่ติดกันนั้น สิ่งหนึ่งที่สำคัญคือการทำ Preprocessing รูปภาพของป้ายทะเบียนที่ได้มาจากขั้นตอนที่แล้วก่อนเพื่อให้สามารถประมวลผลได้ง่ายขึ้น ซึ่งสิ่งที่ต้องทำในขั้นตอนนี้มีดังนี้

- Gaussian Blur เพื่อลด noise ที่เกิดขึ้นในภาพ
- Denoising
- Grayscale การทำให้ภาพเป็นโทนสีเทาเพื่อที่จะทำ Threshold ภาพ
- Thresholding แบ่งส่วนภาพให้เป็นสีขาว-ดำ

หลังจากทำ Preprocessing ภาพแล้วเราจะได้ภาพป้ายทะเบียนเป็นสีขาวดำที่ภาพค่อนกลางชัด พร้อมจะทำการแยกตัวอักษรแล้ว นำภาพที่ได้มาหาตัวหนังสือแต่ละตัวในภาพโดยการใช้อัลกอริทึ่ม Connected Component Labeling

จากการใช้อัลกอริทึ่มนี้จะเป็นการกำหนด Label ให้กับแต่ละส่วนในภาพที่มีพิกเซลติดกัน ซึ่งสิ่งที่จะถูกกำหนดเป็น Label ไม่ได้มีแค่ตัวอักษรในภาพเท่านั้น แต่รวมไปถึงกรอบป้ายทะเบียน แสงสะท้อน และ noise ต่าง ๆ ที่เกิดขึ้นในภาพด้วย ดังนั้นหลังจากที่กำหนด Label แล้วจะต้องหา Label ที่คาดว่าน่าจะเป็นตัวอักษรมาวิเคราะห์ต่อไป ซึ่งวิธีที่ใช้หา Label ที่คาดว่าจะเป็นตัวอักษรนั้น เริ่มจากการหา Label ที่มีพิกเซลติดกันมากที่สุด 10 อันดับแรกมาก่อน ซึ่งเป็น Label ที่คาดว่าน่าจะรวมถึงตัวอักษรอยู่ใน 10 อันดับนี้ทั้งหมดเนื่องจาก Label ที่ถูกหามาทั้งหมดนั้น ส่วนมากจะเป็น noise ของภาพซึ่งมีพิกเซลติดกันน้อยเพราะเป็นแค่ส่วนเล็ก ๆ ภายในภาพ Label ที่จะมีพิกเซลติดกันเยอะ ๆ นั้นจะมีกรอบป้ายทะเบียน หมายเลขป้ายทะเบียน จังหวัด ซึ่งสิ่งที่เราต้องการจะหาคือหมายเลขป้ายทะเบียน 3 – 7 ตัวอักษร เมื่อหาเสร็จแล้วจะหา Label ที่คาดว่าน่าจะเป็นตัวอักษรจาก 10 Label นี้จากการหา Label ที่มีระดับความสูงระดับใกล้เคียงกันหรืออยู่ในบรรทัดเดียวกันและมีลักษณะ Label ที่อยู่ตรงส่วนกลางๆ ของภาพ ซึ่งการคัดกรองแบบนี้จะทำให้ตัดส่วนที่เป็นกรอบป้ายทะเบียน และตัวอักษรที่เป็นจังหวัดออกไปได้ หลังจากขั้นตอนนี้เราจะได้ตัวอักษรที่เป็นหมายเลขทะเบียนอย่างเดียวแล้ว แต่ยังไม่ถูกจัดเรียงว่าตัวไหนมาก่อนตัวไหนมาทีหลังเนื่องจากหมายเลข Label ไม่ได้เรียงตามตำแหน่งซึ่งเป็นผลมาจากการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ใช้ Connected Component Labeling ดังนั้นจึงต้องจัดเรียงตัวอักษรโดยเรียงจากค่าตำแหน่งในแกน x ที่ต่ำที่สุดก็ของแต่ละ Label

หลังจากหาตัวอักษรในภาพได้แล้วก็ตัดตัวอักษรและบันทึกออกมาเป็นไฟล์ภาพแต่ละตัวอักษรแยกกัน

3.5 ทำ Software เพื่อใช้ในการรู้จำตัวอักษร

การรู้จำตัวอักษรนั้น ในตอนแรกของการดำเนินงานได้ทดลองใช้วิธีการทำ Template Matching ซึ่งเป็นการเปรียบเทียบความเหมือนของภาพ โดยเริ่มจากการเตรียมรูปตัวอักษรและตัวเลขที่มีอยู่บนป้ายทะเบียนไว้อย่างละหนึ่งรูป

หลังจากนั้นนำรูปที่ได้จากขั้นตอนการแยกตัวอักษรไปเปรียบเทียบกับภาพทุกรูปที่ได้เตรียมไว้ โดยการใช้ Template matching ของไลบรารี OpenCV เพื่อหาภาพที่ให้คะแนนที่มากที่สุด เมื่อได้ทดลองใช้วิธีการทำ Template Matching มาสักระยะนึงรวมถึงไปค้นหาข้อมูลเพิ่มเติมในบทความวิจัยหลาย ๆ งานทำให้ค้นพบว่าวิธีนี้ใช้กับตัวอักษรภาษาไทยแล้วไม่ได้ผลที่ดีเท่าไรซึ่งจะแสดงให้เห็นในบทที่ 4

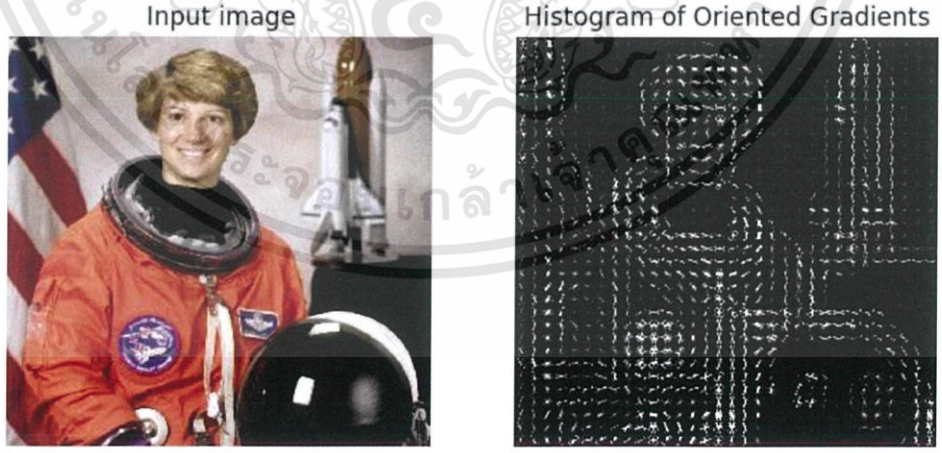
ดังนั้นจึงมาใช้อัลกอริทึม Support Vector Machine (SVM) แทนซึ่งเมื่อเป็น Machine Learning จึงคาดว่าน่าจะมีความแม่นยำมากกว่าการใช้ Image Processing ธรรมดาอย่าง Template Matching

สิ่งแรกที่ต้องเตรียมในการทำ SVM คือ dataset ของตัวอักษรและตัวเลขทุกตัว แต่ไม่เหมือนกับ Template Matching ที่เตรียมแค่อย่างละรูป แต่ต้องเตรียมไว้ยิ่งมากยิ่งดีคล้ายกับตอนทำ YOLO ซึ่งการเก็บ dataset ของตัวอักษรนั้นมีหลายวิธี ยิ่งเป็นตัวอักษรที่เป็นฟอนต์เดียวกับป้ายทะเบียนจะยิ่งดี แต่มีความยากในการทำอย่างมาก วิธีการทำคือสร้าง dataset ขึ้นมาเองและปรับให้แต่ละภาพมีมุม มี noise ที่ต่างกัน หรืออีกวิธีคือการตัดออกมาจากรูปป้ายทะเบียน ซึ่งทั้ง 2 วิธีนี้ใช้เวลานานมากในการเก็บ data ให้เยอะพอที่จะใช้ได้ เพราะฉะนั้นเราจึงเลือกที่จะหา dataset ตัวอักษรจากเว็บของศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ หรือเว็บ https://www.nectec.or.th

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 3.6 dataset ตัวอักษรจากเว็บ NECTEC

หลังจากที่มี dataset ของตัวอักษรครบแล้ว สิ่งต่อไปที่ต้องทำคือการทำ Feature Extraction เพื่อทำให้คอมพิวเตอร์สามารถประมวลผลได้ว่าลักษณะของตัวอักษรแต่ละตัวมีลักษณะเป็นอย่างไร ซึ่งไลบรารี่ OpenCV มีฟังค์ชัน Histogram of Orientation Gradients (HOG) ให้ใช้อยู่แล้ว ดังนั้นเราจึงเลือกที่จะใช้ฟังค์ชันนี้ในการทำ Feature Extraction โดยต้องมีการตั้งค่าตัวแปรแต่ละค่าให้เหมาะสม โดยค่าที่ต้องปรับมีดังนี้

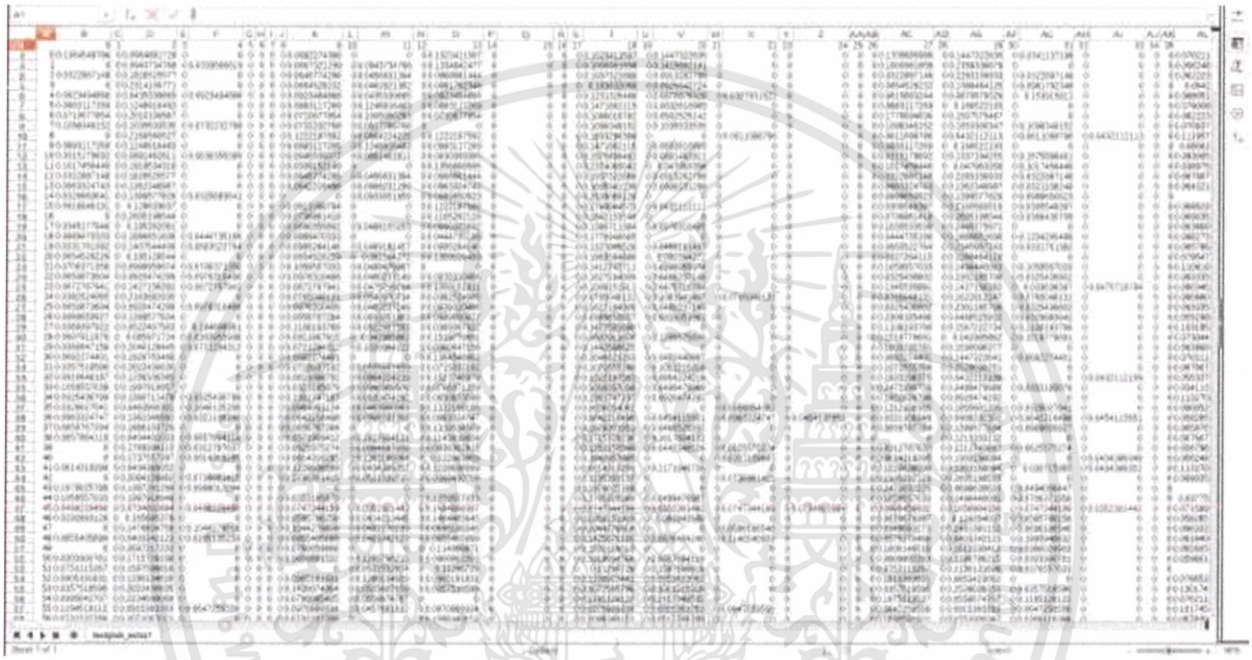


รูปที่ 3.7 ตัวอย่างการทำ HOG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- pixel_per_cell ปรับค่าเป็น (4,4) จะทำให้หน่วยเซลล์แต่ละเซลล์นั้นมี 4x4 พิกเซล
- cell_per_block ปรับค่าเป็น (2,2) เพื่อให้แต่ละบล็อคมีเซลล์อยู่ 2x2 เซลล์
- Orientations ปรับค่าเป็น 9 เพื่อแบ่งช่วงของ Feature ออกเป็น 9 ส่วน

หลังจากรันโปรแกรมจะทำให้ได้เวกเตอร์คุณลักษณะของแต่ละภาพมา และเซฟลงในตารางและเซฟเป็นไฟล์ csv ซึ่งไฟล์นี้จะถูกใช้ไปสร้างโมเดลของ SVM ในการรู้จำตัวอักษร



รูปที่ 3.8 ไฟล์ csv ที่ประกอบไปด้วยเวกเตอร์คุณลักษณะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากในไลบรารี่ของ scikit-learn มีฟังค์ชัน svm เพื่อใช้ในการสร้างโมเดลของ SVM อยู่แล้ว และต้องมีการตั้งค่าตัวแปรในการสร้างโมเดลด้วย ค่าที่จำเป็นต้องตั้งก็คือ multi-class เลือกเป็นแบบ one vs rest วิธีนี้จะทำการจำแนก (classify) ตัวอักษรโดยมีหลักการคือเมื่อคลาสแต่ละคลาส คือ ตัวอักษรแต่ละตัว การจำแนกแต่ละครั้งคือการแยกคลาสหนึ่งกับอีกทุกคลาสที่เหลืออยู่ อย่างเช่น ในครั้งแรกของการจำแนกตัวอักษร คือการแยกตัวอักษร ก.ไก่ สิ่งที่จะเกิดขึ้นในการจำแนกคือจะแยก ก.ไก่ vs ตัวอักษรตัวอื่นทุกตัว ทำให้โมเดลจะเข้าใจว่าลักษณะแบบไหนที่เป็นก.ไก่ และลักษณะแบบไหนที่ไม่ใช่ก.ไก่ ซึ่ง SVM จะทำแบบนี้กับตัวอักษรทุกตัวทำให้เข้าใจลักษณะของตัวอักษรแต่ละตัว สุดท้าย output ที่ได้จะเป็นโมเดลที่ SVM สร้างออกมา และบันทึกโมเดลออกมาเป็นไฟล์ sav เพื่อเอาไปใช้ในการรู้จำตัวอักษร

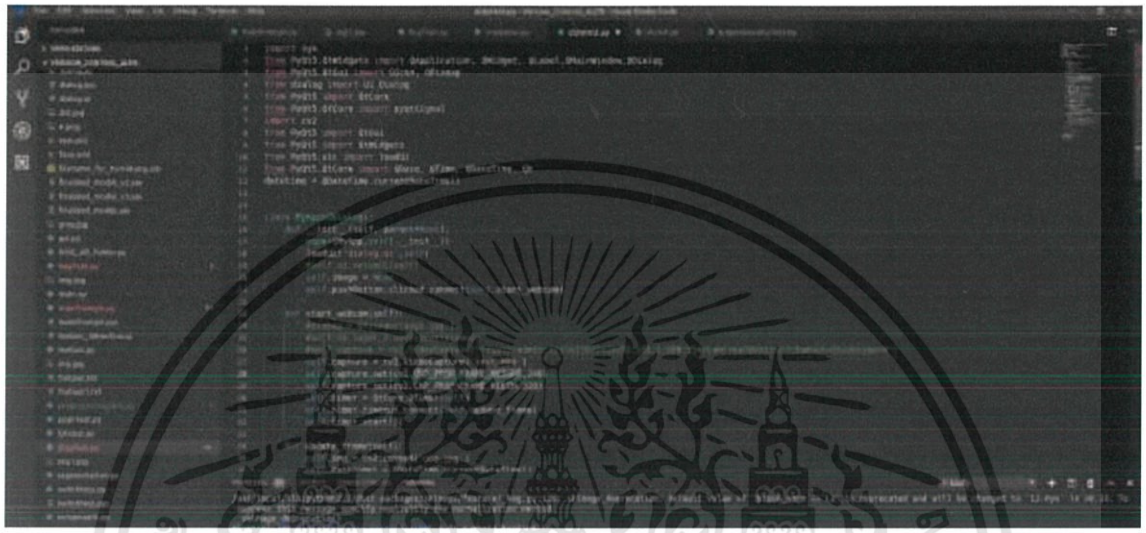
3.6 ทำหน้าต่าง User Interface สำหรับใช้งาน

หน้าต่าง User Interface นั้นทำเพียงเพื่อต้องการแสดงผลจากการรู้จำตัวอักษรเท่านั้น ดังนั้นจึงยังไม่มีความจำเป็นต้องรีบสร้างฟังค์ชันพร้อมใช้งานมากนัก แต่สามารถพัฒนาเพิ่มได้ในอนาคตเพื่อรองรับการใช้งานที่หลากหลายมากขึ้น โดยโปรแกรมที่ใช้ออกแบบหน้าต่าง User Interface คือโปรแกรม Qt Creator

รูปที่ 3.9 โปรแกรม Qt Creator

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ซึ่งโปรแกรมนี้ต้องใช้งานคู่กับไลบรารี่ PyQt5 ของ Python ในการออกแบบต่อว่าต้องการให้ส่วนไหนของหน้าต่าง User Interface มีการใช้งานแบบไหน รวมไปถึงการตั้งค่าการเปิดกล้อง การควบคุมการแสดงผลต่าง ๆ ทำได้โดยการเขียนโปรแกรมเพิ่มในไลบรารี่นี้เช่นกัน



รูปที่ 3.10 ตัวอย่างการเขียนโปรแกรมด้วย PyQt5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.7 ติดตั้งกล้องและทดสอบโปรแกรม

นำกล้องไปทดสอบในมุมต่าง ๆ ของด้านหน้าประตูทางเข้าบริษัทเพื่อทดสอบ Software ทั้งหมดที่ทำมาว่าสามารถใช้งานได้มากน้อยเพียงใด

รูปที่ 3.11 ภาพตัวอย่างการติดตั้งกล้อง

รูปที่ 3.12 ภาพตัวอย่างการติดตั้งกล้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

ผลการดำเนินงาน

4.1 ผลจากการทำ Software เพื่อตรวจจับป้ายทะเบียนจากภาพที่ได้จากกล้อง

จากบทที่ 3 Software ในส่วนนี้เริ่มจากการสอน YOLO ก่อน ซึ่งหลังจากรัน YOLO เพื่อสอนไปทั้งหมด 13300 รอบนั้น จะได้ค่า average loss อยู่ที่ 4.5181 หรือมีค่าแม่นยำอยู่ที่ 95.

รูปที่ 4.1 ภาพตัวอย่างการสอน YOLO เมื่อครบ 13301 รอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อนำค่า weights ไปทดสอบกับรูปภาพและวิดีโอจะแสดงให้เห็นว่าส่วนที่เป็นป้ายทะเบียนจะถูกตีกรอบล้อมรอบเอาไว้

รูปที่ 4.2 ภาพเมื่อทดลองกับคลิปวิดีโอถ่ายรถที่จอดอยู่

และเมื่อ YOLO ตรวจจับป้ายทะเบียนจากกล้องได้นั้น ก็จะทำการตัดภาพส่วนที่ตีกรอบไว้ออกมาอยู่ในโฟลเดอร์ ซึ่งเราจะนำภาพในโฟลเดอร์นี้ไปใช้งานต่อในขั้นตอนถัดไป

รูปที่ 4.3 ภาพป้ายทะเบียนที่ตัดออกมาจากการตรวจเจอป้ายทะเบียน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 ผลการทดสอบ Software จำแนกตัวอักษรที่ติดกัน

ทดลองทำการใช้ Connected Component Labeling กับรูปป้ายทะเบียน 25 รูปเพื่อทดสอบประสิทธิภาพของอัลกอริทึ่มในการแยกตัวอักษรที่ติดกันโดยเริ่มจาก Preprocessing รูปภาพป้ายทะเบียนที่ได้จากขั้นตอนที่แล้วก่อน โดยทำตามลำดับที่ได้อธิบายไว้ในบทที่ 3

รูปที่ 4.6 ภาพป้ายทะเบียนต้นฉบับ

รูปที่ 4.5 ภาพป้ายทะเบียนเมื่อผ่าน Gaussian Blur

รูปที่ 4.4 ภาพป้ายทะเบียนเมื่อผ่าน Denoising

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 4.7 ภาพป้ายทะเบียนเมื่อทำการ Gray Scale

รูปที่ 4.8 ภาพป้ายทะเบียนเมื่อทำ Threshold

หลังจากทำการ Preprocessing ภาพแล้วจึงเริ่มใช้อัลกอริทึ่ม Connected Component Labeling หาตัวอักษรแต่ละตัวในภาพที่ทำ Threshold มา เมื่อกำหนดให้ Label แต่ละ Label มีสีที่แตกต่างกันเพื่อแสดงผลให้เห็นภาพ จะได้ภาพหลังจากหา Label แล้วเป็นดังนี้

รูปที่ 4.9 ภาพเมื่อผ่าน Connected Component Labeling

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หลังจากได้ Label มาก็จะทำการหา Label ที่มีกลุ่มของพิกเซลที่ติดกันมากที่สุด 10 Label เพื่อหาอักษรจะได้ภาพออกมาเป็นดังนี้

รูปที่ 4.10 ภาพเมื่อทำการหากลุ่มพิกเซลเยอะสุด 10 Label

และสุดท้ายเมื่อหา Label ที่อยู่ในระดับเดียวกันคือช่วงส่วนตรงกลางของภาพ ก็จะทำให้เราได้ตัวอักษรของป้ายทะเบียนออกมา ตัดภาพแต่ละตัวอักษรแยกออกมาเป็นภาพเดี่ยว ก็เป็นอันเสร็จสิ้น

รูปที่ 4.11 ภาพเมื่อตัดตัวอักษรออกมาแต่ละตัว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.1 แสดงให้เห็นถึงผลการแยกตัวอักษรที่ติดกันกับป้ายทะเบียนตัวอย่างทั้งหมด 25 รูป

หมายเลขทะเบียนที่นำมาทดสอบ	ตัวอักษรที่สามารถตัดออกมาได้
จจ1888	จ,จ,1,8,8,8
2กม8352	2,ก,ม,8,3,5,2
1กก4640	1,ก,ก,4,6,4,0
2กส7001	2,ก,ส,7,0,0,1
3กพ9486	3,ก,พ,9,4,8,6
กก0123	ก,ก,0,1,2,3
ฎฮ7997	7,9,9
8มร8888	8,ม,ร,8,8,8,8
กจ5946	ก,จ,5,9,4,6
6กผ40	6,ก,ผ,4,0
5กง7358	5,ก,ง,7,3,5,8
ฎผ6557	ผ,6,5,5,7
กก868	ก,ก,8,6,8
ฎภ7	ฎ,ภ,7
8ปป8888	8,ป,ป,8,8,8,8
ฆม910	ฆ,ม,9,1,0
2กพ3759	2,ก,พ,3,7,5,9
ษห2338	ษ,ห,2,3,3,8
รข8450	ร,ข,8,4,5,0
ศย5726	ย,5,7,2,6
ษอ991	ษ,อ,9,9,1
4กจ35	4,ก,จ,3,5
กพ8831	ก,พ,8,8,3,1
สฐ4449	ส,ฐ,4,4,4,9
กจ7585	ก,จ,7,5,8,5
	ป้ายทะเบียนที่ถูกต้อง 22 ภาพ

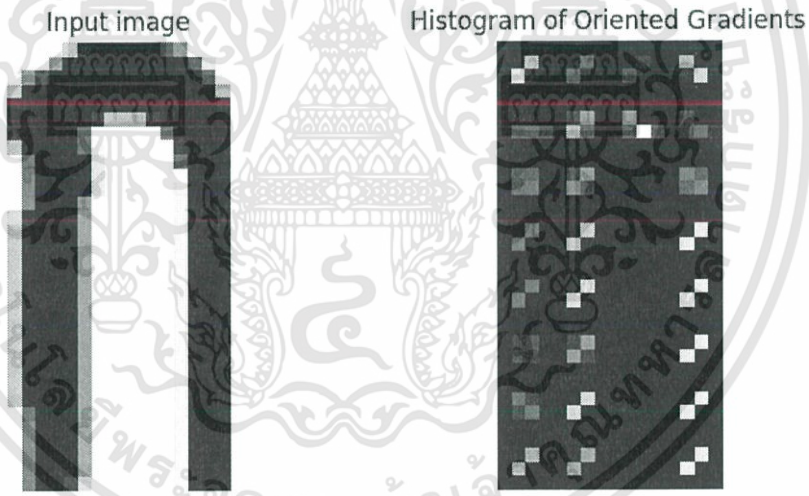
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากตารางจะเห็นได้ว่า ผลของ Software การแยกตัวอักษรที่ติดสามารถแยกได้ถูกต้องทั้งภาพ 22 ภาพจากทั้งหมด 25 ภาพ หรือสามารถบอกได้ว่า Software ส่วนนี้มีความแม่นยำอยู่ที่ 88%

4.3 ผลจากการทำ Software รู้จำตัวอักษร

จากบทที่ 3 หลังจากที่ได้ทำการสร้างโมเดลที่ใช้ในการรู้จำตัวอักษรจาก SVM แล้ว เราจะนำเอารูปตัวอักษรแต่ละตัวที่ได้มาจากการแยกตัวอักษรในขั้นตอนที่แล้วมาผ่านโมเดลนี้ทีละตัว เพื่อให้โมเดลทำนายออกมาว่าตัวอักษรแต่ละตัวนั้นคือตัวอะไร

โดยก่อนที่จะนำตัวอักษรเข้าไปผ่านโมเดลนั้น จะต้องทำ Feature Extraction เหมือนกับภาพตัวอักษรที่นำไปสร้างโมเดลก่อนเช่นกัน เพื่อให้โมเดลสามารถคำนวณเปรียบเทียบลักษณะเด่นของภาพที่จะนำมาทดสอบได้



รูปที่ 4.12 ทำ HOG ตัวอักษร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.2 แสดงผลตัวอย่างของการรู้จำตัวอักษรจากภาพป้ายทะเบียนตัวอย่าง 65 ภาพที่ได้มาจากขั้นตอนการจำแนกตัวอักษรที่ติดกัน

หมายเลขป้ายทะเบียน	ผลการรู้จำตัวอักษร	หมายเลขป้ายทะเบียน	ผลการรู้จำตัวอักษร
จจ1888	จจ1888	ฐง8181	ฐ8181
2กม8352	2กม8352	1กก9998	1ถถ9998
1กก4640	1กก4640	กฉ9926	ดฉ9926
2กส7001	2กส7001	ศย5726	ย5726
3กพ9486	3กพ9486	ศธ8983	ศธ8983
กก0123	กก0123	กฉ3371	กฉ3371
ฎฮ7997	799	ภท8707	ภท8707
8มร888	8มร888	กน564	กน564
กจ5946	กจ5946	ตว5738	ตว5738
6กผ40	6กผ40	ชห9515	ชห9515
5กง7358	5กง7358	ธจ6696	ธจ6696
กก777	กก777	กค9494	กค9494
ฎผ6557	ผ6557	6กฆ2221	6กฆ2221
1กฮ4040	1กฮ4040	3กพ9486	3กพ9486
นน88	นน88	กท9989	กท9989
5กร3963	5กร3963	ฎง9999	ฎง9999
กก868	กก868	ศธ8983	ศธ8983
ฎผ3889	ฎผ3อ89	ญณ9999	ญณ9999
ฎผ7	ฎผ7	ษอ991	ษอ991
1กก1614	1กก1614	ขท99	ขก99
8ปป8888	8ปป8888	กร8582	กร8582
ฆม910	ฆม910	4กจ35	4กจ35
2กพ3759	2กพ3759	2กอ389	2กอ389
ษห2338	ษห2338	สฐ4449	สฐ4449
รข8450	รข8450	สส777	สส777
ฉว5756	ฉว5756	กพ8831	กพ8831

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

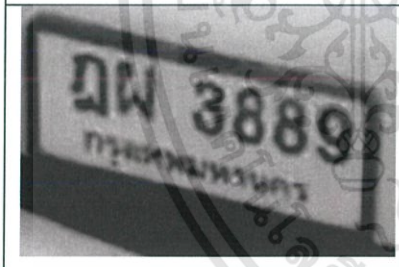
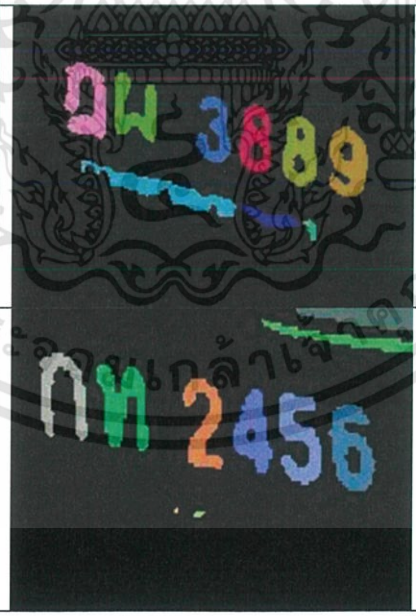
ตารางที่ 4.2 แสดงผลตัวอย่างของการรู้จำตัวอักษรจากภาพป้ายทะเบียนตัวอย่าง 65 ภาพที่ได้มาจากขั้นตอนการจำแนกตัวอักษรที่ติดกัน (ต่อ)

หมายเลขป้ายทะเบียน	ผลการรู้จำตัวอักษร	หมายเลขป้ายทะเบียน	ผลการรู้จำตัวอักษร
กก9999	กก9999	จฉ4444	จฉ4444
4กจ944	4กจ944	6ษ7498	67898
4กถ2824	4กถ2824	1กข2013	1กข2013
9กค9535	9กค9535	9ษ7098	9ษ7098
กท2456	ฐก2ง5ฐ	กฉ5586	กฉ5586
บค1111	บค1111	กจ7585	กจ7585
		สห7272	สห7272

จากตารางจะเห็นได้ว่ามีผลการรู้จำตัวอักษรที่ถูกต้องอยู่ทั้งหมด 55 ภาพจากทั้งหมด 65 ภาพ หรือสามารถบอกได้ว่าความแม่นยำของ Software ส่วนรู้จำตัวอักษรอยู่ที่ 84.62% ซึ่งตารางข้างล่างนี้เป็นตารางแสดงผลของภาพที่ไม่ถูกต้องเพื่อหาสาเหตุของความผิดพลาดของ Software

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.3 แสดงสาเหตุของการรู้จำที่ผิดพลาด

ป้ายทะเบียน	ผลการ Connected Component Labeling	ผลการรู้จำตัวอักษร	สาเหตุ
		799	ป้ายทะเบียนลายกราฟิกส่งผลต่อการแยกตัวอักษร
		ผ6557	ความไม่ชัดของป้ายทะเบียนทำให้ตัวอักษร ฎ ขาดหายไป
		ฎผ3อ89	ความเอียงและความไม่ชัดของป้ายทะเบียนส่งผลต่อการรู้จำ
		ฐก2ง5ฐ	ความเอียงและความไม่ชัดของป้ายทะเบียนส่งผลต่อการรู้จำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.3 แสดงสาเหตุของการรู้จำที่ผิดพลาด (ต่อ)

ป้ายทะเบียน	ผลการ Connected Component Labeling	ผลการรู้จำตัวอักษร	สาเหตุ
		ฐ8181	กลุ่มของพิกเซลของตัวอักษร ง น้อยกว่า 10 ตัวแรก ทำให้แยกออกมาไม่ได้
		1ถถ9998	ความคล้ายกันของตัวอักษร
		ดฉ9926	ความคล้ายกันของตัวอักษร
		ย5726	ป้ายทะเบียนลายกราฟิกส่งผลให้ตอนแยกตัวอักษร ตัว ศ หายไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.3 แสดงสาเหตุของการรู้จำที่ผิดพลาด (ต่อ)

ป้ายทะเบียน	ผลการ Connected Component Labeling	ผลการรู้จำตัวอักษร	สาเหตุ
		ขก99	แสงสะท้อนบนป้ายทะเบียนส่งผลต่อการรู้จำ
		67898	ฟอนต์ที่แตกต่างจากปกติส่งผลต่อการรู้จำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.4 ผลการทดลองกับวิดีโอจากกล้องวงจรปิด

เนื่องจากโครงงานนี้ทำในส่วนของ Software ของบริษัท ทำให้ยังไม่สามารถหาสถานที่ที่จะติดกล้องเพื่อทดสอบที่เหมาะสมได้ตั้งแต่ต้น จึงต้องทำการนำกล้องวงจรปิดไปตั้งในมุมต่าง ๆ ของด้านหน้าบริษัทแล้วได้พบว่าการตั้งกล้องที่บริเวณทางเท้าส่งผลให้ไม่สามารถได้ภาพป้ายทะเบียนที่สามารถนำมาใช้งานได้เนื่องจากมุมที่เอียงจนเกินไป และเมื่อลองย้ายกล้องไปบริเวณที่ใกล้กับรถมากขึ้นจะเริ่มสามารถใช้งานได้ เพราะฉะนั้นเมื่อถึงเวลาที่ Software จะถูกนำไปใช้งานจริงควรติดตั้งในบริเวณที่ยิ่งใกล้กับด้านหน้าของรถเท่าไรยิ่งดี

รูปที่ 4.13 หน้าต่าง User Interface

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปผลการดำเนินงานและข้อเสนอแนะ

5.1 สรุปผลการดำเนินงาน

โครงงาน Automated License Plate Recognition เป็นโครงงานตามความต้องการของแผนก Advanced Manufacturing Engineering เพื่อให้เป็นระบบที่เพิ่มความปลอดภัยในการจัดการรถที่เข้าออกบริษัทโดยจะสามารถสรุปโครงงานนี้ได้ดังนี้

5.1.1 ระบบสามารถจำแนกภาพป้ายทะเบียนออกมาจากภาพต้นฉบับของกล้องได้

ซึ่งจากบทที่ 4 จะเห็นว่าการใช้ YOLO มีประสิทธิภาพสูงในการตรวจหาป้ายทะเบียนจากภาพหรือวิดีโอ แต่ในอีกแง่หนึ่ง YOLO นั้นใช้ทรัพยากรของคอมพิวเตอร์ที่นำมาใช้งานสูงมาก โดยเฉพาะการ์ดจอที่ใช้ในการประมวลผลภาพ

5.1.2 ระบบสามารถจำแนกตัวอักษรที่ติดกันและรู้จำตัวอักษรได้

จากบทที่ 4 จะเห็นว่าผลของการจำแนกตัวอักษรที่ติดกันและรู้จำตัวอักษรสามารถทำได้ผลที่น่าพึงพอใจแม้อาจจะไม่ได้สมบูรณ์ 100% แต่ก็เพียงพอที่จะนำไปใช้งานได้

5.2 ปัญหาที่พบและแนวทางแก้ไข

5.2.1 ปัญหาที่พบ

5.2.1.1 สิ่งแวดล้อมที่ไม่สามารถควบคุมได้ ซึ่งสิ่งแวดล้อมในที่นี้หมายถึงแสงที่เปลี่ยนไปในแต่ละวัน ทำให้ภาพป้ายทะเบียนบางภาพที่ถ่ายมามีแสงที่สว่างเกินไป บางภาพอาจจะมืดเกินไป บางภาพอาจจะเกิดแสงสะท้อนจนส่งผลให้ประมวลผลภาพได้ยาก ซึ่งปัญหานี้ส่งผลอย่างมากในขั้นตอนการแยกตัวอักษรที่ติดกัน

5.2.1.2 ความคล้ายกันของตัวอักษรภาษาไทย เนื่องจากตัวอักษรภาษาไทยมีตัวอักษรหลายตัวที่คล้ายกันมาก ส่งผลให้ขั้นตอนการรู้จำตัวอักษรนั้นมีปัญหาในการประมวลผลมาก ดังที่แสดงให้เห็นในบทที่ 4 ว่าวิธีการเปรียบเทียบความคล้ายกันของภาพไม่สามารถนำมาใช้ได้กับตัวอักษรภาษาไทยแม้ว่าวิธีนี้จะใช้กับภาษาอื่นได้ก็ตาม ซึ่งก็เป็นผลมาจากความคล้ายกันของตัวอักษรนั่นเอง และเมื่อมาใช้ SVM แล้ว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แม้จะมีประสิทธิภาพมากขึ้น แต่ก็ยังส่งผลในการทำนายตัวอักษรอยู่นั่นเอง ซึ่งตัวอักษรที่มีความใกล้เคียงกันที่ส่งผลต่อการรู้จำของ SVM มีดังนี้

1. ก กับ ถ กับ ภ
2. ด กับ ค
3. พ กับ ฟ
4. ท กับ ห
5. บ กับ ป
6. ข กับ ช

5.2.1.3 คุณภาพของภาพป้ายทะเบียนที่ยังไม่ดีมากพอ เนื่องจากภาพป้ายทะเบียนที่ได้จากกล้องนั้นมีคุณภาพที่ไม่สูงมาก ซึ่งอาจจะเกิดจากความเร็วของรถซึ่งส่งผลต่อความไม่ชัดของภาพ หรือภาพที่ได้มาอาจจะเอียงทำให้ประมวลผลได้ยาก และอีกสิ่งหนึ่งที่ยังส่งผลต่อคุณภาพของภาพคือ Software ส่วนที่ตรวจจับป้ายทะเบียนจากภาพในกล้อง ซึ่ง YOLO นั้น ด้วยความที่ใช้ทรัพยากรของคอมพิวเตอร์สูงมาก ทำให้เมื่อภาพมาผ่าน Software ส่วนนี้จะถูกปรับขนาดลงไปเกินครึ่ง ทำให้ภาพป้ายทะเบียนที่ส่งไปยังส่วนถัดไปของ Software มีคุณภาพแย่ลงจากต้นฉบับค่อนข้างเยอะ

5.2.1.4 ป้ายทะเบียนที่มีลักษณะต่างจากป้ายทะเบียนปกติ ลักษณะของป้ายทะเบียนที่ต่างจากป้ายทะเบียนปกติ เช่น ป้ายทะเบียนที่ชำรุด ตัวอักษรบนป้ายทะเบียนขาดหาย หรือใช้กรอบและพื้นหลังป้ายทะเบียนที่ไม่ใช่สีขาวทำให้การรู้จำตัวอักษรมีความผิดพลาดได้ง่าย เพราะภาพจะถูกประมวลผลได้ยากในทุกส่วนของ Software

5.2.2 แนวทางการแก้ไข

5.2.2.1 คุณภาพของภาพป้ายทะเบียน สามารถเพิ่มได้ด้วยการใช้กล้องวงจรปิดที่มีคุณภาพสูงขึ้นในการทดสอบ และเลนส์ของกล้องไม่ควรจะมีทางยาวโฟกัสที่น้อยเกินไป เพราะจะทำให้ภาพที่ได้นั้นโค้งและนำมาประมวลผลยาก

5.2.2.2 ความคล้ายกันของตัวอักษร สามารถเพิ่มความสามารถในการรู้จำของ SVM ได้โดยการเตรียมภาพตัวอักษรที่จะนำมาใช้สร้างโมเดลจากภาพป้ายทะเบียนจริง ใช้ฟอนต์ที่เหมือนฟอนต์ป้ายทะเบียนจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.3 ข้อเสนอแนะ

5.3.1 Software ส่วนที่ตรวจจับป้ายทะเบียนนั้น หากเปลี่ยนเป็นอัลกอริทึ่มอื่นที่ใช้ทรัพยากรน้อยลงกว่านี้ อาจจะทำให้ภาพป้ายทะเบียนที่ได้มามีคุณภาพที่สูงมากขึ้น และอาจจะมีความง่ายในการทำมากยิ่งขึ้นถ้าภาษาที่ใช้ในการเขียนในส่วนนี้เป็นภาษา Python เหมือนกับส่วนอื่น ๆ

5.3.2 ควรจะมีจุดจอดรถก่อนที่รถจะเข้าบริษัทช่วยให้ Software ทำงานได้ดีขึ้น เนื่องจากเราจะสามารถนำกล้องมาติดตั้งใกล้กับบริเวณที่รถจอดได้และทำให้ได้ภาพในมุมตรงมากขึ้น และยังทำให้ไม่เกิดการเบลอของภาพที่ได้อีกด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

[1] Histogram of Orientation Gradient

แหล่งที่มา: https://www.learnopencv.com/histogram-of-oriented-gradients/

[2] OpenCV

แหล่งที่มา: https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_tutorials.html

[3] Support Vector Machine

แหล่งที่มา: https://blog.statsbot.co/support-vector-machines-tutorial-c1618e635e93

[4] You Only Look Once

แหล่งที่มา: https://pjreddie.com/darknet/yolo/

[5] Connected Component Labeling

แหล่งที่มา: https://homepages.inf.ed.ac.uk/rbf/HIPR2/label.htm

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้