

การออกแบบและพัฒนาโมดูลการมองเห็นสำหรับประยุกต์ใช้ในกระบวนการ
ตรวจสอบคุณภาพชิ้นงาน และแยกแยะตำหนิบนผิววัตถุด้วยปัญญาประดิษฐ์



นายไตรศรินทร์ แสงเสียงฟ้า

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมเมคคาทรอนิกส์
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2564

DESIGN AND DEVELOPMENT OF VISION MODULE FOR DETECTION
AND WORKPIECE QUALITY AND CLASSIFICATION OF
DEFECT USING AI



A Thesis Submitted in Partial Fulfillment of the Requirements for the
Degree of Master of Engineering in Mechatronics Engineering
Suranaree University of Technology
Academic Year 2021

การออกแบบและพัฒนาโมดูลการมองเห็นสำหรับประยุกต์ใช้ในกระบวนการตรวจสอบ
คุณภาพชิ้นงาน และแยกแยะตำหนิบนผิววัตถุด้วยปัญญาประดิษฐ์

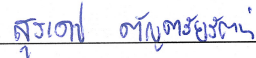
มหาวิทยาลัยเทคโนโลยีสุรนารี อนุมัติให้บัณฑิตวิทยาลัยฉบับนี้เป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาโทบริหารธุรกิจ

คณะกรรมการสอบวิทยานิพนธ์



(รศ. ดร.จิระพล ศรีเสรีสุผล)

ประธานกรรมการ



(อ. ดร.สุรเดช ตัญจรัยรัตน์)

กรรมการ (อาจารย์ที่ปรึกษาวิทยานิพนธ์)



(ผศ. ดร.ชัยยุทธ์ สัมภาวะคุปต์)

กรรมการ

มหาวิทยาลัยเทคโนโลยีสุรนารี



(รศ. ดร.ฉัตรชัย โชติษฐียงกูร)

รองอธิการบดีฝ่ายวิชาการและประกันคุณภาพ



(รศ. ดร.พรศิริ จงกล)

คณบดีสำนักวิชาวิศวกรรมศาสตร์

ไตรศรินทร์ แสงเสียงฟ้า : การออกแบบและพัฒนาโมดูลการมองเห็นสำหรับประยุกต์ใช้ใน
กระบวนการตรวจสอบคุณภาพชิ้นงาน และแยกแยะตำหนิบนผิววัตถุด้วยปัญญาประดิษฐ์
(DESIGN AND DEVELOPMENT OF VISION MODULE FOR DETECTION AND
WORKPIECE QUALITY AND CLASSIFICATION OF DEFECT USING AI)
อาจารย์ที่ปรึกษา : อาจารย์ ดร.สุรเดช ตัญจรัยรัตน์, 90 หน้า.

คำสำคัญ: ปัญญาประดิษฐ์/การประมวลผลภาพ

การผลิตชิ้นส่วนยานยนต์ คุณภาพของผลิตภัณฑ์ถือเป็นสิ่งแรกที่ทางผู้ผลิตต้องให้
ความสำคัญในกระบวนการผลิต งานวิจัยนี้ ได้นำเสนอการออกแบบและพัฒนาโมดูลการมองเห็น
แบบอัจฉริยะเพื่อตรวจสอบคุณภาพชิ้นงานและแยกแยะตำหนิบนผิววัตถุ โดยใช้หลักการประมวลผล
ภาพ (Image Processing) และประยุกต์ใช้ปัญญาประดิษฐ์ (Artificial Intelligence) ที่มีอัลกอริทึม
แบบโครงข่ายประสาทเทียมสังวัตนาการ (Convolutional Neural Network) ช่วยวิเคราะห์
แยกแยะประเภทรอยต่าง ๆ เสมือนการวิเคราะห์จำลองการมองเห็นของมนุษย์ อุปกรณ์ต้นแบบเป็น
แบบระบบประมวลผลที่รองรับโครงข่ายประสาทเทียม (Neural Network) สำหรับแยกประเภท
ปัญหาบนผิววัตถุ โดยทดสอบกับแม่แรงรถยนต์พื้นผิวสีดำ ขนาด 10x10x19 เซนติเมตร (กว้างxยาวx
สูง) ในสภาพแวดล้อมแสงที่ถูกควบคุม ผลลัพธ์จากการทดสอบ ค่าเฉลี่ยความแม่นยำอยู่ที่ 95.6
เปอร์เซ็นต์และประสิทธิภาพการแยกประเภทของโมเดลกับตัวอย่างค่าเฉลี่ยอยู่ที่ 96 เปอร์เซ็นต์
สามารถตรวจจับปัญหาที่เกิดขึ้นบนผิววัตถุทั้ง 4 ชนิดตามที่กำหนดไว้ในขั้นตอนการเรียนรู้เชิงลึก
(Deep Learning) และสามารถตรวจจับได้มากกว่า 1 รอยตำหนิบนวัตถุอันเดียวกันโดยไม่เกิดปัญหา
Underfitting หรือ Overfitting

สาขาวิชา วิศวกรรมเมคคาทรอนิกส์
ปีการศึกษา 2564

ลายมือชื่อนักศึกษา ไตรศรินทร์ แสงเสียงฟ้า
ลายมือชื่ออาจารย์ที่ปรึกษา สุรเดช ตัญจรัยรัตน์

TRAISIRIN SAENGSEANGFA : DESIGN AND DEVELOPMENT OF VISION MODULE FOR DETECTION AND WORKPIECE QUALITY AND CLASSIFICATION OF DEFECT USING AI. THESIS ADVISOR : SURADET TANTRAIRATN, Ph.D., 90 PP.

Keyword: Artificial Intelligence/Image Processing

In automotive parts manufacture, the quality of the products is the most important issue for a manufacturer. This research presents the design and development of an intelligent vision module classifying the defects on object surfaces using image processing, artificial intelligence, convolutional neural network algorithm as if using human vision. The developed prototype was a neural network-supported computing system for classifying the defects on the surface of black car jacks of 10x10x19 cm (W x L x H) in an environment with specific brightness. The test result was the average precision of 95.6%, success rate classification was 96%, showing that it could detect defects on the four types of surfaces previously specified in the Deep Learning process. It could also detect more than one defects on the same object in real time without underfitting or overfitting problem.

มหาวิทยาลัยเทคโนโลยีสุรนารี

School of Mechatronic Engineering

Academic Year 2021

Student's Signature ไตรสินทร์ แสงแสงงฟ้า

Advisor's Signature สุรเดช ตันตรไรรัตน์

กิตติกรรมประกาศ

วิทยานิพนธ์นี้สำเร็จลุล่วงด้วยดี เนื่องจากได้รับความช่วยเหลืออย่างยิ่ง ทั้งด้านวิชาการและด้านการดำเนินงานวิจัย จากบุคคลและกลุ่มบุคคลต่าง ๆ ได้แก่

อาจารย์ ดร.สุรเดช ตัญจรัยรัตน์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ให้โอกาสทางการศึกษา ให้คำแนะนำปรึกษา ช่วยแก้ปัญหาและให้กำลังใจแก่ผู้วิจัยมาโดยตลอด รวมทั้งช่วยตรวจทาน และแก้ไขวิทยานิพนธ์เล่มนี้จนเสร็จสมบูรณ์

สำหรับคุณงามความดีอันใดที่เกิดจากวิทยานิพนธ์เล่มนี้ ผู้วิจัยขอมอบให้กับบิดา มารดา และป้า ยาย ซึ่งเป็นที่รักและเคารพยิ่ง ตลอดจนครูอาจารย์ที่เคารพทุกท่าน ที่ได้ประสิทธิ์ประสาทวิชาความรู้และถ่ายทอดประสบการณ์ที่ดีให้แก่ผู้วิจัยตลอดมา จนทำให้ประสบความสำเร็จในชีวิต

ไตรศรินทร์ แสงเสียงฟ้า



สารบัญ

หน้า

บทคัดย่อ (ภาษาไทย).....	ก
บทคัดย่อ (ภาษาอังกฤษ).....	ข
กิตติกรรมประกาศ.....	ค
สารบัญ.....	ง
สารบัญตาราง.....	ช
สารบัญรูป.....	ซ
บทที่	
1 บทนำ.....	1
1.1 ความสำคัญและที่มาของปัญหาการวิจัย.....	1
1.2 วัตถุประสงค์ของการวิจัย.....	1
1.3 สมมุติฐานของการวิจัย.....	2
1.4 ขอบเขตของการวิจัย.....	2
1.5 ระเบียบวิธีวิจัย.....	2
1.6 ประโยชน์ที่คาดว่าจะได้รับ.....	3
2 ปรัชญาบรรณกรรมและทฤษฎีที่เกี่ยวข้อง.....	4
2.1 กล่าวนำ.....	4
2.2 ปรัชญาบรรณกรรม.....	4
2.3 การประมวลผลภาพ.....	10
2.3.1 ภาพดิจิทัล.....	11
2.3.2 การประมวลผลระดับพิกเซล.....	16
2.3.3 การทำสังวัตนาการ (Convolution).....	17
2.3.4 ฟังก์ชันเกาส์เซียน (Gaussian).....	21
2.4 ปัญญาประดิษฐ์.....	21
2.4.1 ความหมายของปัญญาประดิษฐ์.....	22

สารบัญ (ต่อ)

หน้า

2.4.2	คุณสมบัติและข้อจำกัดของปัญญาประดิษฐ์.....	22
2.4.3	ข้อเปรียบเทียบและข้อจำกัดของปัญญาประดิษฐ์.....	23
2.4.4	ประโยชน์ของปัญญาประดิษฐ์.....	24
2.4.5	การนำปัญญาประดิษฐ์มาประยุกต์ใช้งาน.....	25
2.5	โครงข่ายประสาทเทียม.....	30
2.5.1	ประวัติความเป็นมาของโครงข่ายประสาทเทียม.....	31
2.5.2	หลักการทำงานของระบบโครงข่ายประสาทเทียมในคอมพิวเตอร์.....	33
2.5.3	การเรียนรู้ของโครงข่ายประสาทเทียม.....	34
2.5.4	สถาปัตยกรรมโครงข่ายประสาทเทียม.....	38
2.5.5	ประโยชน์ของโครงข่ายประสาทเทียม.....	41
2.5.6	การประยุกต์ใช้งานโครงข่ายประสาทเทียม.....	42
2.6	อัลกอริทึมพื้นฐานของ YOLOv5.....	42
2.7	ไมโครคอนโทรลเลอร์เบื้องต้น.....	45
2.7.1	โครงสร้างไมโครคอนโทรลเลอร์.....	45
2.7.2	ไมโครคอนโทรลเลอร์และไมโครโพรเซสเซอร์.....	48
3	วิธีการดำเนินงานวิจัย.....	49
3.1	กล่าวนำ.....	49
3.2	ซอฟต์แวร์และอุปกรณ์ที่ใช้ใน.....	49
3.2.1	ระบบปฏิบัติการและภาษาคอมพิวเตอร์.....	49
3.2.2	อุปกรณ์.....	49
3.2.3	การติดตั้งอุปกรณ์บนระบบสายพานลำเลียงวัตถุ.....	51
3.3	ขั้นตอนการดำเนินการวิจัย.....	52
3.3.1	ขั้นตอนการเก็บข้อมูลตัวอย่างที่ทำโมเดล.....	53
3.3.2	ขั้นตอนการเตรียมข้อมูลตัวอย่างสำหรับเทรนนิ่ง.....	55
4	ผลการศึกษาและวิเคราะห์ผล.....	62
4.1	กล่าวนำ.....	62
4.2	การวัดประสิทธิภาพโมเดลแบบ Confusion Matrix.....	62

สารบัญ (ต่อ)

	หน้า
4.3 การวัดประสิทธิภาพการแยกประเภทรอยต่าง ๆ	63
4.4 การทดสอบประสิทธิภาพโมเดลกับวัตถุจริงแบบเรียลไทม์	64
4.5 การทดสอบประสิทธิภาพโมเดลกับข้อจำกัดของอุปกรณ์และอัลกอริทึม	66
4.6 สรุป.....	67
5 สรุปผลการวิจัย.....	68
5.1 สรุปผลการวิจัย	68
5.2 ปัญหาที่พบในการวิจัย	68
5.2.1 ปัญหาแสงสว่างที่ไม่เหมาะสมทำให้มีแสงสะท้อนบนผิววัตถุ.....	68
5.2.2 ข้อจำกัดของงานวิจัย.....	69
5.2 ข้อเสนอแนะ	69
5.3 ข้อสังเกต.....	70
รายการอ้างอิง.....	71
ภาคผนวก	
ภาคผนวก ก. คำสั่งโค้ดตั้งเทรนนิ่งโมเดล.....	74
ภาคผนวก ข. การใช้โปรแกรม Label IMG	76
ภาคผนวก ค. โครงสร้างอุปกรณ์ NVIDIA Jetson Nano	79
ภาคผนวก ง. บทความวิชาการที่ได้รับการตีพิมพ์เผยแพร่ในระหว่างการศึกษา.....	82
ประวัติผู้เขียน.....	90

สารบัญตาราง

ตารางที่	หน้า
3.1	การทำAugmentation Pattern..... 56
3.2	การแบ่งข้อมูลสำหรับเทรนนิ่งโมเดล 57
3.3	การตั้งค่าพารามิเตอร์ของโมเดล 59
3.4	ค่าความแม่นยำจากการทำ transfer learning..... 60
4.1	การทดสอบประสิทธิภาพความถูกต้องการแยกประเภทรอยต่าง ๆ บนวัตถุ 64
4.2	การทดสอบประสิทธิภาพโมเดลกับข้อจำกัดของอุปกรณ์และอัลกอริทึม..... 67



สารบัญรูป

รูปที่		หน้า
2.1	ตัวอย่างรูปภาพของแผ่นวงจรอิเล็กทรอนิกส์.....	5
2.2	การตรวจจ็บรอยตำหนิ	6
2.3	การตรวจจ็บยานพาหนะ.....	7
2.4	ตัวอย่างข้อมูลภาพสัญญาณไฟรถไฟ	8
2.5	การเทรนนิ่งโมเดลการเทรนนิ่งโมเดลจาก Global Wheat Head Detection Dataset	9
2.6	ประสิทธิภาพโมเดลหลังจากทำการเทรนนิ่ง.....	10
2.7	การประมวลผลภาพและผลลัพธ์ในรูปแบบต่าง ๆ	12
2.8	โครงสร้างการทำงานของตามนุษย์(ก) และ(ข)กล็องดิจิตัล	14
2.9	การเปลี่ยนจากภาพสีเป็นภาพระดับเทาโดยการแปลงค่าสี.....	16
2.10	ความสัมพันธ์ระหว่างพิกเซลรอบจุด (x, y)	17
2.11	การทำสังวัตนาการระหว่าง f และ w	19
2.12	การทำสังวัตนาการระหว่าง z และ w	20
2.13	ตัวอย่างโครงสร้างโครงข่ายประสาทในสมองและโครงข่ายประสาทเทียม	28
2.14	โครงสร้างของเซลล์ประสาทในสมองมนุษย์	30
2.15	โครงสร้างของเซลล์ประสาท (Neurons) ในคอมพิวเตอร์.....	33
2.16	การเรียนรู้แบบมีผู้สอน (Supervised Learning)	37
2.17	การเรียนรู้แบบไม่มีผู้สอน Unsupervised Learning	38
2.18	สถาปัตยกรรมของ Feed-Forward Network	39
2.19	สถาปัตยกรรมของ Feed-Back Network.....	39
2.20	สถาปัตยกรรมของ Network Layer.....	40
2.21	Grid Cellในภาพเพื่อทำการค้นหาวัตถุ (15 x 12)	43
2.22	สมการ Intersection over Union.....	43
2.23	พารามิเตอร์ของ bounding box	44
2.24	โมเดล Yolo ขนาดgrid cell 3 x3 ทำนายวัตถุประกอบด้วย 3 คลาส	45
2.25	โครงสร้างของโมโครคอนโทรลเลอร์	46

สารบัญรูป (ต่อ)

รูปที่	หน้า
2.26	การสื่อสารของไมโครคอนโทรลเลอร์ 47
3.1	อุปกรณ์ต้นแบบ 50
3.2	ระบบการทำงานของอุปกรณ์..... 51
3.3	การติดตั้งอุปกรณ์ 52
3.4	กระบวนการพ่นสีวัตถุ..... 53
3.5	แผนภูมิเก้าปลา 7 Tools 54
3.6	ผลลัพธ์หลังจากปรับความสว่างและทำ Contrast 55
3.7	การใช้โปรแกรม Label IMG สร้าง Bounding Box 57
3.8	ค่าตัวแปรหลังจากทำ Bounding box 58
3.9	การ Batch size ข้อมูลเพื่อทำการ Training..... 59
3.10	ตัวแปร weight ของโมเดล 60
3.11	กราฟการเทรนนิ่งข้อมูล 100 epoches 61
4.1	กราฟค่าความแม่นยำของผลลัพธ์การสร้างโมเดล..... 63
4.2	การตรวจสอบรอยสีลอกกับแม่แรงรถยนต์ขนาด 2 ตัน 65
4.3	การตรวจสอบรอยขีดข่วนและรอยกระแทกกับแม่แรงรถยนต์ขนาด 3 ตัน 65
4.4	การตรวจสอบรอยกระแทกและรอยวัตถุดิบราวกับแม่แรงรถยนต์ขนาด 5 66
5.1	องศาของแสงแนวตั้งฉากกับวัตถุ..... 69
ก.1	โค้ดดึงป้อนข้อมูลเข้าโมเดล..... 75
ข.1	การเรียกใช้โปรแกรม Label IMG 77
ข.2	การใช้โปรแกรม Label IMG 78
ค.1	โครงสร้างอุปกรณ์ NVIDIA Jetson Nano ด้านหน้าและหลัง 80
ค.2	โครงสร้างอุปกรณ์ NVIDIA Jetson Nano ด้านบนและด้านล่าง 81

บทที่ 1

บทนำ

1.1 ความสำคัญและที่มาของปัญหาการวิจัย

ในประเทศไทยภาคอุตสาหกรรมการผลิตยานยนต์ถือเป็นหนึ่งในอุตสาหกรรมที่ขับเคลื่อนและพัฒนาประเทศมาอย่างยาวนาน ยอดการผลิตแต่ละปีเฉลี่ย 1.5 ล้านคัน รวมการผลิตขายภายในประเทศและส่งออกต่างประเทศ ในแต่ละโรงงานของกลุ่มอุตสาหกรรมยานยนต์ กระบวนการผลิตจำเป็นต้องมีกระบวนการควบคุมคุณภาพการผลิตของชิ้นงาน

การควบคุมคุณภาพการผลิตของชิ้นงาน เป็นการตรวจสอบชิ้นงานตามคุณสมบัติสเปคการออกแบบชิ้นงาน ในภาคอุตสาหกรรมส่วนใหญ่ยังคงใช้มนุษย์ในการตรวจสอบชิ้นงาน เพราะเป็นวิธีที่ง่ายแต่ยังไม่สามารถการันตีได้ว่ามีประสิทธิภาพ 100% เนื่องจากมนุษย์ยังคงมีข้อจำกัดในทางด้านกายภาพของร่างกาย สภาพจิตใจ รวมถึงต้องใช้ทักษะความรู้ในการตรวจสอบ เพื่อเพิ่มประสิทธิภาพและลดปัญหา Human Error การนำเทคโนโลยีการประมวลผลภาพเป็นตัวเลือกถูกพิจารณานำมาประยุกต์ใช้ ในปัจจุบันมีอัลกอริทึมที่พัฒนาขึ้นตามการพัฒนาของอุตสาหกรรมคอมพิวเตอร์

งานวิจัยฉบับนี้จึงนำเทคโนโลยีการประมวลผลภาพและอัลกอริทึมโครงข่ายประสาทเทียมแบบสังวัตนาการของการเรียนรู้เชิงลึกปัญญาประดิษฐ์มาประยุกต์ใช้ในการสร้างโมเดลอุปกรณ์ต้นแบบ เพื่อตรวจจับและแยกแยะรอยตำหนิวัตถุเสมือนการมองเห็นมนุษย์ โดยเลือกใช้อุปกรณ์ NVIDIA Jetson Nano ที่รองรับระบบการประมวลผลโครงข่ายประสาทเทียม ทำการเทรนนิ่งอัลกอริทึมของโมเดลโดยแบ่งประเภทของตำหนิวัตถุชิ้นงานทั้ง 4 ประเภทได้แก่ 1) รอยตำหนิขีดข่วน 2) รอยตำหนิสีลอก 3) รอยตำหนิกระแทก และ 4) รอยตำหนิวัตถุบิดเบี้ยว รวมถึงประยุกต์ใช้ไมโครคอนโทรลเลอร์รับคำสั่งเพื่อส่งสัญญาณแจ้งถึงความผิดปกติ

1.2 วัตถุประสงค์ของงานวิจัย

1.2.1 เพื่อศึกษาอัลกอริทึมการประมวลผลรอยตำหนิวัตถุ นำมาใช้เพิ่มประสิทธิภาพกระบวนการตรวจสอบคุณภาพในอุตสาหกรรมยานยนต์

1.2.2 เพื่อพัฒนาอุปกรณ์ประมวลผลภาพประยุกต์ใช้ร่วมกับปัญญาประดิษฐ์การเรียนรู้เชิงลึก โดยสามารถระบุรอยตำหนิวัตถุแบบเรียลไทม์อัตโนมัติ

1.2.3 เพื่อพัฒนาอุปกรณ์ที่สามารถแยกแยะรอยตำหนิวัตถุแทนการมองเห็นมนุษย์

1.3 สมมติฐานของการวิจัย

การประมวลผลภาพของวัตถุ มีสภาพแวดล้อมแหล่งกำเนิดแสงที่แตกต่างกัน จึงทำให้ค่าความสว่างไม่คงที่ รวมถึงแสงสะท้อนกับชิ้นงานมีผลต่อการประมวลผลภาพ

1.4 ขอบเขตของการวิจัย

1.4.1 วิเคราะห์ตรวจสอบค่าความสว่างในกระบวนการตรวจสอบคุณภาพชิ้นงาน โดยกำหนดค่าความสว่างจากแหล่งกำเนิดแสงตามมาตรฐาน 1,800 ลักซ์

1.4.2 กำหนดตัวอย่างวัตถุเป็นแม่แรงรถยนต์ ขนาด 16 x 11.8 x 21.5 เซนติเมตร น้ำหนัก 3.46 กิโลกรัม พื้นผิววัตถุสีดำ เพื่อใช้วัดประสิทธิภาพของอัลกอริทึมโมเดล

1.4.3 ศึกษาปัญหาสาเหตุการเกิดรอยตำหนิวัตถุโดยใช้หลักการวิเคราะห์ด้วยเครื่องมือคุณภาพ 7 ชนิด ของทฤษฎี 7 QC Tools

1.4.4 กำหนดตำแหน่งที่ติดตั้งกล้องประมวลผลภาพห่างจากวัตถุ 25 เซนติเมตร จับภาพตรงตำแหน่งจากศูนย์กลางของวัตถุ

1.4.5 กำหนดลักษณะรูปแบบของตัวอย่างที่ทำการเทรนนิ่งอัลกอริทึมโมเดล โดยแบ่งออกเป็น 4 ประเภท ได้แก่ รอยขีดข่วน รอยบุบ รอยสีลอก และรอยแตกของผิวโลหะ

1.4.6 กำหนดอุปกรณ์ประมวลผลโดยใช้อุปกรณ์ NVIDIA Jetson Nano Developer Device ใช้ แหล่งจ่ายไฟ 5 โวลต์ กระแส 4 แอมป์ และใช้อุปกรณ์กล้อง Logitech ความละเอียด Full HD 1080 พิกเซล

1.5 ระเบียบวิธีวิจัย

- 1.5.1 รวบรวมข้อมูลวรรณกรรมงานวิจัยที่เกี่ยวข้อง
- 1.5.2 เก็บข้อมูลตัวอย่างปัญหาารอยตำหนิ
- 1.5.3 วิเคราะห์ทดสอบอัลกอริทึมที่เกี่ยวข้องงานวิจัย
- 1.5.4 ทำการเทรนนิ่งข้อมูลเพื่อสร้างอุปกรณ์โมเดลต้นแบบ
- 1.5.5 ทดสอบประสิทธิภาพของอุปกรณ์โมเดลต้นแบบกับวัตถุจริง
- 1.5.6 รวบรวมข้อมูลและสรุปผลการทดลอง
- 1.5.7 จัดทำวิทยานิพนธ์
- 1.5.8 สอบวิทยานิพนธ์

1.6 ประโยชน์ที่คาดว่าจะได้รับ

1.6.1 ระบบกล้องประมวลผลภาพที่สามารถตรวจสอบรอยตำหนิวัตถุ

1.6.2 อัลกอริทึมการประยุกต์ใช้โครงข่ายประสาทเทียมแบบสังวัตนาการการเรียนรู้เชิงลึก
ปัญญาประดิษฐ์สำหรับแยกแยะรอยตำหนิวัตถุ

1.6.3 อุปกรณ์โมเดลต้นแบบที่สามารถตรวจสอบและแยกแยะรอยตำหนิวัตถุแสดงผลแบบ
เรียลไทม์อัตโนมัติ



บทที่ 2

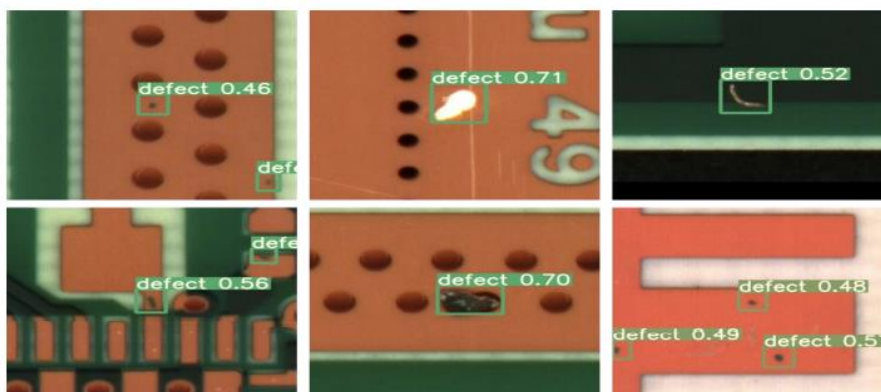
ปริทัศน์วรรณกรรมและทฤษฎีที่เกี่ยวข้อง

2.1 กล่าวนำ

ในบทนี้จะเป็นการกล่าวถึงปริทัศน์วรรณกรรมและทฤษฎีที่เกี่ยวข้องกับการประมวลผลภาพ การรับข้อมูลจากเซ็นเซอร์ เครื่องมือพื้นฐานสำหรับการประมวลผลภาพ การหาขอบภาพเป็นต้น รวมถึงความเป็นมาของปัญญาประดิษฐ์ การเรียนรู้เชิงลึกและทฤษฎีโครงข่ายประสาทเทียม ที่ทางผู้วิจัยนำมาประยุกต์ใช้กับงานวิจัยฉบับนี้

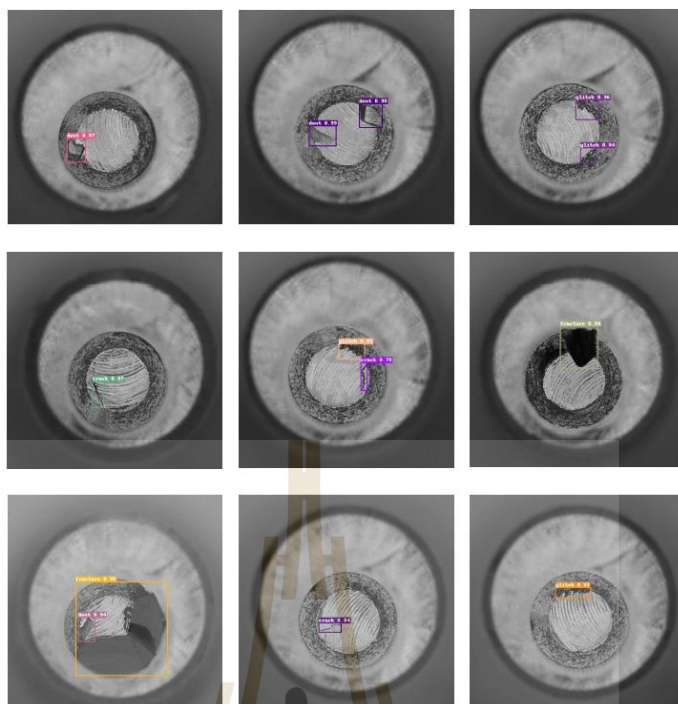
2.2 ปริทัศน์วรรณกรรม

Venkart Anil Adibhatla และคณะ (2021) ทำการศึกษาการเรียนรู้เชิงลึกเพื่อตรวจจับรอยตำหนิแผ่นวงจรอิเล็กทรอนิกส์โดยใช้โมเดลอัลกอริทึม YOLO (You Look Only Once) ที่มาของงานวิจัยคือ ในปัจจุบันมีอัลกอริทึมที่ถูกพัฒนาขึ้นหลากหลายแต่ยังมีข้อจำกัดในเรื่องความเร็วและความแม่นยำสำหรับการตรวจจับรอยตำหนิแผ่นวงจรอิเล็กทรอนิกส์ เพื่อแก้ปัญหาส่วนนี้จึงทดสอบโดยทดลองอัลกอริทึมของ YOLO ขั้นตอนการวิจัยมีดังต่อไปนี้ เก็บรวบรวมข้อมูลตัวอย่างรอยตำหนิ 23,000 รูป ขนาดของภาพ 400 x 400 พิกเซล แบ่งข้อมูลออกเป็น 2 ส่วน ส่วนแรก 20,700 ภาพ นำข้อมูลส่วนนี้ทำการเทรนนิ่งโมเดล และ 2,300 ภาพ นำไปใช้ทดสอบโมเดล ผลการวิจัยโมเดลมีความแม่นยำ 99.74 เปอร์เซ็นต์



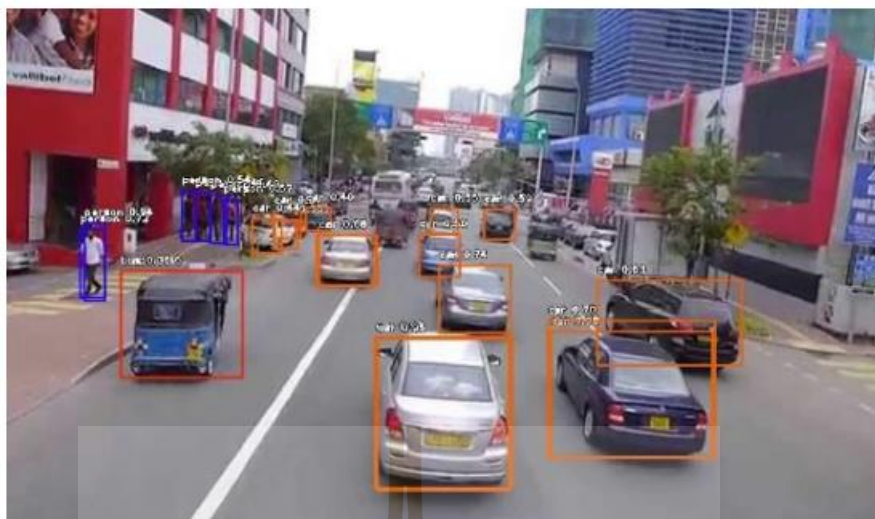
รูปที่ 2.1 ตัวอย่างรูปภาพของแผ่นวงจรอิเล็กทรอนิกส์
(ที่มา: Venkat Anil Adibhatla, 2021)

Yiting Li และคณะ (2021) ทำการศึกษาการตรวจสอบคุณภาพผลิตภัณฑ์แบบเรียลไทม์โดยประยุกต์ใช้อัลกอริทึม Multi Scale PP-YOLO เพื่อเพิ่มประสิทธิภาพในการตรวจสอบของกระบวนการผลิตและลดค่าใช้จ่ายที่สูญเปล่า ดำเนินการวิจัยโดยทำการเก็บรวบรวมข้อมูล 1,200 ภาพ ได้แก่ รอยยุบ รอยร้าว รอยทั่วไป และรอยแตก ทั้งหมดเป็นภาพขาวดำ (Gray scale) กำหนดขนาดของภาพให้เล็กลงเป็น 416×416 พิกเซล เพื่อเพิ่มความเร็วของการประมวลผลโมเดลและลดปัญหาการเกิดข้อผิดพลาดขณะทำการเทรนนิ่ง นอกจากนี้ทำการเพิ่มอภิพันธุ์ข้อมูล 5 รูปแบบคือ 1) การหมุนทวนเข็มนาฬิกา 20 องศา 2) ขยับภาพไปทางด้านขวา 10 เปอร์เซ็นต์ 3) ชุม 10 เปอร์เซ็นต์ 4) พลิกแกนแนวตั้ง และ 5) พลิกแกนแนวนอน ตั้งค่าพารามิเตอร์เทรนนิ่งข้อมูล 35 epochs ผลลัพธ์ประสิทธิภาพโมเดลจากการคำนวณ Confusion Matrix ของทั้ง 4 คลาส อยู่ที่ 97.68 เปอร์เซ็นต์



รูปที่ 2.2 การตรวจจ็บบรอยตำหนิ (ที่มา: Yiting Li และคณะ, 2021)

Dr. S.V. Virkatmath และคณะ (2021) ทำการศึกษาการตรวจจ็บบวัตถุและการแยกประเภท โดยใช้อัลกอริทึม YOLOv3 (You Look Only Once version 3) วัตถุประสงค์งานวิจัย คือการทดสอบอัลกอริทึมเพื่อตรวจจ็บบานพาทนะและการแยกแยะวัตถุอื่น ๆ โดยใช้ฐานข้อมูล COCO Dataset ผลลัพธ์การทดสอบอัลกอริทึมสามารถตรวจจ็บบานพาทนะและแยกแยะสิ่งของได้ตามวัตถุประสงค์ที่กำหนดไว้ สามารถนำไปประยุกต์ใช้กับงานอื่น ๆ ได้หลากหลาย



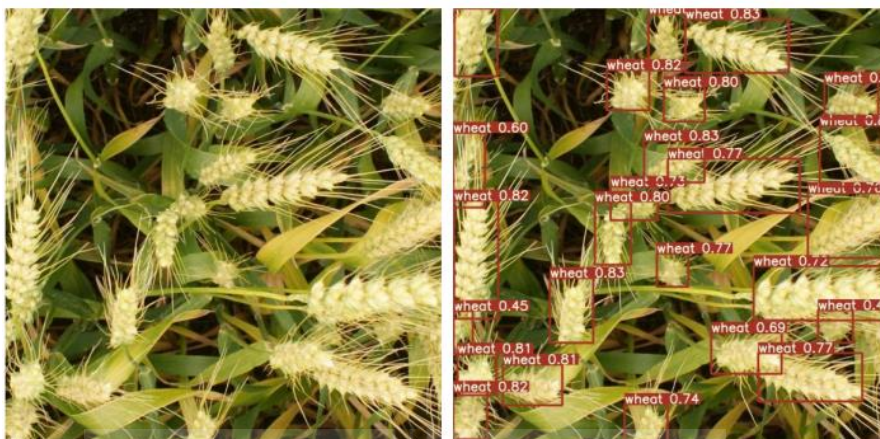
รูปที่ 2.3 การตรวจจับยานพาหนะ (ที่มา: Dr. S.V. Virkatmath และคณะ, 2021)

Wentao Liu และคณะ (2021) ทำการศึกษาการตรวจจับสัญญาณเตือนรถไฟแบบเรียลไทม์ โดยใช้อัลกอริทึม YOLOv5 (You Look Only Once version 5) เพื่อช่วยการสังเกตการณ์ของพนักงานขับรถไฟเมื่อวิ่งผ่านตรงจุดที่มีสัญญาณเตือน ขั้นตอนการดำเนินการวิจัยทำการศึกษาสภาพแวดล้อมและเก็บตัวอย่างข้อมูลแบ่งเป็น 4 รูปแบบ ได้แก่ 1) ไฟแดง จำนวน 2,460 ภาพ 2) สัญญาณเตือนไฟแดง จำนวน 3,409 ภาพ 3) ไฟเขียว จำนวน 361 ภาพ และ 4) สัญญาณเตือนไฟเขียว จำนวน 676 ภาพ โดยจะตัดภาพที่เป็นพื้นหลังออกเพื่อการทำ Label กำหนดขนาดภาพ 1,280 x 720 พิกเซล ตั้งค่าพารามิเตอร์ 300 epochs ใช้ไลบรารี python3.8 และ pytorch 1.6 ผลการศึกษา YOLOv5 มีค่าความถูกต้องในการแยกแยะสัญญาณไฟ 97.2 เปอร์เซ็นต์



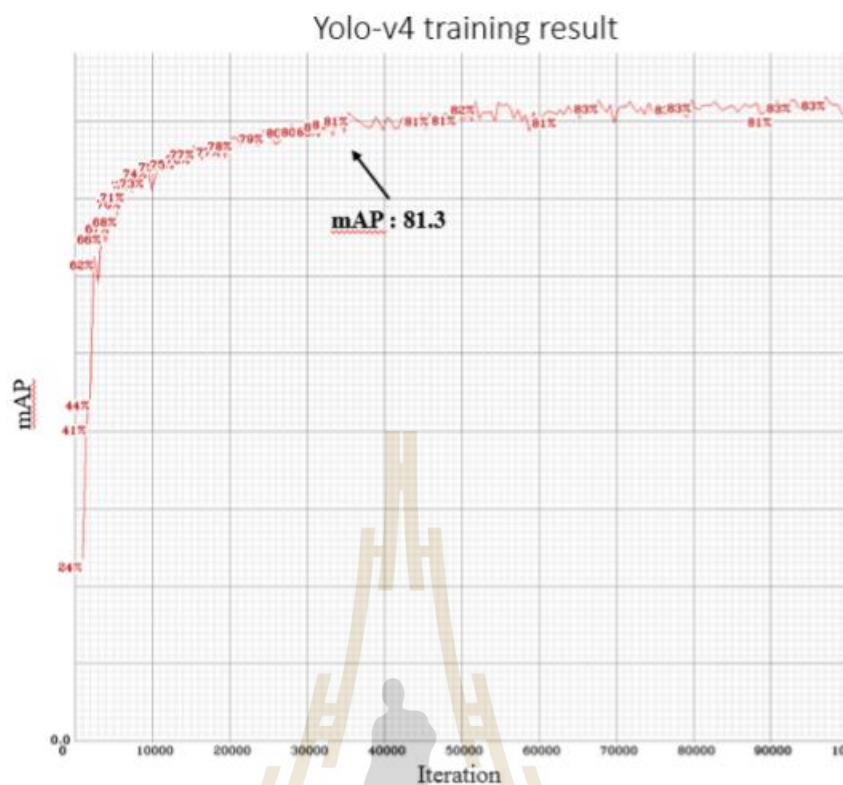
รูปที่ 2.4 ตัวอย่างข้อมูลภาพสัญญาณไฟรถไฟ (ที่มา: Wentao Liu และคณะ, 2021)

Do Thuan (2021) ทำการศึกษาประเมินประสิทธิภาพอัลกอริทึมYOLOv4 (You Look Only Once version 4) และ YOLOv5 (You Look Only Once version 5) โดยใช้เครื่องมือ Google Collaboratory และการประมวลผลแบบ Graphic Process Unit มีไลบรารี เช่น PyTorch, Tensorflow, Keras และ OpenCV ฟังก์ชันอยู่ในโปรแกรม ใช้ Global Wheat Head Detection Datasetสำหรับการเทรนนิ่งโมเดล กำหนดค่าพารามิเตอร์ 416 x 1416 พิกเซล 86 batch size และ 100 epochs สรุปผลการทดลอง YOLOv5 ที่ได้รับการพัฒนาเป็นเวอร์ชันล่าสุดมีความเร็วและความแม่นยำเท่ากับ YOLOv4 และมีจุดเด่นในเรื่องของเฟรมเวิร์ค Pytorch ที่ใช้งานง่ายขึ้นกว่าเวอร์ชันเดิม



รูปที่ 2.5 การเทรนนิ่งโมเดลจาก Global Wheat Head Detection Dataset
(ที่มา: Do Thuan, 2021)

OMAR MOURED (2020) ทำการศึกษาการเรียนรู้เชิงลึกสำหรับการตรวจจับวัตถุที่หลากหลาย เพื่อประเมินประสิทธิภาพของอัลกอริทึม YOLOv4 (You Look Only Once version 4) อุปกรณ์ที่เลือกใช้เป็นบอร์ด Nvidia RTX 2080 GPU ใช้ข้อมูล REID สำหรับเทรนนิ่ง กำหนดค่าพารามิเตอร์ขนาดเลเยอร์สูงสุด 250 เลเยอร์ ค่าiteration 35,000 ครั้ง จากผลการทดลองโมเดลสามารถตรวจจับวัตถุตามวัตถุประสงค์ที่กำหนดไว้ไม่เกิดปัญหา Overfitting ผลลัพธ์ค่า mAP อยู่ที่ 81.3 เปอร์เซ็นต์



รูปที่ 2.6 ประสิทธิภาพโมเดลหลังจากทำการเทรนนิ่ง (ที่มา: OMAR MOURED, 2020)

2.3 การประมวลผลภาพ

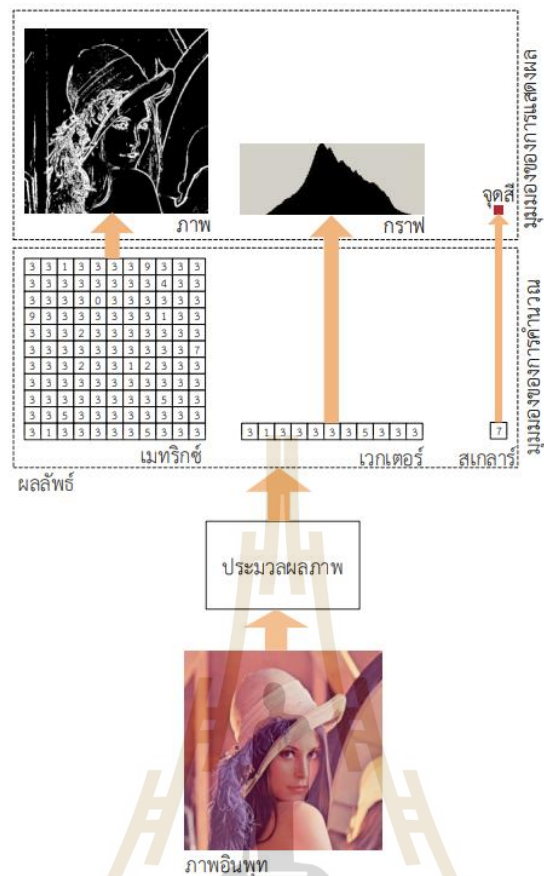
การประมวลผลภาพคือ วิธีการหรืออัลกอริทึมในการนำภาพมาประมวลผลแล้วส่งผลลัพธ์ออกเป็นอีกภาพ ซึ่งประกอบด้วย

- 1) การแสดงผล เช่น การแสดงผลออกที่จอ CRT/LCD การพิมพ์ออกที่เครื่องพิมพ์แบบเลเซอร์ หรือแบบหมึกพิมพ์
- 2) การเปลี่ยนแปลงและการแก้ไขภาพ เช่นการตกแต่งภาพ การตัดต่อแก้ไขภาพ
- 3) การปรับปรุงภาพให้ดีขึ้น เช่น การปรับความสว่าง การทำให้ภาพคมชัด การทำให้ภาพมีรายละเอียด มากขึ้นกว่าเดิม
- 4) การค้นหาลักษณะหรือลักษณะเด่นภายในภาพ เช่น การค้นหาขอบภาพ การค้นหามุมของวัตถุ การสกัดเอาคุณลักษณะของพื้นผิวของภาพ
- 5) การบีบอัดภาพ เช่นการบีบอัดข้อมูลภาพแบบ JPG, J2K ซึ่งวิธีการเหล่านี้ถูกสร้างขึ้นมา เพื่อให้ภาพที่นำมาประมวลผลมีคุณภาพดีขึ้น หรือแย่ง เช่นการทำ ภาพถ่ายให้คมชัดหรือสว่างมากขึ้นกว่าเดิม หรือการทำให้บางส่วนของภาพ (ขวดเหล้าหรือบุหรี) เสียหายหรือพรมัว หากภาพ

นั้นไม่สามารถฉายในโทรศัพท์นี้ได้ การทำให้ไฟล์ภาพมีขนาดเล็กลง การสกัด (Extract) เอาลักษณะเด่นหรือองค์ประกอบของภาพออกมา เช่นขอบภาพ มุมของวัตถุ เพื่อใช้ในการรู้จำต่อไป

2.3.1 ภาพดิจิทัล

ภาพดิจิทัลเกิดจากการเปลี่ยนภาพที่ได้จากการสะท้อนแสง (Light Reflection) ลงบน เซ็นเซอร์รับภาพ (Image Sensor) ซึ่งเป็นอุปกรณ์ที่ไวต่อการรับแสง (Photo Sensitive Devices) เช่น CCD/CMOS ให้เป็นข้อมูลภาพทางด้านดิจิทัล ซึ่งภาพที่เรามองเห็นบนจอภาพหรือบนกระดาษ เกิดจากการผสมสีเข้าด้วยกัน เช่นจอภาพใช้การผสมสีจากโมเดลสี RGB ซึ่งเป็นการผสมสีแดง เขียว และน้ำเงิน เข้าด้วยกัน เพื่อให้ได้สีตามที่ต้องการ เครื่องพิมพ์แบบฉีดหมึก (Ink Jet) ใช้การผสมสีจาก โมเดลสี CMYK โดยอาศัยหลักการดูดซับแสง (Light Absorption) ซึ่งจุดภาพแต่ละจุด ถูกเรียกว่า พิกเซล (Pixel) โดยเมื่อนำพิกเซลมาเรียงกันเป็นรูปทรงสี่เหลี่ยม (หรือรูปทรงใด ๆ) ก็จะทำให้เกิดเป็น ภาพขึ้นมา แสดงว่าภาพเกิดจากการนำพิกเซลมาเรียงต่อกันเพื่อให้เกิดเป็นภาพ ดังนั้นคุณภาพของ ภาพที่สร้างขึ้นมา จึงขึ้นอยู่กับคุณภาพของแต่ละพิกเซล การประมวลผลภาพเป็นการนำตัวเลขในแต่ละพิกเซลมาวิเคราะห์และประมวลผล โดยอาจมีการวิเคราะห์เพียงแค่จุดเดียว หรือการวิเคราะห์เป็นกลุ่มพิกเซล หรือการวิเคราะห์ทั้งภาพ ซึ่งผลลัพธ์ที่ได้จากการประมวลผล หากมองในมุมมองของการคำนวณ ผลลัพธ์อาจเป็นเมทริกซ์ (Matrix) เวกเตอร์ (Vector) หรือสเกลาร์ (Scalar) ในมุมมองของการแสดงผล ผลลัพธ์ที่ได้ อาจเป็นภาพอีกหนึ่งภาพ อาจเป็นกราฟ หรืออาจเป็นค่าสีก็ได้ ทั้งนี้ ขึ้นอยู่กับวิธีการที่ใช้ในการคำนวณและการแสดงผล รูปที่ 2.7 แสดงตัวอย่างผลลัพธ์จากการประมวลผล ภาพในรูปแบบ เมทริกซ์= ภาพ เวกเตอร์= กราฟ และ สเกลาร์= ค่าสี



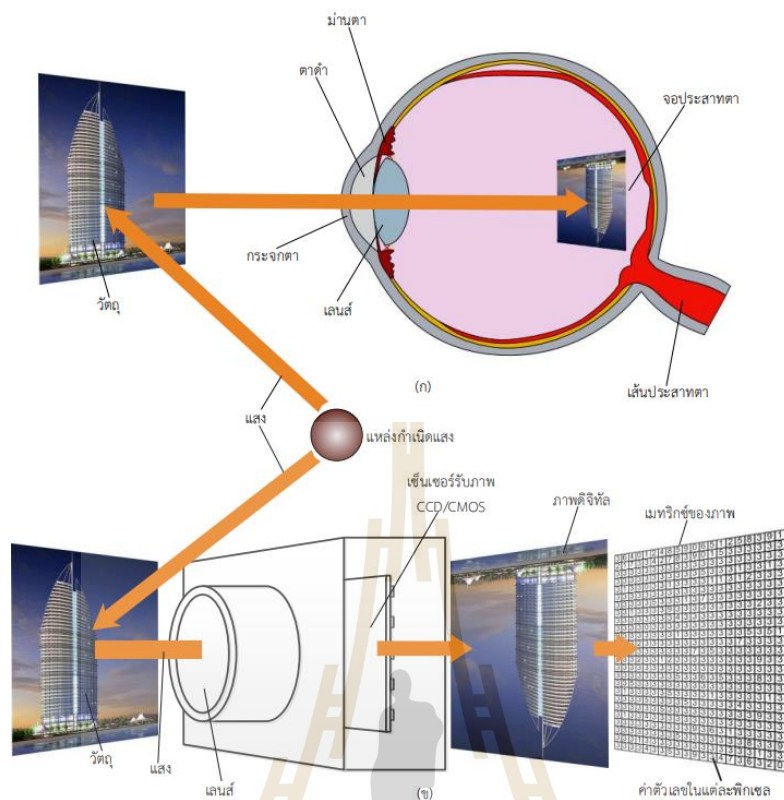
รูปที่ 2.7 การประมวลผลภาพและผลลัพธ์ในรูปแบบต่าง ๆ

(ที่มา: รองศาสตราจารย์ ดร.สนั่น ศรีสุข หนังสือการประมวลผลภาพขั้นสูง)

ภาพดิจิทัลที่สร้างขึ้นมาจากกล้องดิจิทัล มีการทำงานคล้ายตามนุษย์ ตัวอย่างการทำงานและส่วนประกอบต่าง ๆ ของตามนุษย์แสดงดังรูปที่ (รูปที่ 2.8 (ก)) เมื่อมีแสงจากแหล่งกำเนิดแสงมาตกกระทบบที่วัตถุ วัตถุนั้น จะสะท้อนแสงออกมา แสงจะผ่านไปยังกระจกตา (Cornea) ซึ่งกระจกตาจะทำหน้าที่ปกป้องตาโดยเคลือบผิวด้านนอกไว้ ซึ่งทำให้เรามองเห็นภาพได้และยังได้รับการปกป้องในเวลาเดียวกันด้วย นอกจากนั้นกระจกตาสามารถปรับรูปร่างได้ เพื่อให้ตาสามารถโฟกัสภาพได้ดีที่สุด จากนั้น แสงจะผ่านไปยังตาดำ (Pupil) โดยตาดำทำหน้าที่ในการควบคุมแสงที่ส่งผ่านเข้าไปยัง ลูกตา หากแสงมีปริมาณมาก ตาดำจะหดเล็กลงเพื่อให้แสงเข้าน้อยลง ในทางตรงข้าม หากแสงมีปริมาณน้อย ตาดำจะขยายมากขึ้น เพื่อให้แสงผ่านเข้าได้มากขึ้น กล้ามเนื้อที่ทำหน้าที่ในการหดขยายตาดำ คือ ม่านตา (Iris) ซึ่งเป็นกล้ามเนื้อที่อยู่ล้อมรอบตาดำ โดยกล้ามเนื้อเหล่านี้ สามารถเปลี่ยนรูปได้ เมื่อกล้ามเนื้อม่านตาเกิดการเปลี่ยนรูป จะทำให้ตาดำเกิดการหดขยาย กล้ามเนื้อม่านตาของคนมีเอกลักษณ์เฉพาะตัวที่ไม่ซ้ำกัน สามารถใช้ในการระบุเอกลักษณ์ของแต่ละบุคคลได้ แสงที่ผ่านตาดำจะไปตกกระทบบที่เลนส์ (Lens) ซึ่งเลนส์จะทำหน้าที่ในการรวมแสงให้ตกกระทบบไปยังจอประสาทตา

(Retina) จอประสาทตาเป็นส่วนที่ไวต่อแสง และทำหน้าที่ในการเปลี่ยนพลังงานแสง เป็นพลังงานทางเคมี และพลังงานทางไฟฟ้า ซึ่งพลังงานนี้จะถูกส่งต่อไปยังสมอง เพื่อให้สมองแปลความหมายของภาพที่เรามองเห็นต่อไป อย่างไรก็ตาม ภาพที่เรามองเห็น ไม่ได้อยู่ในรูปแบบการแสดงผลบนจอภาพ หรือการพิมพ์ออกที่เครื่องพิมพ์ เป็นภาพที่เกิดจากจินตนาการที่สร้างขึ้นมาโดยสมองมนุษย์ ดังนั้นภาพที่เกิดจากมโนภาพนี้จึงไม่มีหน่วยวัดความละเอียด ไม่มีขนาดความกว้างความสูง ไม่มีพิกเซล

กล้องดิจิทัล (รูปที่ 2.8 (ข)) มีหลักการทำงานคล้ายกันกับตาของมนุษย์ โดยเมื่อมีแสงจากแหล่งกำเนิดแสงมาตกกระทบบที่วัตถุ แสงนั้นจะสะท้อนออกมาจากวัตถุ ผ่านเข้าไปยังเลนส์ โดยเลนส์จะทำหน้าที่ในการรวมแสงให้ตกกระทบบยังเซ็นเซอร์รับภาพ ซึ่งในปัจจุบันมีอุปกรณ์ที่ไวต่อแสงอยู่สองชนิด คือ CCD และ CMOS อุปกรณ์ทั้งสองชนิดนี้ทำหน้าที่ในการเปลี่ยนพลังงานแสงไปเป็นพลังงานไฟฟ้า โดยเซ็นเซอร์รับภาพจะประกอบด้วยจุดพิกเซล ที่จะให้แรงดันไฟฟ้าออกมา โดยแรงดันไฟฟ้าจะเปลี่ยนไปตามแสงที่มาตกกระทบบ เมื่อนำค่าแรงดันไฟฟ้าไปแปลความหมายแล้วแสดงผล ก็จะเป็นภาพดิจิทัลขึ้นมา เนื่องจากภาพเกิดจากการนำจุดพิกเซล มาเรียงต่อกันเพื่อสร้างขึ้นมาเป็นภาพ ดังนั้นความละเอียดของภาพดิจิทัลจึงขึ้นอยู่กับคุณภาพของแต่ละพิกเซล อย่างไรก็ตาม ค่าแรงดันไฟฟ้าในแต่ละจุดพิกเซลจะถูกเปลี่ยนไปเป็นตัวเลขทางด้านดิจิทัล เมื่อนำตัวเลขมาเรียงต่อกันให้เป็นที่เรียบร้อย ผลลัพธ์ที่ได้ก็จะเป็นเมทริกซ์นั่นเอง ดังนั้นการประมวลผลภาพจึงเป็นการคำนวณบนตัวเลขของแต่ละพิกเซลของภาพที่ถูกจัดเก็บในรูปแบบเมทริกซ์นั่นเอง (การจัดเก็บในรูปแบบเมทริกซ์นี้เป็นการจัดเก็บในทางตรรกะ หากมองถึงการจัดเก็บทางกายภาพแล้ว การจัดเก็บข้อมูลจะเป็นเวกเตอร์ หรือสายของข้อมูล เนื่องจากข้อจำกัดทางกายภาพของอุปกรณ์ในการบันทึกข้อมูล)



รูปที่ 2.8 โครงสร้างการทำงานของ (ก) ตามมนุษย์ และ (ข) กล้องดิจิทัล

(ที่มา: Creative Commons http://en.wikipedia.org/wiki/File:Human_eye_cross_section_detached_retina.svg)

ภาพที่สามารถมองเห็นจะเป็นภาพที่กลับหัว เมื่อแสงมาตกกระทบ ที่จอประสาทตาของมนุษย์ หรือเซ็นเซอร์รับภาพ ดังนั้นภาพดิจิทัลที่ถูกจัดเก็บลงในไฟล์ จึงเป็นภาพ ที่กลับหัวด้วย ในการแสดงผลที่จอภาพ คำสั่งมาตรฐานที่แสดงภาพ จะทำการกลับภาพให้เอง แต่ถ้าเขียนโปรแกรมแสดงภาพเองที่ละพิกเซล จะได้ภาพกลับหัว เพราะฉะนั้น ในการเขียนโปรแกรม ทางด้านการประมวลผลภาพ ภาพทำการประมวลผลอยู่กลายเป็นภาพกลับหัว ถ้าอ่านข้อมูลจากบนลงล่าง จะเป็นการประมวลผลภาพแบบกลับหัว ถ้าอ่านข้อมูลจากล่างขึ้นบน จึงจะเป็นการประมวลผลภาพแบบไม่กลับหัว

หน้าที่ ของเซ็นเซอร์รับภาพคือ ทำหน้าที่ในการเปลี่ยนโฟตอน (Photon) ไปเป็น อิเล็กตรอน หรืออีกอย่างหนึ่งคือ เซ็นเซอร์รับภาพทำหน้าที่ในการเปลี่ยนพลังงานแสงไปเป็น พลังงานไฟฟ้าเนื่องจากแสงมีคุณสมบัติพิเศษ ที่เป็นทั้งอนุภาค (Particle) และคลื่นแม่เหล็กไฟฟ้า

(Electromagnetic Radiation) อนุภาคของแสงคือโฟตอน เมื่อแสงมาตกกระทบที่เซ็นเซอร์รับภาพ โฟตอนบางส่วนก็จะถูกดูดซับโดยเซ็นเซอร์รับภาพ ทำให้พลังงานจากโฟตอนถ่ายเทไปยังเซ็นเซอร์รับภาพ ซึ่งโดยทั่วไปจะเป็นอุปกรณ์ที่ทำจากสารกึ่งตัวนำ (Semiconductor) พลังงานของโฟตอนนี้ จะทำให้อิเล็กตรอนเกิดการเคลื่อนที่ทำให้เกิดกระแสไฟฟ้าไหลผ่านสารกึ่งตัวนำ ผลลัพธ์ที่ออกจากเซ็นเซอร์รับภาพคือแรงดันไฟฟ้า แรงดันไฟฟ้านี้จะเปลี่ยนไปตามแสงที่มาตกกระทบที่เซ็นเซอร์รับภาพ เนื่องจากสีต่าง ๆ เช่น น้ำเงิน หรือเหลือง จะมีจำนวนโฟตอนที่ไม่เท่ากัน ซึ่งวัดโดยค่า พลังงานโฟตอน ดังนั้น เมื่อโฟตอนของแต่ละสีไปกระทบที่อิเล็กตรอนในเซ็นเซอร์รับภาพ ก็จะได้แรงดันไฟฟ้าไม่เท่ากัน แรงดันไฟฟ้านี้จะถูกเปลี่ยนให้เป็นตัวเลขทางดิจิทัล

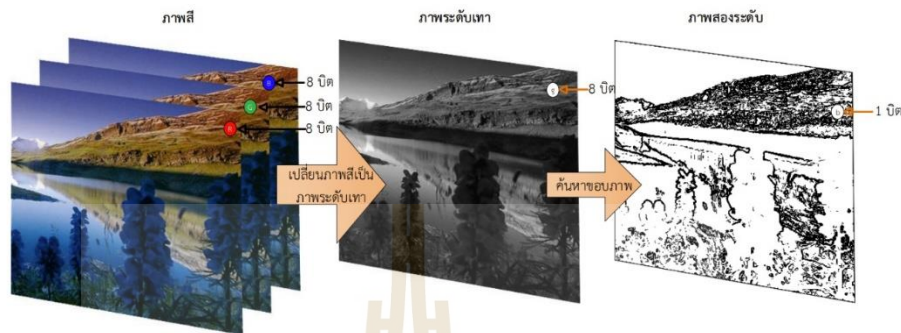
ภาพที่ถูกนำมาคำนวณด้วยการประมวลผลภาพมีหลายรูปแบบ ซึ่งโดยทั่วไปที่พบมีได้สามแบบ คือภาพสี (Color Image) ภาพระดับเทา (Gray Level Image) และภาพสองระดับ (Binary Image)

ภาพสีคือภาพในแต่ละพิกเซลที่ประกอบด้วยแม่สี แดง เขียว และน้ำเงิน แต่ละแม่สีจะใช้พื้นที่ในการ จัดเก็บ 8 บิต ดังนั้นแต่ละพิกเซลของภาพสีจึงใช้พื้นที่ในการจัดเก็บ 24 บิต (3 ไบต์) และสามารถ แสดงสีได้ประมาณ 2²⁴ ล้านสี(16,777,215 ล้านสี)

ภาพระดับเทา คือภาพที่ไม่มีสี เป็นภาพที่แสดงถึงระดับความเข้มของแสง ซึ่งแต่ละพิกเซลจะประกอบด้วยตัวเลขเพียงหนึ่งตัว โดยทั่วไป แต่ละพิกเซลจะมีขนาด 8 บิต (1 ไบต์) และสามารถแสดงระดับความเข้มของแสงได้ 256 ระดับ จากมืด (สีดำ) ซึ่งมีค่าเป็น 0 ไปจนถึงสว่าง (สีขาว) ซึ่งมีค่าเป็น 255 ภาพระดับเทาเป็นรูปแบบภาพที่ถูกนำมาใช้ในการประมวลผลภาพมากที่สุด เนื่องจากมีขนาดเล็กเมื่อเทียบกับภาพสี และยังสามารถเก็บรายละเอียดของพื้นผิว (Texture) ของวัตถุในภาพได้เป็นอย่างดี ซึ่งพื้นผิวของวัตถุภายในภาพ เป็นรายละเอียดที่จำเป็นต่อการวิเคราะห์ภาพ

ภาพสองระดับ คือภาพขาวดำ ซึ่งแต่ละพิกเซลจะมีเพียง 2 รหัส คือ ขาว และ ดำ โดยแทนเป็นรหัส 0 หรือ 1 ภาพสองระดับก็ถูกนำมาใช้ในการวิเคราะห์ภาพเช่นเดียวกัน แต่เนื่องจากมีแค่สองระดับ ดังนั้น รายละเอียดของภาพจึงหายไป มีเพียงแต่ภาพโครงร่างของวัตถุ เท่านั้น อย่างไรก็ตาม ภาพสอง ระดับอาจเป็นผลลัพธ์ที่เกิดจากการทำการประมวลผลภาพก็ได้ เช่นการหาขอบภาพ และการแยก วัตถุออกจากพื้นหลัง เป็นต้น

ทั้งนี้ ภาพระดับเทาอาจเกิดจากการเปลี่ยนจากภาพสี และภาพสองระดับอาจเกิดจากการนำภาพระดับเทามาสกัดเอาลักษณะเด่น จนได้ผลลัพธ์เป็นภาพสองระดับ ก็ได้ รูปที่ 2.9 แสดงตัวอย่างการเปลี่ยนภาพสีเป็นภาพระดับเทา และการเปลี่ยนจากภาพระดับเทาเป็นภาพสองระดับ



รูปที่ 2.9 การเปลี่ยนจากภาพสีเป็นภาพระดับเทาโดยการแปลงค่าสี และการเปลี่ยนจากภาพ ระดับเทาไปเป็นภาพสองระดับโดยการค้นหาขอบภาพ (ภาพทิวทัศน์ ลิขสิทธิ์: สมบัติสาธารณะ แหล่งที่มา <http://en.wikipedia.org/wiki/File:Bachalpseefflowers.jpg>)

2.3.2 การประมวลผลระดับพิกเซล

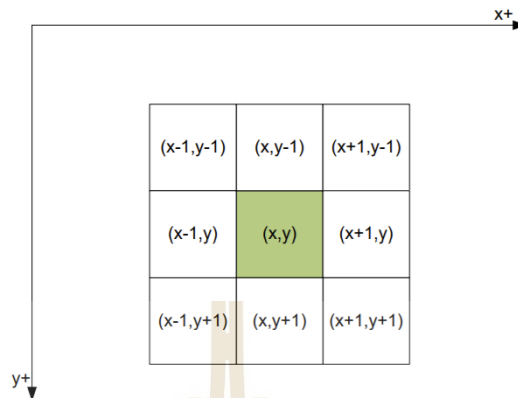
การประมวลผลภาพ $f(x, y)$ เป็นการนำค่าตัวเลขในระดับพิกเซลมาคำนวณ ซึ่งอาจเป็น การนำค่าตัวเลขที่ตำแหน่ง (x, y) เพียงตำแหน่งเดียวมาคำนวณ หรือมีการนำค่าพิกเซลที่อยู่รอบ ๆ มาคำนวณร่วมด้วย เพื่อคำนวณหาลักษณะเฉพาะของแบบรูป ณ ตำแหน่ง นั้น อย่างไรก็ตาม ภาพที่ เรานำมาประมวลผล นั้น อยู่ในพิกัดจอภาพ (Screen Coordinate) ซึ่งมีจุดเริ่มต้นที่ตำแหน่งบนซ้าย (ซึ่งตรงกันข้ามกับพิกัดคาร์ทีเซียน ที่มีจุดเริ่มต้นที่ตำแหน่งล่างซ้าย) ดังนั้น สำหรับพิกัดจอภาพ แกน x เลื่อนไปทางขวาเป็นบวก เลื่อนไปทางซ้ายเป็นลบ แกน y เลื่อนลงล่างเป็นบวก เลื่อนขึ้นบนเป็นลบ ดังนั้น หากเราอยู่ที่ตำแหน่ง (x, y) ตำแหน่งทางด้านแนวนอนและแนวตั้ง สามารถเขียนได้ดังนี้

$$(x + 1, y), (x - 1, y), (x, y + 1), (x, y - 1) \quad (2.1)$$

ซึ่งเป็นการประมวลผลพิกเซลแบบ 4 จุดใกล้เคียง (4-Neighbors) ที่อยู่รอบๆ ตำแหน่ง (x, y) และนิยามได้เป็น $N_4(x, y)$ นอกจากนี้ พิกเซลที่อยู่รอบ ๆ ในแนวทะแยงสามารถเขียนได้ดังนี้

ซึ่งนิยามเป็น $ND(x, y)$ ดังนั้น เมื่อรวม $N_4(x, y)$ และ $ND(x, y)$ จะได้เป็น $N_8(x, y)$

$$(x + 1, y + 1), (x + 1, y - 1), (x - 1, y + 1), (x - 1, y - 1) \quad (2.2)$$



รูปที่ 2.10 ความสัมพันธ์ระหว่างพิกเซลรอบจุด (x, y)
(ที่มา: รองศาสตราจารย์ ดร.สนั่น ศรีสุข หนังสือการประมวลผลภาพขั้นสูง)

2.3.3 การทำสังวัตนาการ (Convolution)

เทคนิคหนึ่งในการประมวลผลภาพที่สำคัญคือ การทำสังวัตนาการ ซึ่งสามารถอธิบายได้ ง่ายๆ ว่า เป็นการนำผลคูณจุด (Dot Product) ในแต่ละตำแหน่งของภาพ วิธีการดังกล่าวเป็นการ สร้างฟังก์ชัน g แล้วนำไปทำผลคูณจุดกับภาพ f ดังนั้นในแต่ละตำแหน่ง (x, y) สามารถเขียนสมการ ในการทำสังวัตนาการระหว่าง f กับ g ในรูปของฟังก์ชันต่อเนื่องได้ดังนี้

$$h(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x', y') g(x - x', y - y') dx' dy' \quad (2.3)$$

ซึ่งจะเห็นได้ว่าเป็นการนำค่าในแต่ละตำแหน่ง $f(x', y')$ ไปคูณกันกับ $g(x - x', y - y')$ แล้วนำไปหาผลรวม (การทำอินทิกรัลคือการหาผลรวม ซึ่งในที่นี้มีขอบเขตจาก $-\infty$ ถึง ∞) ผลลัพธ์ที่ได้เป็น $h(x, y)$ ดังนั้น แต่ละตำแหน่งจึงเป็นการทำผลคูณจุด เนื่องจากการนำสมาชิกของ f และ g แต่ละตัวมาคูณกัน จากนั้น นำทั้งหมดมารวมกัน เมื่อทำซ้ำทั้งภาพทุกตำแหน่ง (x, y) ก็จะได้ภาพอีกภาพ หนึ่ง ซึ่งเป็นผลลัพธ์ของการทำสังวัตนาการ และสามารถเขียนในอีกรูปหนึ่งได้

สัญลักษณ์ $*$ หมายถึงการทำสังวัตนาการ มีความหมายอีกนัยว่า การทำสังวัตนาการสามารถสลับได้ ฟังก์ชัน $g(x, y)$ เมื่อนำมาใช้ในการทำสังวัตนาการ จะถูกเรียกได้หลายชื่อ เช่น ตัวกรอง (Filter) หน้ากาก (Mask) แก่น (Kernel) แม่แบบ (Template) และหน้าต่าง (Window) โดยในหนังสือเล่มนี้ จะใช้คำว่า หน้ากาก เป็นส่วนใหญ่ ค่าในหน้าฉากเรียกว่าค่าสัมประสิทธิ์

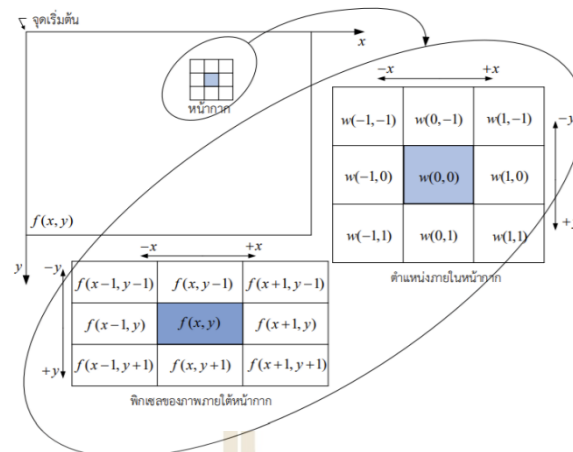
(Coefficient) โดยทั่วไป หน้ากาก $g(x, y)$ อาจมีขนาด 3×3 หรือ 5×5 หรือ 7×7 ตามแต่ผู้นำไปใช้งานเป็นผู้กำหนด ดังนั้น มันจึงเป็นการนำพิกเซลรอบ ๆ ตำแหน่ง (x, y) ไปประมวลผล การทำสังวัตนาการเป็นเทคนิคที่สำคัญมาก เนื่องจาก มันถูกนำไปใช้งานหลากหลาย เช่น การกรองสัญญาณรบกวน การหาขอบภาพ การทำให้ภาพคมชัด เป็นต้น

$$h(x, y) = f(x, y) * g(x, y) = g(x, y) * f(x, y) \quad (2.4)$$

ในกรณีของภาพดิจิทัล การทำสังวัตนาการ จะเป็นการทำผลคูณจุด ในแต่ละตำแหน่ง (x, y) แล้วทำการเลื่อนจุดไปยังจุดต่อไป จุดต่อจุด ผลลัพธ์ที่ได้แต่ละตำแหน่ง เรียกว่าเป็นผลตอบสนอง (Response) เนื่องจากแนวความคิดพื้นฐานในการทำสังวัตนาการ มาจากการทำการแปลงแบบฟูเรียร์ (Fourier Transform) ซึ่งอยู่ในสาขาการประมวลผลสัญญาณ ดังนั้น สำหรับฟังก์ชันไม่ต่อเนื่อง การทำสังวัตนาการ สามารถเขียนเป็นสมการได้ดังนี้

$$h(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x + s, y + t) \quad (2-5)$$

เมื่อ $a = (m - 1)/2$ และ $b = (n - 1)/2$ คือขอบเขตในการประมวลผลพิกเซลรอบ ๆ ตำแหน่ง (x, y) ตัวแปร m และ n คือขนาดของหน้ากาก มีหน่วยเป็นพิกเซล ตัวแปร w คือหน้ากาก ดังนั้น แต่ละตำแหน่ง (x, y) มันจึงเป็นการนำค่าสัมประสิทธิ์ $w(s, t)$ ไปคูณกันกับ $f(x + s, y + t)$ แล้ว นำไปหาผลรวม



รูปที่ 2.11 การทำสังวัตนาการระหว่าง f และ w

(ที่มา: รองศาสตราจารย์ ดร.สนั่น ศรีสุข หนังสือการประมวลผลภาพขั้นสูง)

รูปที่ 2.11 แสดงวิธีการในการทำสังวัตนาการ โดยการวางหน้ากาก w ไปบนภาพ f ตำแหน่งจุดกึ่งกลางที่กำลังคำนวณคือตำแหน่ง (x, y) ตำแหน่งกึ่งกลางของหน้ากากคือตำแหน่ง $(0,0)$ ดังนั้นจะเห็นได้ว่า เมื่อดึงข้อมูลออกมาจากภาพขนาดเท่ากับหน้ากาก w เราจะได้เมทริกซ์ 2 ชุด ที่มีขนาดเท่ากัน แล้วนำเมทริกซ์ทั้ง 2 ชุดมาทำผลคูณจุด สามารถเขียนสมการได้ดังนี้

$$R = w(-1, -1)f(x - 1, y - 1) + w(-1, 0)f(x - 1, y) + \dots + w(0, 0)f(x, y) + \dots + w(1, 0)f(x + 1, y) + w(1, 1)f(x + 1, y + 1) \quad (2.6)$$

ซึ่งเป็นการนำข้อมูลในแต่ละตำแหน่งที่ตรงกันมาคูณกัน จากนั้น นำผลคูณทั้งหมดมารวมกัน ขนาดของหน้ากากคำนวณได้จาก $m = 2a + 1$ และ $n = 2b + 1$ ซึ่งสมการนี้มีความหมายว่า m และ n เป็นเลขคี่

นอกจากนั้น เพื่อให้อยู่ในรูปง่าย สามารถสรุปสมการให้อยู่ในรูปเวกเตอร์ได้ โดยการนำ ค่าในภาพมาใส่ในเวกเตอร์ $z = \{z_i \mid 1 \leq i \leq n\}$ และสร้างหน้ากาก $w = \{w_i \mid 1 \leq i \leq n\}$ ดังนั้น จึงเขียนในรูปของการทำผลคูณจุดได้

$$R = w \cdot z \quad (2.7)$$

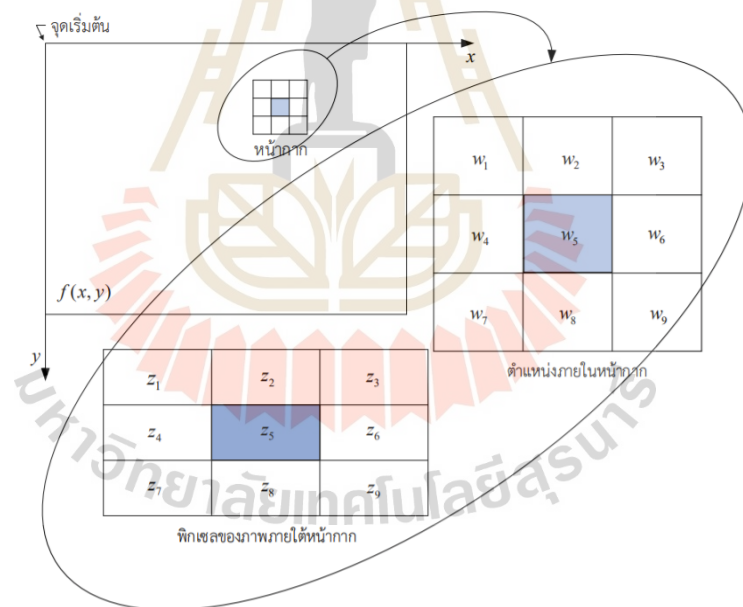
ซึ่งเขียนในรูปฟังก์ชันได้

$$R = \sum_{i=1}^{mn} w_i z_i = w_1 z_1 + w_2 z_2 + \dots + w_{mn} z_{mn} \quad (2.8)$$

m และ n คือขนาดของหน้าภาพ สมมติให้ $m = 3$ และ $n = 3$ ดังนั้น ขอบเขตของการประมวลผล พิกเซลรอบตำแหน่ง (x, y) คือ $a = 1, b = 1$ และสามารถเขียนสมการให้อยู่ในรูปง่ายได้เป็น

$$R = \sum_{i=1}^9 w_i z_i = w_1 z_1 + w_2 z_2 + \dots + w_9 z_9 \quad (2.9)$$

การเลื่อนหน้าภาพไปบนภาพ แล้วนำตำแหน่งของพิกเซลที่ตรงกันจับมาคูณกับ แล้วนำทั้งหมดมารวมกัน จากนั้น ทำซ้ำไปกับทุกจุดบนภาพ เมื่อทำครบทุกตำแหน่งแล้ว ก็ได้ผลลัพธ์ของการทำสังวัตนาการ รูปที่ 2.12 แสดงการทำสังวัตนาการระหว่าง Z และ W



รูปที่ 2.12 การทำสังวัตนาการระหว่าง Z และ W

(ที่มา: รองศาสตราจารย์ ดร.สนั่น ศรีสุข หนังสือการประมวลผลภาพขั้นสูง)

ปัญหาหนึ่งของการทำสังวัตนาการ คือเมื่อเลื่อนหน้าภาพจนทำให้จุดกึ่งกลางของหน้าภาพ อยู่ที่ตำแหน่งซ้ายสุด ขวาสุด บนสุด หรือล่างสุดของภาพ จะทำให้บางตำแหน่งในหน้าภาพอยู่นอกขอบเขตของภาพ (ตกขอบ) ทำให้ตำแหน่งนั้นที่ตกขอบ ไม่สามารถคำนวณได้ หากพิจารณาจาก

หน้าภาพขนาด $m \times n$ จะต้องมิตำแหน่งใดตำแหน่งหนึ่งที่ตกขอบ เมื่อจุดกึ่งกลางของหน้าภาพอยู่ห่างจาก ตำแหน่งซ้ายสุด ขวาสุด บนสุด หรือล่างสุดของภาพ เป็นระยะทาง $(n - 1)/2$ พิกเซล ดังนั้น วิธีแก้ไข คือจำกัดการเลื่อนหน้าภาพ ไม่ให้จุดกึ่งกลางของหน้าภาพ มีระยะทางน้อยกว่า $(n - 1)/2$ พิกเซล จากขอบของภาพ ซึ่งวิธีการนี้จะทำให้ผลลัพธ์การทำสังวัตนาการ มีขนาดเล็กลง หรืออีกวิธี หนึ่งคือ ขยายภาพออกไป (Padding) ขนาด $(n - 1)/2$ ทั้งทางซ้าย ขวา บน และล่าง โดยการเติม ค่าคงที่ แล้วคำนวณเฉพาะภายในขอบเขตของภาพเดิม วิธีการนี้จะทำให้ได้ภาพขนาดเท่าเดิม แต่การ คำนวณบริเวณขอบซ้าย ขวา บน และล่าง อาจไม่ถูกต้องมากนัก เนื่องจากบริเวณที่ขยายออกไป ไม่ใช่ค่าตัวเลขของภาพจริง ๆ

2.3.4 ฟังก์ชันเกาส์เซียน (Gaussian Function)

ฟังก์ชันที่สำคัญมากเป็นอันดับต้น ๆ ของการประมวลผลภาพคือ ฟังก์ชันเกาส์เซียน โดยตั้ง ตามชื่อคาร์ล เฟรดริช เกาส์ (Carl Friedrich Gauss) ซึ่งนับได้ว่าเป็นนักคณิตศาสตร์ที่สำคัญมากคน หนึ่งของโลก สมการของฟังก์ชันเกาส์เซียนคือ

$$g(x) = ae^{-\frac{(x-b)^2}{2c^2}} \quad (2.10)$$

เมื่อ $e = 2.71828$ คือตัวเลขจากออยเลอร์ Euler's number ตัวแปร a คือค่าความสูงสูงสุดของเส้นโค้งของกราฟ ตัวแปร b คือ ตำแหน่งจุดศูนย์กลางของกราฟเกาส์เซียน และตัวแปร c ควบคุมความกว้างของกราฟเกาส์เซียน เมื่อเราวาดกราฟของฟังก์ชันเกาส์เซียน ผลลัพธ์ที่ได้จะเป็นรูประฆังคว่ำ โดยจะมีค่าสูงที่สุด ณ จุดกึ่งกลางของกราฟ ซึ่งกำหนดโดยตัวแปร b แล้วจะมีค่าลดลงทั้งสองทางในทิศทาง $\pm\infty$ ในมุมมองทางสถิติ ฟังก์ชันเกาส์เซียนเป็นการแจกแจงความน่าจะเป็นแบบปกติของตัวแปรที่เกิดแบบสุ่ม ซึ่งมีแนวโน้มที่จะเกิดใกล้กับค่าเฉลี่ย

2.4 ปัญญาประดิษฐ์

ปัจจุบันระบบการทำงานต่าง ๆ รวมถึงเทคโนโลยีสมัยใหม่พัฒนาตามการพัฒนาของมนุษย์ มีฟังก์ชันซึ่งสามารถเข้าใจ และเรียนรู้สิ่งต่าง ๆ ได้ เช่น การสอนให้เครื่องจักรเรียนรู้ในสิ่งต่าง ๆ การหาเหตุและผล รวมถึงการแก้พวกปัญหาต่าง ๆ เพื่อสนับสนุนการทำงานหลาย ๆ อย่างให้กับมนุษย์ ถ้าเทคโนโลยีสามารถคิดวิเคราะห์อย่างมีเหตุผล ตัดสินใจ และ เรียนรู้ได้เอง จะสามารถประยุกต์เพื่อแก้ปัญหาต่าง ๆ โดยจะมีประโยชน์ต่อมนุษย์ สิ่งเหล่านี้เป็นสิ่งที่ท้าทายตรรกะความคิดของมนุษย์ของการสร้างความฉลาดในการเรียนรู้ให้กับเทคโนโลยีกับแนวคิดที่จะสร้างการประมวลผลคล้ายสมองหรือเรียกอีกชื่อหนึ่งว่า ปัญญาเทียม ที่มีความฉลาดเทียบเท่ากับความคิดของมนุษย์ ทำให้เกิดศาสตร์

ความรู้แขนงหนึ่งทางด้านคอมพิวเตอร์ขึ้นมาเรียกว่า ปัญญาประดิษฐ์ ในชื่อเรียกภาษาอังกฤษ Artificial Intelligence หรือ AI

2.4.1 ความหมายปัญญาประดิษฐ์

ปัญญาประดิษฐ์ (Artificial Intelligence) คือวิชาการความรู้ส่วนหนึ่งของสาขาวิทยาการคอมพิวเตอร์ เพื่อพัฒนาอุปกรณ์ต่างให้สามารถทำงานเปรียบเสมือนมนุษย์ เช่น เครื่องคอมพิวเตอร์ หรือหุ่นยนต์ที่ใช้งานในปัจจุบัน เพื่อให้มีความสามารถด้านการวิเคราะห์ประมวลผลในการตัดสินใจ ทางหลักการความคิดและพฤติกรรมที่คล้ายกับความคิดมนุษย์ โดยแก้ปัญหาหรือหาเหตุผลจากข้อมูล หรือองค์ความรู้ที่มีอยู่แล้วนำมาวิเคราะห์ และแปลความหมายให้ได้ผลสรุปหรือผลลัพธ์ที่ถูกต้องและมีความใกล้เคียงเสมือนกับการตัดสินใจของมนุษย์ ปัญญาประดิษฐ์เป็นศาสตร์ที่ช่วยสนับสนุนระบบการทำงานที่อาศัยการทำงานของเครื่องจักร หรือเครื่องคอมพิวเตอร์ เมื่อมีการพัฒนาปัญญาประดิษฐ์ที่มี ประสิทธิภาพจะทำให้เกิดข้อผิดพลาดน้อย และสามารถตัดสินใจเร็วกว่าการประมวลผลสมองของมนุษย์ เนื่องจากมนุษย์ยังต้องอาศัยพื้นฐานทางกายภาพทางร่างกายทำให้การทำงานอาจเกิดความผิดพลาดขึ้นได้เสมอ

2.4.2 คุณสมบัติต่าง ๆ และข้อจำกัดปัญญาประดิษฐ์

ถ้ากล่าวถึงหลักการการทำงานปัญญาประดิษฐ์ จากการมองภาพรวมที่เกี่ยวกับการสร้างเครื่องคอมพิวเตอร์ให้มีประสิทธิภาพที่สามารถทำงานเทียบเท่ากับการประมวลผลของมนุษย์หรือมากกว่า ปัญญาประดิษฐ์เมื่อถูกพัฒนา จะมีคุณสมบัติตามด้านล่างดังต่อไปนี้

- การเข้าใจและเรียนรู้จากข้อมูลที่ผ่านมา
- การตอบสนองที่รวดเร็ว
- การใช้หลักเหตุและผล เพื่อแก้ปัญหา พร้อมทั้งแนะนำวิธีการ หรือ แนวทาง
- เข้าใจปัญหาที่มีความซับซ้อน รวมถึงวิเคราะห์เพื่อหาข้อสรุปที่เหมาะสมได้
- การทำความเข้าใจกับเนื้อหาข้อมูลที่คลุมเครือ
- การนำองค์ความรู้มาประยุกต์ใช้ได้เหมาะสมกับสถานการณ์ และ สภาพแวดล้อมใน

แบบต่าง ๆ

- การจำแนกข้อมูล ความสำคัญ และปัญหาที่มีลักษณะองค์ประกอบไม่เหมือนกันตามสถานการณ์

ตามที่กล่าวด้านบน ในเนื้อ นั้น เป็นส่วนหนึ่งของความสามารถหรือคุณสมบัติของปัญญาประดิษฐ์บางส่วนเท่านั้น ความเป็นจริงเมื่อนำไปพัฒนาใช้งานจะมีคุณสมบัติตามงานที่นำไปใช้ความสามารถส่วนใหญ่ของปัญญาประดิษฐ์มักต้องพึ่งพาการทำงานของเครื่องคอมพิวเตอร์ที่มีประสิทธิภาพสูงที่มีการประมวลผลที่รวดเร็วในการคำนวณ สำหรับเป็นตัวควบคุมและสนับสนุนในการทำงาน ให้มีความถูกต้องตามวัตถุประสงค์ที่ตั้งไว้

2.4.3 ข้อเปรียบเทียบและข้อจำกัดของปัญญาประดิษฐ์

การเปรียบเทียบปัญญาประดิษฐ์กับการทำงานของมนุษย์ สรุบบื้องต้นของปัญญาประดิษฐ์ได้ดังนี้ ในด้านการคำนวณ ทั่วไปใช้คอมพิวเตอร์ประมวลผลและวิเคราะห์ข้อมูล หรือตัวเลขที่ซับซ้อนมีปริมาณมากกว่าสมองมนุษย์สามารถคำนวณประมวลได้ ปัญญาประดิษฐ์มีส่วนช่วยให้วิเคราะห์ประมวลขั้นตอนการคำนวณได้ถูกต้องและแม่นยำขึ้น ยิ่งไปกว่านั้น สามารถประมวลผลการคำนวณในปริมาณมาก ๆ ได้ อย่างรวดเร็ว และเกิดความผิดพลาดน้อย ดังนั้น จึงเป็นข้อเปรียบเทียบการประยุกต์ใช้ ปัญญาประดิษฐ์ทำให้ประหยัดเวลาในด้านอื่น ๆ ค่อนข้างเยอะ

ในเรื่องการจัดเก็บข้อมูล ถ้าในระบบมีข้อมูลปริมาณมาก การเก็บข้อมูลในสมองมนุษย์จะไม่ใช้ตัวเลือกที่ปฏิบัติ ดังนั้นเครื่องคอมพิวเตอร์จะเป็นตัวเลือกสำหรับการเก็บข้อมูลเหล่านี้ เมื่อมีการนำปัญญาประดิษฐ์ใช้ในระบบจัดเก็บข้อมูล จะช่วยให้เครื่องคอมพิวเตอร์มีความสามารถมากขึ้น ไม่ได้แค่เป็นเพียงพื้นที่จัดเก็บข้อมูลอย่างเดียว ปัญญาประดิษฐ์จะช่วยเพิ่มประสิทธิภาพการแบ่งจำแนกข้อมูลต่าง ๆ ส่งผลให้เครื่องคอมพิวเตอร์สามารถจัดระเบียบการจัดเก็บข้อมูลได้ด้วยตนเอง โดยไม่ต้องใช้แรงงานมนุษย์ในการทำงาน ในกรณีข้อมูลเหล่านั้น เป็นความรู้หรือประสบการณ์ที่ต้องอาศัยการตีความก่อนของข้อมูลที่รับมา เครื่องคอมพิวเตอร์ไม่สามารถทำความเข้าใจได้ จะทำได้แค่เพียง จัดเก็บสิ่งเหล่านี้ในลักษณะของข้อมูลเท่านั้น

ระบบรูปแบบการทำแบบวนซ้ำ ที่เป็นลักษณะการทำงานแบบเดิม ๆ สิ่งเหล่านี้ส่งผลต่อการคิดวิเคราะห์ความสามารถในการตัดสินใจของมนุษย์ เพราะว่ามันจะมีข้อจำกัดทางร่างกายและความเหนื่อยล้าของสมอง ถ้าทำงานแบบเดิมในระยะเวลาที่นานอาจจะเกิดความผิดพลาดเนื่องจากความเบื่อหน่ายในการทำงานและอาการเหนื่อยล้าของระบบร่างกาย อันนี้คือปัจจัยที่มีผลกระทบต่อการทำงานโดยตรง ทำให้ขาดความต่อเนื่องและใช้เวลาการทำงานมากขึ้น ในอุตสาหกรรมที่เรียกกันบ่อย ๆ ว่า การใช้ Tag-time ในการประกอบชิ้นงาน ถ้าเปรียบเทียบการใช้เครื่องคอมพิวเตอร์ที่มีปัญญาประดิษฐ์เหล่านี้นำไปใช้กับเครื่องจักรกลในสายการผลิตด้านต่าง ๆ เช่น อุตสาหกรรมอิเล็กทรอนิกส์ อุตสาหกรรมการบิน หรืออาหารและยา เป็นต้น ถ้านำไปใช้จะได้คุณภาพงานที่คงที่ และยังมีความสม่ำเสมอด้วยไม่แปรผัน ไม่เหมือนการใช้แรงงานมนุษย์ แต่ถ้ามองเรื่องค่าใช้จ่าย การใช้ปัญญาประดิษฐ์จะใช้ต้นทุนสูงแต่ถ้ามองการลงทุนที่มีการผลิตจำนวนมากหรือระยะยาว ถือเป็นตัวเลือกที่ค่อนข้างคุ้มค่ากับการลงทุน

นอกจากหน้าที่ของปัญญาประดิษฐ์ที่มีจุดเด่นในการนำมาใช้แทนแรงงานมนุษย์แล้ว ยังมี ความสามารถของปัญญาประดิษฐ์ที่แตกต่างจากเครื่องคอมพิวเตอร์ทั่วไป ซึ่งเป็นส่วนหนึ่งที่ทำให้เห็นได้ อย่างชัดเจนว่า เหตุใดปัญญาประดิษฐ์จึงมีความสำคัญและมีประสิทธิภาพมากกว่าและการใช้เครื่อง คอมพิวเตอร์ ข้อเปรียบเทียบระหว่างเครื่องคอมพิวเตอร์ทั่วไปกับปัญญาประดิษฐ์

จากที่กล่าวมาส่วนใหญ่ประสิทธิภาพของคอมพิวเตอร์ทั่วไปยังอิงตาม Algorithm ที่สร้างไว้ เนื่องจากการประมวลผลความรู้ของคอมพิวเตอร์ทั่วไปจะดำเนินตามลำดับขั้นตอนที่ถูกกำหนดไว้ใน Algorithm ที่สร้างขึ้นในโปรแกรม ในขณะที่คอมพิวเตอร์ปัญญาประดิษฐ์อาศัยการประมวลผลจากองค์ความรู้ที่เรียกว่า “การประมวลผลจากสัญลักษณ์ (Symbolic Processing)” เป็นการประมวลผลจากการ เชื่อมโยงของข้อมูลในรูปสัญลักษณ์ที่แทนตัวอักษร คำ ข้อความ ประโยค หรือ ออบเจ็ค เพื่อใช้ค้นหาคำตอบที่เหมาะสมสำหรับแก้ไขปัญหาที่ต้องการ ดังนั้น การทำงานของคอมพิวเตอร์ทั่วไปกับคอมพิวเตอร์ ปัญญาประดิษฐ์จึงมีความแตกต่างกันอย่างเห็นได้ชัด ทั้งนี้ ควรพิจารณาความเหมาะสมของระบบงานว่าควร เลือกใช้คอมพิวเตอร์แบบใดเพื่อให้ได้ประโยชน์อย่างสูงสุดและไม่สิ้นเปลืองงบประมาณอย่างเปล่าประโยชน์

ทั้งหมดที่กล่าวมานี้คือการเปรียบเทียบการทำงานของปัญญาประดิษฐ์กับมนุษย์ และ คอมพิวเตอร์ทั่วไป ในความเป็นจริงมีอีกหลายปัจจัยของการนำปัญญาประดิษฐ์มาใช้แต่ยังไม่สามารถตอบสนองตามวัตถุประสงค์ที่ตั้งไว้ได้ดีเท่ากับการทำงานของมนุษย์ เพราะว่ามีมนุษย์มีความรู้หรือประสบการณ์ที่สะสมมาแตกต่างกัน หรือในบางคนมีประสบการณ์มากและสามารถเข้าใจปัญหาที่เจอและแก้ไขได้ตรงจุดและรวดเร็ว ปัญญาประดิษฐ์ยังมีข้อจำกัดในการทำงานหลาย ๆ ด้าน ดังนี้

1) ปัญญาประดิษฐ์ นั้น ไม่มีการสร้างไอเดียหรือการริเริ่มการสร้างสิ่งใหม่ ๆ ดังนั้น จะไม่สามารถสร้างสิ่งต่าง ๆ โดยปราศจากข้อมูลองค์ความรู้ ถ้ามองกลับกัน มนุษย์มักมีจินตนาการ ไอเดียเป็นตัวขับเคลื่อนให้คิดค้นสิ่งใหม่ ๆ ที่อยู่นอกเหนือจากข้อมูลหรือองค์ความรู้ได้อีก รวมทั้งมีสภาพแวดล้อมเป็นแรงผลักดันและแรงบันดาลใจ

2) ปัญญาประดิษฐ์ต้องใช้ข้อมูลเป็นองค์ประกอบทางด้านความรู้เพื่อให้ตัดสินใจได้ถูกต้องดังนั้น ข้อมูลแต่ละอย่างที่รับเข้ามาต้องผ่านกระบวนการแปลงข้อมูลเพื่อให้เหมาะสมก่อนนำไปประมวลผลใช้งาน กลับกัน มนุษย์สามารถเข้าใจข้อมูลได้ทันที เพราะว่ามีประสบการณ์การเรียนรู้ข้อมูลมาก่อน

3) ปัญญาประดิษฐ์จะถูกพัฒนาให้มีหน้าที่เฉพาะด้านใดด้านหนึ่ง หากเป็นการทำงานที่ นอกเหนือจุดประสงค์แล้ว การให้เหตุผลอาจมีประสิทธิภาพไม่เทียบเท่ากับมนุษย์ ในขณะที่มนุษย์สามารถ รับรู้ระบบการทำงานในหลาย ๆ ด้านได้ตามประสบการณ์ ถึงแม้จะมีข้อจำกัดเรื่องความจำก็ตาม

2.4.4 ประโยชน์ของปัญญาประดิษฐ์

การนำปัญญาประดิษฐ์มาใช้ในระบบงานต่าง ๆ ขององค์กร นอกจากจะช่วยเพิ่มประสิทธิภาพของระบบงานที่ต้องการได้แล้ว ปัญญาประดิษฐ์ยังส่งผลต่อการทำงานในภาพรวมขององค์กร ได้เป็นอย่างดี และยังสร้างประโยชน์ในด้านต่าง ๆ อีกมากมาย ได้แก่

- ช่วยให้กระบวนการแก้ไขปัญหาที่เกิดขึ้นทำได้รวดเร็ว และทำให้มีความสอดคล้องกันมากขึ้น
- ช่วยแก้ไขปัญหาต่าง ๆ ในกรณีที่ยากเกินความสามารถของมนุษย์และคอมพิวเตอร์ทั่วไปที่จะวิเคราะห์ได้
- รongรับข้อมูลและองค์ความรู้จำนวนมาก
- สามารถจัดเก็บข้อมูลได้แบบถาวรและถ่ายโอนได้
- การเผยแพร่องค์ความรู้ต่าง ๆ ภายในองค์กรทำได้ง่าย และทั่วถึงมากขึ้น
- การวิเคราะห์ปัญหาทำได้ละเอียดรอบคอบมากขึ้นเพราะว่าต้องทำตามขั้นตอนที่กำหนดให้ครบ โดยเฉพาะปัญญาประดิษฐ์ที่ถูกพัฒนามาใช้งานเฉพาะด้าน เช่น ทางวิทยาศาสตร์
- มีความยืดหยุ่นต่อสถานการณ์และตอบสนองได้อย่างมีประสิทธิภาพ เมื่อถูกนำมาประยุกต์ใช้
- การควบคุมเครื่องจักรกลต่าง ๆ รวมถึงเครื่องคอมพิวเตอร์สามารถทำได้ง่ายขึ้น

คุณสมบัติ ข้อจำกัด และประโยชน์ของปัญญาประดิษฐ์เป็นอีกสิ่งหนึ่งที่แสดงให้เห็นว่าประสิทธิภาพและความสามารถการทำงานของปัญญาประดิษฐ์ นั้น มีความสำคัญต่อระบบงานในองค์กรมากขนาดไหน และเป็นปัจจัยสำคัญที่ชี้ให้เห็นได้อย่างชัดเจนของการนำปัญญาประดิษฐ์มาประยุกต์งานในด้านต่าง ๆ และช่วยให้ทราบถึงทิศทางความก้าวหน้าและทิศทางขององค์กรได้ในระดับหนึ่ง นอกจากนี้ยังเป็นแนวทางสำหรับการพัฒนาปัญญาประดิษฐ์ให้มีความสามารถมากขึ้น และนำไปประยุกต์ใช้งานร่วมกับแขนงวิชาต่าง ๆ ได้อย่างมีประสิทธิภาพ

2.4.5 การนำปัญญาประดิษฐ์มาประยุกต์ใช้งาน

ขอบเขตในการประยุกต์ใช้งานทางปัญญาประดิษฐ์ไม่ได้จำกัดเฉพาะเพียงในสาขาที่เกี่ยวข้องกับปัญญาประดิษฐ์ เท่านั้น แต่ยังขยายไปถึงสิ่งที่เกี่ยวข้องกับพฤติกรรมความชาญฉลาดและสิ่งต่าง ๆ ที่สามารถดึงความสามารถปัญญาประดิษฐ์มาใช้ ในบทความนี้จะนำเสนอ สาขาต่าง ๆ ที่ประสบผลสำเร็จและได้ผลลัพธ์ในการพัฒนาและการประยุกต์ใช้งาน ได้แก่

- เกมส์ เป็นการนำปัญญาประดิษฐ์ไปใช้ในซอฟต์แวร์ด้านบันเทิงในแวดวงอุตสาหกรรมความบันเทิง ส่วนใหญ่จะเป็นการนำไปใช้ในการเพิ่มความเสมือนจริงให้กับระบบการทำงานของเกมให้มีระดับความยากมากขึ้น เป็นการเพิ่มความท้าทายให้กับผู้เล่นที่ต้องการชนะ การเพิ่มประสิทธิภาพในด้านนี้จะช่วยให้ผู้เล่นรู้สึกเหมือนกับเกมมีระบบความคิดเป็นของตนเอง สามารถที่จะโต้ตอบ ตอบสนองการสั่งการของผู้ใช้โดยราวกับมีตัวตนอยู่ หรือมีมนุษย์ควบคุมอยู่ เปรียบเสมือนผู้เล่นมนุษย์อีกคนหนึ่ง การพัฒนาของปัญญาประดิษฐ์สามารถพัฒนาไปพร้อมร่วมกับการพัฒนาเกม

เนื่องจากการนำเสนอลักษณะที่เกี่ยวข้องกับ ปัญหาขนาดเล็ก และมีกฎที่ชัดเจน ง่ายต่อการ นำเสนอด้วยรูปแบบของสัญลักษณ์ ซึ่งคุณลักษณะที่กล่าวมา นี้ทำให้เกมเป็นแบบทดสอบที่ดีเยี่ยมทาง ปัญญาประดิษฐ์ทั้งในแง่แนวคิด และแง่เทคนิค ส่งผลให้เกม และปัญญาประดิษฐ์ต่างพัฒนาตนเองไป อย่างมาก

- ระบบผู้เชี่ยวชาญ (Expert System : ES) ถือเป็นอีกระบบหนึ่งที่พัฒนามาจาก ปัญญาประดิษฐ์ โดยรู้จักกันอย่างกว้างขวางและนำไปใช้งานอย่างแพร่หลายในหลาย ๆ ส่วน โดย ระบบปัญญาประดิษฐ์จะช่วยในการตัดสินใจเกี่ยวกับปัญหาหรือข้อมูลที่ซับซ้อนและต้องใช้การ ประมวลผลค่อนข้างเยอะ จึงต้องอาศัยองค์ประกอบของความรู้ต่าง ๆ ที่เกี่ยวข้อง เพื่อที่จะใช้เป็น ข้อมูลสนับสนุนในการแก้ไขปัญหาที่เกิดขึ้น จนกระทั่งได้คำตอบที่ดีที่สุดและเหมาะสมกับสถานการณ์ ที่นำปัญญาประดิษฐ์มาประยุกต์ใช้ ระบบผู้เชี่ยวชาญเป็นการนำปัญญาประดิษฐ์มาใช้ในอีกรูปแบบ หนึ่งโดยให้เครื่องคอมพิวเตอร์ทำหน้าที่เสมือนเป็นที่ปรึกษาคล้าย ๆ กับมนุษย์ ซึ่งเดิมทีอาจใช้แรงงาน มนุษย์ แต่เนื่องจากปัญหามีความซับซ้อน ข้อมูลแวดล้อมมีปริมาณมาก การอาศัยแรงงานมนุษย์อาจ ทำได้ล่าช้างบประมาณที่ใช้สำหรับจ้างผู้เชี่ยวชาญหรือพนักงานค่อนข้างสูง การนำปัญญาประดิษฐ์มา ประยุกต์ใช้เป็นระบบผู้เชี่ยวชาญจึงเป็นทางเลือกหนึ่งที่สามารถแก้ไขปัญหาดังกล่าวและลดต้นทุน ของบริษัทได้

- การให้เหตุผลเชิงกรณีศึกษา (Case-based Reasoning) หรือมีชื่อเรียกสั้น ๆ ว่า CBR คือการ ประยุกต์ใช้งานปัญญาประดิษฐ์อีกรูปแบบหนึ่งที่อาศัยพื้นฐานทางความคิดและ ประสบการณ์การเรียนรู้จากมนุษย์สำหรับเป็นแนวทางในการพัฒนาระบบการประมวลผล โดยอ้างอิง วิธีการแก้ไขปัญหาต่างจากกรณีศึกษาที่เคยผ่านมา ซึ่ง CBR เป็นระบบที่มีกระบวนการแก้ไขปัญหา ด้วยการสรุปความและหาเหตุผลจากพวกปัญหาหรือกรณีศึกษาแต่เดิมที่เคยผ่านมา CBR เป็นระบบที่ มีกระบวนการแก้ไขปัญหาด้วยการสรุปความ และหาเหตุผลจากปัญหาที่เกิดขึ้นหรือ กรณีศึกษาเดิมที่ เคยพบในอดีต กรณีศึกษาเหล่านี้จะถูกจัดเก็บเป็นองค์ความรู้ไว้ สำหรับเมื่อระบบต้องเผชิญกับ ปัญหา จะนำกรณีศึกษาที่มีความใกล้เคียงกันนำมาเป็นแนวทางในการแก้ไขปัญหาดังกล่าว โดย กรณีศึกษาที่ถูกจัดเก็บไว้จะนำมาใช้แก้ปัญหาได้โดยตรง หรือทางอ้อมซึ่งอาจต้องดัดแปลงตามความ สอดคล้อง และเรียนรู้เพื่อหาวิธีการแก้ไขปัญหาดังกล่าวจากกรณีศึกษามากที่สุด หลังจากระบบแก้ไข ปัญหาเรียบร้อยแล้ว จึงดำเนินการจัดเก็บไว้ในฐานกรณีศึกษาทุกครั้ง โดยเฉพาะกรณีศึกษาใหม่ๆเพื่อ เป็นแหล่งองค์ความรู้ใหม่ ๆ สำหรับใช้งานในอนาคตต่อไป จากที่กล่าวมา CBR นิยมนำมาใช้กับ รูปแบบงานที่มีโอกาสเผชิญกับปัญหาที่คล้ายกันอยู่เสมอ เพื่อให้ระบบหรือซอฟต์แวร์ที่ประมวลผล สามารถแก้ไขปัญหาได้ด้วยตนเอง โดยไม่ต้องพึ่งพาอาศัยแรงงานการวิเคราะห์แก้ปัญหาจากมนุษย์ ทั้งนี้ ระบบจะทำการเรียนรู้จากกรณีศึกษาที่จัดเก็บไว้ในฐานองค์ความรู้ การประยุกต์ใช้งานระบบ CBR มีอยู่ หลายประเภท เช่น ระบบการตรวจจับข้อผิดพลาดที่เกิดขึ้นในสายงานผลิตภายในโรงงาน

ซึ่งเป็นการตรวจสอบโดย อาศัยกรณีศึกษา เพื่อลดข้อผิดพลาดที่อาจเกิดขึ้นซ้ำ ๆ ให้น้อยลงจากการเรียนรู้และหาแนวทางป้องกันจากฐานกรณีศึกษาที่ถูกบันทึกไว้ เป็นต้น การแก้ไขปัญหาต่าง ๆ ด้วยกรณีศึกษา ขึ้นอยู่กับประสบการณ์ที่ผ่านมาว่าพบกับรูปแบบปัญหามากน้อยเพียงใด ยิ่งพบเจอปัญหา มากก็ยิ่งมีกรณีศึกษาที่มีความหลากหลายของข้อมูล การแก้ไขปัญหามีประสิทธิภาพมากขึ้นตาม เช่นเดียวกับมนุษย์ที่มีประสบการณ์มากย่อมได้เปรียบผู้ที่มีประสบการณ์น้อยนั่นเองในการตัดสินใจต่าง ๆ

- กลไกการเรียนรู้ (Machine Learning) เป็นการนำปัญญาประดิษฐ์มาประยุกต์ใช้ในด้านการเรียนรู้ เพื่อเพิ่มประสิทธิภาพให้เครื่องจักรกลต่าง ๆ และระบบเครื่องคอมพิวเตอร์มีความรู้ในระบบงานที่มีหน้าที่รับผิดชอบ ซึ่งเป็นการเรียนรู้จาก สถานการณ์ หรือองค์ความรู้ใหม่ๆที่เกิดขึ้นตามสภาพแวดล้อมต่าง ๆ โดยอาจเป็นสิ่งที่เครื่องคอมพิวเตอร์ไม่เคยพบมาก่อน ทำให้การทำงานมีประสิทธิภาพมากขึ้น อีกทั้งยังลดความผิดพลาดลงได้อีกทางหนึ่ง เนื่องจากเครื่องคอมพิวเตอร์หรือเครื่องจักรมีองค์ความรู้ใหม่ที่มีข้อมูลครอบคลุมทุกสถานการณ์หรือปัญหาที่เกิดขึ้น ปัญญาประดิษฐ์เป็นตัวช่วยให้ระบบของเครื่องคอมพิวเตอร์หรือเครื่องจักรเกิดการเรียนรู้สิ่งต่าง ๆ ได้ด้วยตนเองโดยไม่ต้องอาศัยพึ่งพาการสั่งจากมนุษย์ กลไกการเรียนรู้มีจุดมุ่งหมายเพื่อค้นคว้าและเข้าถึงฐานความรู้ เพื่อเพิ่มความรู้จากสถานการณ์ที่เกิดขึ้นได้

- การคำนวณเชิงวิวัฒนาการ (Evolutionary Computation) เป็นการเรียนรู้อีกชนิดสำหรับประยุกต์ใช้งานกับเครื่องจักรหรือเครื่องคอมพิวเตอร์ประมวลผล เพื่อเพิ่มประสิทธิภาพในการเรียนรู้ให้สูงขึ้น ซึ่งเลียนแบบวิวัฒนาการของสิ่งมีชีวิตต่าง ๆ การเรียนรู้รูปแบบนี้ทำงานอยู่บนพื้นฐานทฤษฎีวิวัฒนาการของ Charles Darwin ที่ได้นำเสนอเมื่อปี ค.ศ. 1858 โดยนำมาประยุกต์ใช้กับการเรียนรู้ของเครื่องจักร เช่น ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm: GA) และ โปรแกรมเชิงพันธุกรรม (Genetic Programming: GP) เป็นต้น เป็นทฤษฎีการดำรงอยู่ของสิ่งมีชีวิตที่เกิดการเปลี่ยนแปลงตลอดเวลา ส่งผลให้สิ่งมีชีวิตมีวิวัฒนาการ (Evolution) โดยปรับเปลี่ยนลักษณะต่าง ๆ ทั้งทางกายภาพ และพฤติกรรมให้มีความเหมาะสม เพื่อให้สามารถอยู่รอดในสภาพแวดล้อมขณะนั้นได้ การคำนวณเชิงวิวัฒนาการจะมุ่งเน้นในการศึกษาด้านพันธุกรรมของสิ่งมีชีวิต เพื่อเป็นแนวทางในการพัฒนาการเรียนรู้ของเครื่องจักร เนื่องจากสิ่งมีชีวิตจะมีลักษณะทางพันธุกรรม (Genetic Character) ที่สามารถถ่ายทอดจากพ่อแม่ไปสู่ลูกหลานได้ด้วยหน่วยพันธุกรรมที่เรียกว่า “ยีน (Gene)” สำหรับลักษณะทางพันธุกรรมทั่วไป เช่น สีผม สีผิว สติปัญญา และความสูง เป็นต้น ยีนเหล่านี้เป็นส่วนหนึ่งของ DNA ที่ปรากฏอยู่บนตำแหน่งโครงสร้างของโครโมโซม (Chromosome) โดยมีหน้าที่ถ่ายทอดและ ควบคุมลักษณะทางพันธุกรรม ด้วยเหตุนี้รูปแบบของการวิวัฒนาการของสิ่งมีชีวิตจึงถูกนำมาประยุกต์ใช้ใน กระบวนการเรียนรู้ ที่สามารถพัฒนาสติปัญญาของเครื่องจักรให้มีกระบวนการคิดที่ดีขึ้นและเรียนรู้สิ่งต่าง ๆ ได้จากการเปลี่ยนแปลงดังกล่าว การเรียนรู้ในการแก้ไข

ปัญหาต่าง ๆ จึงมีประสิทธิภาพสูงขึ้น การจัดการกับปัญหาจะพัฒนาขึ้นจากเดิม ช่วยให้รูปแบบการเรียนรู้และกระบวนการคิดใกล้เคียงกับมนุษย์มากยิ่งขึ้น

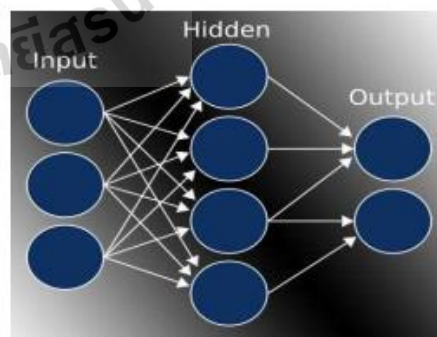
- ระบบโครงข่ายประสาท (Neural Network) เป็นการประยุกต์ใช้ปัญญาประดิษฐ์ที่มีสถาปัตยกรรมที่ค่อนข้างแตกต่างจากสาขาอื่น ๆ โดยเป็นสาขาที่พยายามใช้หลักโครงสร้างการทำงานที่มีลักษณะเลียนแบบระบบประสาทในสมองของ มนุษย์ดังภาพที่ 2.13 (ก) ซึ่งแต่ละส่วนของ

โครงข่ายประสาทจะเชื่อมต่อเข้าด้วยกันและทำงานคล้ายกับระบบประสาทที่มีการประมวลผลได้อย่างรวดเร็ว การประมวลผลองค์ความรู้นี้เรียกว่า การประมวลผลแบบขนาน (Parallel Processing) ซึ่งมีความสามารถในการประมวลองค์ความรู้จำนวนมากได้ในเวลาเดียวกัน

โครงข่ายประสาทนี้สามารถเรียนรู้และจดจำสารสนเทศในรูปแบบที่ปัญญาประดิษฐ์ทั่วไปไม่สามารถทำได้ ซึ่งก็คือ “ประสบการณ์” เนื่องจากการเลียนแบบจากโครงข่ายประสาทของมนุษย์ทำให้สามารถเชื่อมโยงข้อเท็จจริงเข้าด้วยกัน เพื่อใช้หาข้อสรุปโดยอาศัยประสบการณ์ที่ได้จัดเก็บไว้มาเป็นข้อมูล ในการเรียนรู้ และทำความเข้าใจจนได้ความเกี่ยวข้องสัมพันธ์กันระหว่างองค์ความรู้ทั้งหมด ซึ่งสามารถนำไปปรับปรุงองค์ความรู้เดิมให้มีประสิทธิภาพมากยิ่งขึ้น เพื่อความทันสมัยและรองรับสถานการณ์ที่จะเกิดขึ้นในอนาคต ระบบโครงข่ายประสาทประกอบด้วยโหนด (Node) ต่าง ๆ ที่เปรียบเสมือนเป็นเซลล์ ประสาทของมนุษย์ โดยแต่ละโหนดจะเชื่อมโยงกันเป็นเครือข่าย และถูกแบ่งออกเป็นระดับชั้น (Layer) ในแต่ละระดับชั้นจะมีหน้าที่แตกต่างกันโดยอาศัยองค์ประกอบอื่น ๆ เพื่อเริ่มกระบวนการ เช่น ข้อมูลการ นำเข้า (Input) และฟังก์ชันการแปลง (Transformation Function) เป็นต้น ดังภาพที่ 2.13 (ข) จะมีหน้าที่แตกต่างกันโดยอาศัยองค์ประกอบอื่น ๆ เพื่อเริ่มกระบวนการ เช่น ข้อมูลการ นำเข้า (Input) และฟังก์ชันการแปลง



(ก) โครงข่ายประสาทในสมอง



(ข) โครงข่ายประสาทเทียม

รูปที่ 2.13 ตัวอย่างโครงสร้างโครงข่ายประสาทในสมองและโครงข่ายประสาทเทียม
(ที่มา: ผู้ช่วยศาสตราจารย์ ดร. ชูพันธุ์ รัตนโกคา เอกสารความรู้เบื้องต้นทางปัญญาประดิษฐ์)

- ตัวแทนปัญญา (Intelligent Agent) เป็นอีกหนึ่งเทคโนโลยีของปัญญาประดิษฐ์ที่ได้รับการพัฒนาอย่างกว้างขวาง โดยนำปัญญาประดิษฐ์มาประยุกต์ใช้งาน ในอนาคตจะกลายเป็นเทคโนโลยีที่สำคัญทางด้านระบบสารสนเทศ และงานด้านการวิจัยที่เกี่ยวกับคอมพิวเตอร์ เนื่องจากการพัฒนา ตัวแทนปัญญาที่สามารถแก้ไขปัญหาต่าง ๆ และเพิ่มประสิทธิภาพในการทำงานได้ ตัวแทนในที่นี้อาจเรียกว่า “Agent (เอเจนต์)” หรือ “โปรแกรมตัวแทน” ซึ่งหมายถึง ผู้ที่ได้รับมอบหมายให้ทำหน้าที่ตามที่มนุษย์ ต้องการ ซึ่ง Agent สามารถรับข้อมูลต่าง ๆ ได้อย่างเป็นลำดับ จากสภาพแวดล้อมผ่านทางตัว Sensors หรือระบบที่สร้างไว้เพื่อการรับรู้ โดยสามารถตอบสนองกลับ ด้วยการกระทำแบบต่าง ๆ ตามรูปแบบของการ โต้ตอบที่ระบบกำหนดไว้

- การวางแผน (Planning) เป็นอีกเทคนิคหนึ่งทางด้านปัญญาประดิษฐ์ ที่ใช้สำหรับกระบวนการวิเคราะห์ตัดสินใจล่วงหน้า เพื่อใช้ออกแบบแผนการที่จะนำไปสู่จุดมุ่งหมายตามที่ต้องการ การวางแผนจะเป็นลำดับขั้นตอนที่ใช้ในการแก้ไขปัญหโดยต้องมีความสอดคล้องกันระหว่างการกระทำและผลลัพธ์ การวางแผนอาจถูกนำไปใช้ควบคู่กับตัวแทนปัญญา (AI) เพื่อให้ตัวแทนปัญญาทำหน้าที่เป็น ผู้วางแผน (Planner) ปัญหาได้ การวางแผนปัญหา (Planning Problem) เป็นจุดเริ่มต้นของการออกแบบ แผนการด้วยกระบวนการวิเคราะห์ตัดสินใจเพื่อรับมือกับปัญหา ซึ่งวิธีการค้นหา (Searching) ก็เป็นการ วางแผนชนิดหนึ่ง เนื่องจากการหาเส้นทางเพื่อนำไปสู่การแก้ไขปัญหตามที่ต้องการนั่นเอง

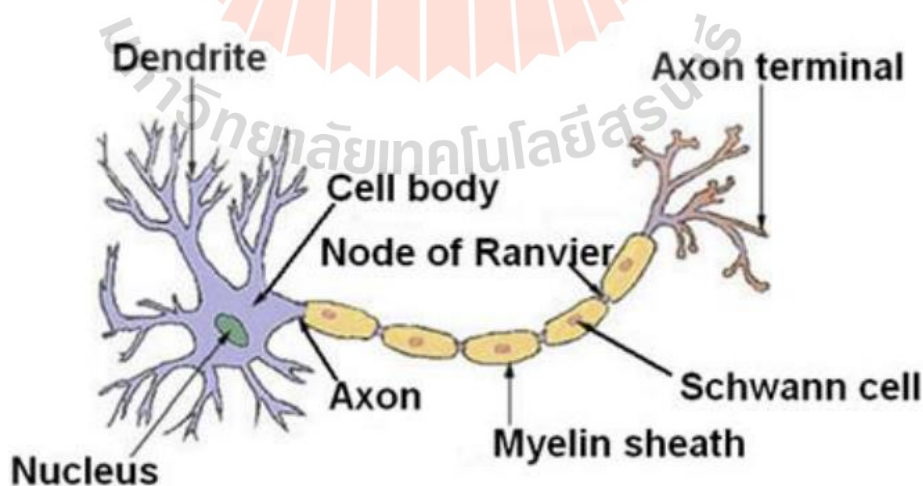
- การประมวลผลภาษาธรรมชาติ (Natural Language Processing : NLP) การประมวลผลภาษาธรรมชาติ คือ รูปแบบการประมวลผลภาษาด้วยการวิเคราะห์จาก เครื่องคอมพิวเตอร์ โดยมีจุดประสงค์ให้ผู้ใช้สามารถโต้ตอบกับคอมพิวเตอร์ได้โดยใช้ภาษาธรรมชาติ ซึ่งผู้ใช้เข้าใจเป็นอย่างดีโดยเครื่องคอมพิวเตอร์จะมีการจัดเก็บองค์ความรู้ที่จำเป็นสำหรับวิเคราะห์และแปลเสียง ไว้เรียบร้อยแล้ว

- หุ่นยนต์ (Robotic) การใช้ปัญญาประดิษฐ์เพื่อพัฒนาเครื่องจักรกลให้มีทักษะ ประสิทธิภาพที่สามารถลาดพอกที่จะทำงานแทนมนุษย์ ถือเป็นการสร้างระบบเครื่องจักรมีลักษณะการทำงานคล้ายคลึงกับมนุษย์ซึ่งสามารถทำงานแบบอัตโนมัติ ที่พบเห็นโปรเจ็คหลาย ๆ อย่าง นำหลักการนี้ไปประยุกต์ใช้กับเครื่องจักรในโรงงานอุตสาหกรรมต่าง ๆ วัตถุประสงค์คือเพื่อความปลอดภัยจากงานที่เสี่ยงอันตราย ลดต้นทุนค่าแรงงานและเพิ่มปริมาณการผลิต เนื่องจากหุ่นยนต์ไม่ต้องการวันหยุดและมีลักษณะการทำงานที่คงที่ การนำปัญญาประดิษฐ์เข้ามาบทบาทในตรงนี้ถือเป็นทักษะศาสตร์ที่เกี่ยวข้องกับการควบคุมพฤติกรรมของหุ่นยนต์หรือเครื่องจักรในลักษณะรูปแบบต่าง ๆ ดังเช่น การมองเห็นวัตถุ การตรวจจับ หลักการเหล่านี้ถูกพัฒนาขึ้นตามวัตถุประสงค์ที่ตั้งไว้ ส่วนประกอบขององค์ความรู้ที่สร้างขึ้นจะถูกผนวกเข้าศาสตร์ความรู้ของระบบปัญญาประดิษฐ์ เช่นทางกลการขยับเคลื่อนที่ไปยังวัตถุเป้าหมายที่กำหนดตัวไว้ จะมีศาสตร์ความรู้ในด้านต่าง ๆ เช่น

กลศาสตร์ หรือการประมวลผลภาพ เพื่อพัฒนาหุ่นยนต์ให้สามารถหลบวัตถุสิ่งกีดขวางและมองเห็นวัตถุเป้าหมายได้

2.5 โครงข่ายประสาทเทียม

โครงข่ายประสาทเทียม (Artificial Neural Network: ANN) มีชื่อเรียกสั้น ๆ ว่า ระบบโครงข่ายประสาท (Neural Network หรือ Neural Net) หมายถึง ระบบคอมพิวเตอร์ที่สร้างขึ้นเลียนแบบ เส้นใยประสาทของมนุษย์ ที่เรียกว่า นิวรอน (Neurons) โดยโครงข่ายประสาท คือโมเดลทางคณิตศาสตร์ สำหรับประมวลผลสารสนเทศด้วยการคำนวณแบบคอนเนคชันนิสต์ (Connectionist) เพื่อจำลองการทำงานของเครือข่ายประสาทในสมองมนุษย์ ด้วยวัตถุประสงค์ที่จะสร้างเครื่องมือซึ่งมีความสามารถในการ เรียนรู้การจดจำแบบรูป (Pattern Recognition) และการอุปมาความรู้ (Knowledge Deduction) เช่นเดียวกับความสามารถที่มีในสมองมนุษย์แนวคิดเริ่มต้นของเทคนิคนี้ได้มาจากการศึกษาข่ายงานไฟฟ้า ชีวภาพ (Bioelectric Network) ในสมอง ดังภาพที่ 2.14 ซึ่งประกอบด้วย เซลล์ประสาท ชื่อเรียกอีกอย่างหนึ่งที่ใช้กันในอธิบายบ่อย ๆ คือ นิวรอน (neurons) และ จุดประสานประสาท (Synapses) แต่ละเซลล์ประสาทประกอบด้วยปลายในการรับกระแสประสาท เรียกว่า เดนไดรต์ (Dendrite) ซึ่งเป็น Input และปลายในการส่งกระแสประสาทเรียกว่า แอคซอน (Axon) ซึ่งเป็นเหมือน Output ของเซลล์ เซลล์เหล่านี้ทำงานด้วยปฏิกิริยาไฟฟ้าเคมีเมื่อมีการ กระตุ้นด้วยสิ่งเร้าภายนอกหรือกระตุ้นด้วยเซลล์ด้วยกัน กระแสประสาทจะวิ่งผ่านเดนไดรต์เข้าสู่นิวเคลียส (Nucleus) ซึ่งจะเป็นตัวตัดสินใจว่าต้องกระตุ้นเซลล์อื่น ๆ ต่อหรือไม่ ถ้ากระแสประสาทแรงพอ นิวเคลียสก็ จะกระตุ้นเซลล์อื่น ๆ ต่อไปผ่านทางแอกซอน



รูปที่ 2.14 โครงสร้างของเซลล์ประสาทในสมองมนุษย์

(ที่มา: <http://wiki.bethanycrane.com/artificial-neural-networks>)

2.5.1 ประวัติความเป็นมาของโครงข่ายประสาทเทียม

ในปี พ.ศ. 2486 เป็นปีแห่งการกำเนิดของสาขาโครงข่ายประสาทเทียมในวงการวิทยาศาสตร์ โดย แม็กคัลลอค (McCulloch) และพิตส์ (Pitts) ได้เสนอแบบจำลองของเซลล์ประสาท และได้แสดงให้เห็นว่า ในทางทฤษฎีแล้วโครงข่ายของแบบจำลองเซลล์ประสาทดังกล่าวสามารถทำงานเป็นโปรแกรม คอมพิวเตอร์ใด ๆ ก็ได้

ปี พ.ศ. 2492 โดแนลด์ เฮบบ์ (Donald Hebb) ได้เสนอผลงานวิจัยว่า การเรียนรู้ของสมอง สามารถอธิบายได้ด้วยรูปแบบของการประกอบเซลล์ประสาทเข้าด้วยกันเป็นโครงข่าย และได้เสนอกฎการ เรียนรู้ของเฮบบ์ (Hebb's rule) ที่ทำให้โครงข่ายของเซลล์ประสาทเทียมที่แม็กคัลลอค และพิตส์ได้เสนอไว้ สามารถเรียนรู้ปัญหาง่าย ๆ ได้สำเร็จ การเรียนรู้ในแบบของเฮบบ์บนเซลล์ประสาทเทียมของแม็กคัลลอคและ พิตส์นั้นเป็นการเรียนรู้แบบ "ไม่มีผู้สอน" ซึ่งในทางปฏิบัติแล้วโครงข่ายประสาทเทียมที่ทำการเรียนรู้จะพยายามทำการจัดกลุ่มข้อมูลที่โครงข่ายมองว่าคล้ายคลึงกันนำไปไว้ในกลุ่มเดียวกัน ซึ่งไม่เหมาะสมกับ ปัญหาประเภทที่ต้องมีการควบคุมกระบวนการเรียนรู้

ในช่วง พ.ศ. 2490 คอมพิวเตอร์ที่ทำงานเลียนแบบสมองเครื่องแรกของโลกถูกสร้างและ ทดสอบโดยมินสกี(Minsky) ซึ่งได้เสนอผลงานดังกล่าวในปี พ.ศ. 2511 เมื่อคอมพิวเตอร์ดังกล่าวได้รับการ บ้อนตัวอย่างสำหรับการเรียนรู้เข้าไป ก็จะสามารถปรับอัตราการขยายสัญญาณในการ เชื่อมโยงหรือ "ความ แข็งแรงของการเชื่อมโยง" ระหว่างเซลล์ประสาทเทียมได้เองโดยอัตโนมัติ ซึ่งเป็นการแสดงการเรียนรู้ ตัวอย่างที่ถูกบ้อนเข้าไป

ในปี พ.ศ. 2501 แฟรงค์ โรเซนแบลตต์ (Frank Rosenblatt) ได้พัฒนา สถาปัตยกรรมโครงข่าย ประสาทเทียมขึ้น โดยใช้แบบจำลองของแม็กคัลลอคและพิตส์เป็นแนวทาง รวมทั้งเสนอวิธีการเรียนรู้แบบ ใหม่สำหรับสถาปัตยกรรมโครงข่ายประสาทเทียมดังกล่าวด้วย โครงข่ายประสาทเทียมดังกล่าวเรียกว่า เพอร์เซปตรอน (Perceptron) ซึ่งมีการเรียนรู้แบบ "มีผู้สอน" (supervised learning) โดยการปรับความ แข็งแรงของการเชื่อมโยง ซึ่งจะพิจารณาได้จากการ เปรียบเทียบความรู้ของโครงข่ายประสาทเทียมกับ ความรู้ของ "ผู้สอน" (teacher) เพอร์เซปตรอนมีความเหมาะสมกับงานประเภท "การระบุชนิด" ซึ่งใน ระหว่างการเรียนรู้นั้น เพอร์เซปตรอนจะถูก สอนว่าข้อมูลตัวอย่างที่สอนเข้าไปแต่ละแบบนั้น จัดเป็นชนิด ไตบ้าง หากปัญหาและข้อมูลตัวอย่างมีความเหมาะสม เพอร์เซปตรอนจะสามารถระบุชนิดของข้อมูลที่ไม เคยเห็นมาก่อนได้ถูกต้อง

ในช่วงต้น พ.ศ. 2500 เบร์นาร์ด วิโดรว (Bernard Widrow) และมาร์เซียนฮอฟฟ์ (Marcian Hoff) ได้พัฒนาอุปกรณ์ที่เรียกว่าอดาไลน์ (ADALINE; Adaptive Linear combiner) และ กฎ การเรียนรู้แบบใหม่ที่มีประสิทธิภาพสูงเรียกว่า กฎการเรียนรู้ของวิโดรว-ฮอฟฟ์ (Windrow-Hoff learning rule) ที่เป็นการเรียนรู้แบบ "มีผู้สอน" ซึ่งในเวลาต่อมาอุปกรณ์ดังกล่าวได้รับการขยาย แนวคิดไปเป็นมา ดาไลน์ (MADALINE; Many ADALINES) และได้ถูกนำไปประยุกต์ใช้ในการ

รู้จำรูปแบบ (Pattern Recognition) การพยากรณ์อากาศ และระบบควบคุมที่จำเป็นต้องมีการปรับเปลี่ยนระบบไปตาม สภาพแวดล้อมต่าง ๆ

ในขณะที่การค้นคว้าเพื่อพัฒนาคอมพิวเตอร์ที่มี "ปัญญา" ได้ดำเนินไปในแนวทางของการศึกษา และได้พยายามอธิบายการทำงานของระบบประสาทของสิ่งมีชีวิต อีกแนวทางหนึ่งที่สามารถดำเนินควบคู่ กันไปได้ คือแนวทางการค้นคว้าที่พยายามจะอธิบายพฤติกรรมที่เกี่ยวข้องกับปัญญาของมนุษย์ในเชิง โครงสร้างของเหตุและผล โดยใช้สัญลักษณ์ในการแทนแนวคิดของมนุษย์ และดำเนินการกับสัญลักษณ์ เหล่านั้นด้วยกระบวนการในลักษณะของคณิตศาสตร์ผลลัพธ์ที่ได้จากกระบวนการดังกล่าวอาจถือได้ว่าเป็น การตอบสนองที่มนุษย์น่าจะทำได้เงื่อนไขเดียวกัน แนวทางดังกล่าวนำไปสู่พัฒนาการของสาขา ปัญญาประดิษฐ์ (Artificial Intelligence) และจิตวิทยาการเรียนรู้ (Cognitive Psychology) ในช่วง พ.ศ. 2510 และ พ.ศ. 2520

ในปี พ.ศ. 2512 มินสกี (Minsky) และพาเพิร์ต (Papert) ได้เสนอผลงานตีพิมพ์ในรูปแบบของหนังสือ ซึ่งในหนังสือดังกล่าวมีการระบุข้อจำกัดทางทฤษฎีของเพอร์เซปตรอน ผลงานตีพิมพ์ดังกล่าวก่อให้เกิด แนวคิดด้านลบต่อการพัฒนาระบบประมวลผลแบบโครงข่ายของเซลล์ประสาทและเป็นจุดเริ่มต้นจุดหนึ่ง ของกระแสแนวความคิดที่ว่า การคิดของมนุษย์เป็นการประมวลผลแบบลำดับขั้น (Serial Processing) เนื่องจากในช่วงเวลาดังกล่าวไม่มีผู้ที่สามารถแก้ไขข้อจำกัดทางทฤษฎีของโครงข่ายประสาทเทียมได้ วิทยาการด้านโครงข่ายประสาทเทียมจึงไม่ได้รับความสนใจจากวงการคอมพิวเตอร์ในช่วง พ.ศ. 2520 ผลงานวิจัยด้านโครงข่ายประสาทเทียมในช่วงเวลาดังกล่าวจึงเป็นที่รู้จักในวงแคบ ๆ ของผู้ที่สนใจ เท่านั้น

อย่างไรก็ตาม ในช่วงพ.ศ. 2510 มีผลงานวิจัยด้านโครงข่ายประสาทเทียมจากประเทศ ญี่ปุ่น โดย ชุนอิชิ อามาริ (Shun-ichi Amari) ในปี พ.ศ. 2515 และ พ.ศ. 2520 ต่อมาในปี พ.ศ. 2523 คูนิฮิโกะ ฟูกุชิมะ (Kunihiko Fuku-shima) ได้พัฒนาสถาปัตยกรรมโครงข่ายประสาทเทียมที่มีชื่อว่า นีโอค็อกนิตรอน (neocognitrons) สำหรับการรู้จำรูปแบบภาพลักษณ์ โดยเลียนแบบแนวทางการมองเห็น ของสิ่งมีชีวิต

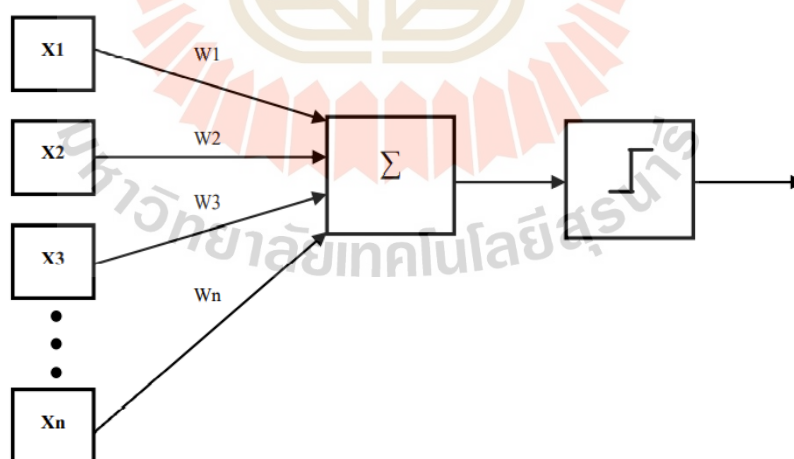
ในปี พ.ศ. 2525 จอห์น ฮอปฟิลด์ (John Hopfield) ได้เสนอสถาปัตยกรรมโครงข่ายประสาท เทียมสำหรับเป็นหน่วยความจำแบบแอสโซซิเอทีฟ (Associative memory) ที่ดึงความจำออกมา โดยใช้ ข้อมูลที่มีความเกี่ยวข้องกับความจำนั้นเป็นตัวชี้แนะ ซึ่งเป็นจุดเริ่มต้นที่ทำให้ นักวิทยาศาสตร์หันกลับมาสนใจ โครงข่ายประสาทเทียมอีกครั้ง ในช่วงเวลาใกล้เคียงกัน สตีเฟน กรอสเบอร์ก (Stephen Grossberg) และ เกล คาร์เพนเตอร์ (Gail Carpenter) ก็ได้เสนอทฤษฎีของการกำทอนแบบอะแดปทีฟ (Adaptive resonance) และได้พัฒนาโครงข่ายประสาทเทียมอาร์ต (ART Network) ซึ่งในปัจจุบันถือว่าเป็นโครงข่าย ประสาทเทียมที่มีสมรรถนะสูงที่สุดแบบหนึ่ง

จากนั้น โครงข่ายประสาทเทียมกลับมาเฟื่องฟูเต็มที่เนื่องจากผลงานตีพิมพ์ของเจมส์ แม็กคลีแลนด์ (James McClelland) และเดวิด รูเมลฮาร์ท (David Rumelhart) ในปีพ.ศ. 2529 ซึ่งเสนอทฤษฎีการเรียนรู้แบบใหม่สำหรับสถาปัตยกรรมโครงข่ายประสาทเทียมที่มีความซับซ้อนขึ้น โดยอิงกับเพอร์เซปตรอน โครงข่ายประสาทเทียมดังกล่าวสามารถก้าวข้ามขีดจำกัดที่เคยถูกเสนอไว้โดยมินสกีและพาเพิร์ต ตั้งแต่ปี พ.ศ. 2512 ได้สำเร็จ อย่างไรก็ตาม มีการค้นพบว่า แนวทางคล้าย ๆ กันได้เคยถูกเสนอแล้วโดย พอล เวอร์โบส (Paul Werbos) ตั้งแต่ปี พ.ศ. 2517 แต่ไม่มีผู้ใดสนใจในช่วงเวลาดังกล่าว ตั้งแต่นั้นมางานวิจัยค้นคว้าด้านโครงข่ายประสาทเทียมก็เกิดขึ้นอีกอย่างมากมายมาจนถึงปัจจุบัน

2.5.2 หลักการทำงานของระบบโครงข่ายประสาทเทียมในคอมพิวเตอร์

จากความคิดเห็นของนักวิจัยที่ทำงานเกี่ยวกับด้านนี้มีความคิดเห็นพ้องตรงกันว่าโครงข่ายประสาทเทียมมีลักษณะโครงสร้างแตกต่างจากโครงข่ายในสมองมนุษย์ แต่ยังคงมีความคล้ายคลึงกัน ในหลักการที่ว่าโครงข่ายประสาทเทียมเป็นการรวมกลุ่มแบบขนานของหน่วยประมวลผลย่อย ๆ ในตรงจุดนี้ทำให้เกิดทักษะสติปัญญาของข่ายงาน แต่ถ้าจะพิจารณาขนาดของสมองจะมีขนาดใหญ่กว่าโครงข่ายประสาท

เทียมอย่างมาก รวมทั้งเซลล์ประสาทยังมีความซับซ้อนกว่า หน่วยย่อยของโครงข่าย อย่างไรก็ตามหน้าที่สำคัญของสมอง เช่น การเรียนรู้ยังคงสามารถถูกจำลองขึ้นอย่างง่ายด้วยโครงข่ายประสาทนี้



รูปที่ 2.15 โครงสร้างของเซลล์ประสาท (Neurons) ในคอมพิวเตอร์
(ที่มา: ผู้ช่วยศาสตราจารย์ ดร. ชูพันธุ์ รัตนโกคา เอกสารความรู้เบื้องต้นทางปัญญาประดิษฐ์)

สำหรับในคอมพิวเตอร์ Neurons ประกอบด้วย Input และ Output เหมือนกัน โดยจำลองให้ Input แต่ละอันมี Weight เป็นตัวกำหนดน้ำหนักของ Input โดย Neuron แต่ละหน่วยจะมีค่า Threshold เป็นตัวกำหนดว่าน้ำหนักรวมของ Input ต้องมากขนาดไหนจึงจะสามารถส่ง Output ไปยัง Neurons ตัวอื่นได้ เมื่อนำ Neuron แต่ละหน่วยมาต่อกันให้ทำงานร่วมกันการทำงานนี้ในทางตรรกะแล้วก็ จะเหมือนกับปฏิกิริยาเคมีที่เกิดในสมอง เพียงแต่ในคอมพิวเตอร์ทุกอย่างเป็นตัวเลขเท่านั้นเอง

การทำงานโครงข่ายประสาท (Neural Network) มีหลักการโดยมี Input บ้อนไปยังเครือข่าย (Network) หลังจากนั้นนำ Input มาคูณกับ Weight ของและหา ผลที่ได้จาก Input ทุก ๆ ค่าของ Neuron จะเอามารวมกันแล้วก็เอามาเทียบกับ Threshold ที่กำหนดไว้ ถ้าผลรวมมีค่ามากกว่า Threshold แล้ว Neuron ก็ จะส่ง Output ออกไป Output นี้ก็จะถูกส่งไปยัง Input ของ Neuron อื่น ๆ ที่เชื่อมกันใน Network ถ้า ค่าน้อยกว่า Threshold ก็จะไม่เกิด Output ซึ่งแสดงได้ดังตัวโปรแกรมเทียมนี้

$$\text{if } (\text{sum}(\text{input} * \text{weight}) > \text{threshold}) \text{ then output} \quad (2.11)$$

สิ่งสำคัญคือต้องทราบค่า Weight และ Threshold สำหรับสิ่งที่ใช้ต้องการเพื่อให้คอมพิวเตอร์ รู้จำซึ่งเป็นค่าที่ไม่แน่นอน แต่สามารถกำหนดให้คอมพิวเตอร์ปรับค่าเหล่านั้น ได้โดยการสอนให้ คอมพิวเตอร์รู้จัก Pattern ของสิ่งที่ต้องการให้คอมพิวเตอร์รู้จำเรียกว่า "Back Propagation" ซึ่งเป็น กระบวนการย้อนกลับของการรู้จำในการฝึก Feed-Forward Neural Networks จะมีการใช้อัลกอริทึม แบบ Back-Propagation เพื่อใช้ในการปรับปรุงน้ำหนักคะแนนของเครือข่าย (Network Weight) หลังจากใส่รูปแบบข้อมูลสำหรับฝึกให้แก่เครือข่ายในแต่ละครั้งแล้ว ค่าที่ได้รับ (Output) จากเครือข่ายจะถูกนำไปเปรียบเทียบกับผลที่คาดหวัง แล้วทำการคำนวณหาความผิดพลาด ซึ่งค่าความผิดพลาดนี้จะถูก ส่งกลับเข้าสู่เครือข่ายเพื่อใช้แก้ไขค่าน้ำหนักคะแนนต่อไป

2.5.3 การเรียนรู้ของโครงข่ายประสาทเทียม

ผลลัพธ์ของการดำเนินงานวิจัยที่เกี่ยวข้องกับพฤติกรรมศาสตร์ พบได้ว่าการเรียนรู้ของสิ่งมีชีวิตแต่ละชนิด มีกระบวนการ แตกต่างหลายๆแบบ ซึ่งขึ้นอยู่กับความเหมาะสมของสิ่งมีชีวิตแต่ละสายพันธุ์ ในสาขาโครงข่ายประสาทเทียม แนวคิดเกี่ยวกับกระบวนการเรียนรู้จะประยุกต์มาจากผลการศึกษาทางพฤติกรรมศาสตร์ อาจกล่าวอีกอย่างหนึ่งว่าการเรียนรู้คือ กระบวนการที่มีระบบประสาทสามารถปรับตัวเองพร้อมทั้งตอบสนองตามสิ่งเร้าที่ถูกกระตุ้น จนกระทั่งสามารถให้ผลตอบได้ตามต้องการ

การเรียนรู้พิจารณาได้ว่า เป็นขั้นตอนกระบวนการจัดชนิดสิ่งเร้าที่กระตุ้น เมื่อได้สัมผัสกับสิ่งเร้าโดยที่ระบบประสาทคุ้นเคยรู้จักกับสิ่งนั้นแล้ว สามารถให้การตอบสนองตามที่เคยได้

พบและเรียนรู้ก่อนหน้านี้ ในกรณีที่มีสิ่งเร้ากระตุ้นแต่ยังไม่เกิดปฏิกิริยาการตอบสนอง ประสาทเทียม จะทำความเข้าใจในการแบ่งแยกชนิด ถ้ากล่าวถึงในทางปฏิบัติ ระบบประสาทของ สิ่งมีชีวิตจะปรับ ความเหนียวแน่นของการเชื่อมต่อที่ซินแนปส์ จนสร้างผลตอบสนองต่อสิ่งเร้าได้ตามที่ต้องการ สถานะที่ กระบวนการของการเรียนรู้ก็จะสิ้นสุดลง เป็นสถานะที่ถือว่าระบบประสาทได้รับความรู้ไปแล้วคำ จำกััดความของกระบวนการเรียนรู้หมายถึงขั้นตอนต่อไป

ขั้นที่ 1 โครงข่ายประสาทถูกกระตุ้นด้วยสิ่งแวดลอม

ขั้นที่ 2 โครงข่ายประสาทเกิดการเปลี่ยนแปลง อันเป็นผลมาจากการกระตุ้นดังกล่าว

ขั้นที่ 3 โครงข่ายประสาทตอบสนองต่อสิ่งแวดลอมในแนวทางใหม่ อันเป็นผลมาจากการ เปลี่ยนแปลงที่เกิดขึ้นในโครงสร้างภายในโครงข่าย

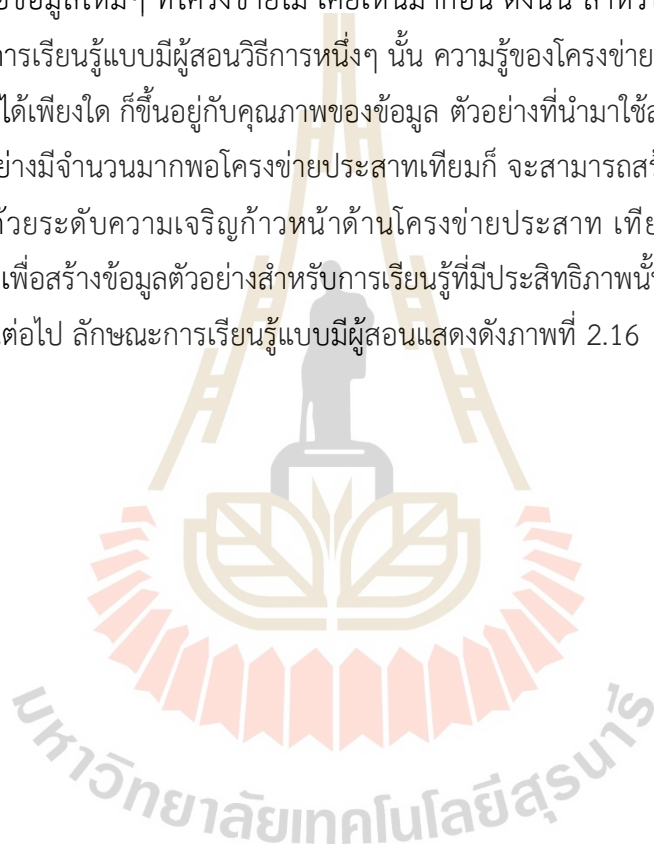
เมื่อพิจารณาไปที่เฉพาะบริเวณหนึ่ง ๆ ของโครงข่ายประสาทจะพบว่า การเชื่อมต่อ ของเซลล์ ประสาทที่บริเวณต่าง ๆ นั้น สามารถมีรูปแบบที่แตกต่างกันได้หลายๆ แบบ และ กระบวนการเรียนรู้ของแต่ละบริเวณก็ไม่เหมือนกันด้วย ในทำนองเดียวกัน เทคนิคการเรียนรู้ของ โครงข่ายประสาทเทียมก็แตกต่างกัน ไปสำหรับแต่ละชนิดของโครงข่าย

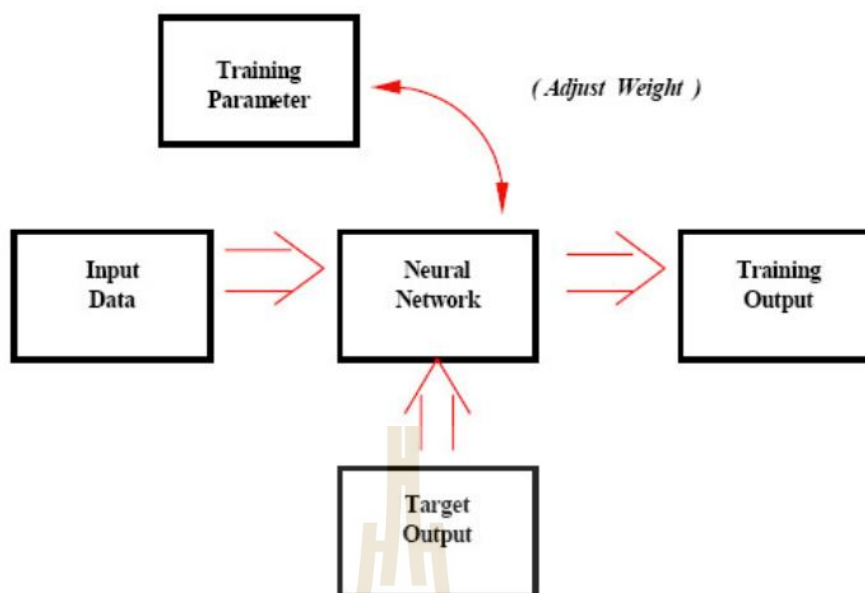
การสอนโครงข่ายให้ทำการคำนวณข้อมูล Output พร้อมปรับปรุ้งค่าน้ำหนักโดยใช้ ข้อมูล Input ที่ ป้อนให้กับโครงข่ายโดยอาศัยกระบวนการทำซ้ำสามารถแบ่งเป็น 2 ประเภท คือ

(1) การเรียนรู้แบบมีผู้สอน (Supervised Learning) รูปแบบการเรียนรู้แบบมีผู้สอน เริ่มด้วยการส่งสิ่งเร้าที่ใช้ในการสอนเข้าไปเป็นอินพุต (Input) ในโครงข่ายประสาทเทียม เพื่อให้ โครงข่ายประสาทเทียมสร้างผลตอบออกมาเป็นเอาต์พุต (Output) ซึ่งผลตอบจะเป็นอย่างไร ก็ขึ้นอยู่กับ สถานะในตอนเริ่มต้นเรียนรู้ของโครงข่ายประสาทเทียม ผลตอบดังกล่าวจะถูกนำมาเปรียบเทียบกับ ผลตอบเป้าหมาย (Target Response) ซึ่งผู้สอน (Teacher) จะ เป็นผู้สร้างขึ้น หากผลตอบทั้งสองมี ความแตกต่างกัน นั่นคือมีความคลาดเคลื่อน (Error) เกิดขึ้นความคลาด เคลื่อนดังกล่าวจะถูกนำไป คำนวณการปรับแต่งค่าน้ำหนักต่าง ๆ ในโครงข่ายประสาทเทียม เพื่อลดความ คลาดเคลื่อนลงให้ เหลือน้อยที่สุด

การปรับแต่งค่าน้ำหนักโดยพิจารณาจากความคลาดเคลื่อนนี้ จะขึ้นอยู่กับกฎการ เรียนรู้ หรือขั้นตอนการคำนวณซึ่งเรียกว่า "อัลกอริทึม" (Algorithm) ที่แตกต่างกันโดยแต่ละ อัลกอริทึมจะมี คุณสมบัติและสมรรถนะแตกต่างกัน วิธีการส่วนใหญ่ของการเรียนรู้แบบมีผู้สอนนี้ จะดัดแปลงมาจาก วิธีการทางคณิตศาสตร์ในเรื่องของเทคนิคการหาค่าเหมาะสม (Optimization Technique)

เนื่องจากเจตนาของมนุษย์ในการพัฒนาเครื่องมือขึ้นมาใช้งาน นั้น จะอิงอยู่กับการที่มนุษย์ ต้องการสั่งการและควบคุมเครื่องมือ นั้น ๆ ให้ทำงานได้ตามที่ต้องการ จึงทำให้โครงข่ายประสาทเทียม ประเภทที่ใช้การเรียนรู้แบบมีผู้สอนได้รับความนิยมในการนำไปประยุกต์ใช้มากที่สุด เนื่องจากเป็นแบบที่ สามารถควบคุมได้ การสั่งการโครงข่ายประสาทเทียมจะเป็นไปโดยทางอ้อม ในลักษณะของการฝึกสอน โครงข่ายประสาทเทียม โดยการสร้างข้อมูลตัวอย่าง(รวมทั้งค่าเป้าหมาย) ที่จะให้โครงข่ายเรียนรู้ เมื่อโครงข่ายประสาทเทียมเรียนรู้ข้อมูลตัวอย่างได้ถูกต้องหมดแล้ว ความรู้ที่โครงข่ายประสาทเทียมได้เก็บไว้ใน ลักษณะของค่าน้ำหนักต่าง ๆ จะเป็นสิ่งที่ถูกนำไปใช้งานจริงเพื่อสร้างผลตอบต่อข้อมูลใหม่ๆ ที่โครงข่ายไม่เคยเห็นมาก่อน ดังนั้น สำหรับโครงข่ายประสาทเทียมหนึ่ง ๆ และวิธีการเรียนรู้แบบมีผู้สอนวิธีการหนึ่งๆ นั้น ความรู้ของโครงข่ายประสาทเทียมจะสามารถนำไปใช้งานจริงได้เพียงใด ก็ขึ้นอยู่กับคุณภาพของข้อมูล ตัวอย่างที่นำมาใช้สอนนั้น อาจจะสามารถกล่าวได้ว่า หากข้อมูลตัวอย่างมีจำนวนมากพอโครงข่ายประสาทเทียมก็ จะสามารถสร้างความรู้ได้อย่างถูกต้อง อย่างไรก็ตามด้วยระดับความเจริญก้าวหน้าด้านโครงข่ายประสาท เทียมในปัจจุบันเทคนิคการวิเคราะห์ปัญหาเพื่อสร้างข้อมูลตัวอย่างสำหรับการเรียนรู้ที่มีประสิทธิภาพนั้น ยังคงเป็นงานวิจัยที่ต้องมีการค้นคว้ากันต่อไป ลักษณะการเรียนรู้แบบมีผู้สอนแสดงดังภาพที่ 2.16





รูปที่ 2.16 การเรียนรู้แบบมีผู้สอน (Supervised Learning)

(ที่มา: ผู้ช่วยศาสตราจารย์ ดร.ชุพันธ์ รัตนโกคา เอกสารความรู้เบื้องต้นทางปัญญาประดิษฐ์)

(2) การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) การเรียนรู้แบบนี้ไม่จำเป็นต้องมีค่าเป้าหมายของตัวอย่างข้อมูล ในช่วงทำการเรียนรู้โครงข่ายประสาทเทียมจะรับข้อมูลกระตุ้นในรูปแบบต่าง ๆ และจัดทำกลุ่มรูปแบบต่าง ๆ เหล่านั้นด้วยตนเอง ผลลัพธ์ของโครงข่ายประสาทเทียมที่ได้รับจะเป็นการระบุกลุ่มข้อมูลที่ถูกป้อนเข้าไป ซึ่งอ้างอิงกับวิธีการจัดกลุ่มกับข้อมูลโครงข่ายที่เคยเรียนรู้

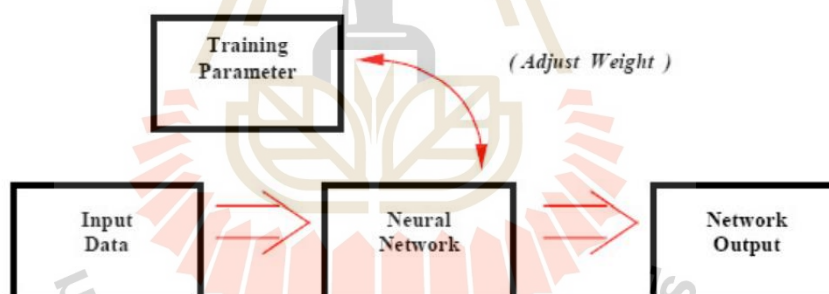
ตัวอย่างการเรียนรู้แบบนี้ในมนุษย์คือ การให้เด็กเล็ก ๆ จัดเก็บสิ่งของไว้บนชั้นวางของให้ เป็นระเบียบเรียบร้อย สมมุติว่า เด็กคนหนึ่งเลือกเก็บหนังสือต่าง ๆ ไว้ที่ชั้นบน เก็บตุ๊กตาไว้ที่ชั้นล่างและเก็บ ของเล่นอื่น ๆ ไว้ที่ชั้นกลาง ๆ หลังจากนั้น หากเด็กคนนั้นซื้อตุ๊กตามาใหม่ ก็นำไปเก็บไว้ที่ชั้นล่าง เป็นต้น

แม้ว่าการเรียนรู้แบบไม่มีผู้สอนนี้จะไม่ต้องการผู้สอน แต่ก็ต้องการแนวทางในการจัดกลุ่ม เช่น การจัดกลุ่มอาจจะจัดตามรูปทรง สีหรือวิธีการใช้งานของวัตถุต่าง ๆ ที่จะนำมาจัด เป็นต้น ดังนั้น หากไม่ มีการให้แนวทางที่ชัดเจนว่า การจัดกลุ่มควรเป็นไปตามคุณลักษณะใด การจัดกลุ่มอาจไม่ประสบความสำเร็จในแง่การนำมาใช้งานจริงก็ได้ ตัวอย่างเช่น การให้เด็กจัดของไว้บนชั้นวางของนั้น เด็กอาจจะจัด ตามใจชอบและไม่เป็นหมวดหมู่ ทำให้ไม่สะดวกต่อการนำสิ่งของใหม่ๆ เข้าไปเก็บรวมด้วยก็ได้ การใช้งาน โครงข่ายประสาทเทียมที่ใช้การเรียนรู้แบบนี้ จึงมักต้องมีการดำเนินการปรับแต่งข้อมูล เพื่อให้เกิดการเน้น สภาพของคุณลักษณะสำคัญ ที่ต้องการนำมาเป็น

แนวทางในการจัดกลุ่มให้เด่นชัดขึ้นหรืออาจเป็นการปรับ กฎการเรียนรู้ เพื่อให้เน้นไปที่คุณลักษณะที่ต้องการก็ได้

เนื่องจากขั้นตอนการเรียนรู้แบบไม่มีผู้สอนนี้จะมีการระบุกลุ่มของข้อมูลตัวอย่างก่อน เมื่อตัดสินใจได้แล้วว่า ข้อมูลใหม่มีลักษณะที่ควรจัดรวมเข้ากลุ่มใด (หรืออาจถือเป็นกลุ่มใหม่ก็ได้ในกรณีที่ เห็นว่าไม่ควรจัดเข้ากลุ่มใดเลย) จะทำการปรับคุณลักษณะกลุ่มโดยนำข้อมูลใหม่ที่ได้รับเข้ามาเป็นตัวแปรช่วยกำหนดแนวทางจัดกลุ่มด้วย การตัดสินใจว่าข้อมูลใหม่ให้อยู่ในกลุ่มไหน หลักการทำงานของโครงข่ายประสาทเทียมบางชนิดใช้วิธีการแข่งขันของกลุ่มว่ากลุ่มใดควรได้ข้อมูลดังกล่าว หลักการแบบนี้เรียกว่า การเรียนรู้แบบแข่งขันกัน (Competitive Learning)

ในอีกแง่มุมหนึ่งนับจากจุดเริ่มต้นของการเรียนรู้ซึ่งไม่มีการจัดกลุ่มข้อมูลในแบบใด ๆ เลย จนถึงเวลาที่มีการจัดเสร็จสิ้นแล้ว จะพบว่า การจัดกลุ่มข้อมูลเกิดขึ้นตามคุณลักษณะบางอย่างของข้อมูล ตัวอย่าง ซึ่งการจัดกลุ่มนี้เกิดจากการที่โครงข่ายประสาทเทียมประเมินข้อมูลต่าง ๆ ที่ถูกป้อนเข้าไปใน ระหว่างการเรียนรู้ จนสร้างเป็นวิธีการจัดกลุ่มขึ้นมาได้ ดังนั้น การเรียนรู้ในลักษณะดังกล่าวจึงถูกเรียกว่า การเรียนรู้แบบจัดตัวเอง (Self-Organizing) ด้วย ลักษณะการเรียนรู้แบบไม่มีผู้สอนแสดงดังรูปที่ 2.17



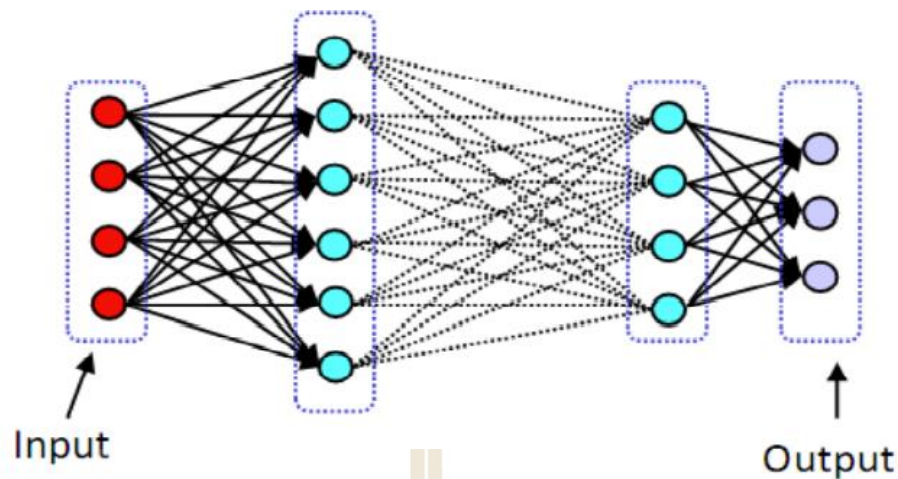
รูปที่ 2.17 การเรียนรู้แบบไม่มีผู้สอน Unsupervised Learning

(ที่มา: ผู้ช่วยศาสตราจารย์ ดร. ชูพันธุ์ รัตนโกศา เอกสารความรู้เบื้องต้นทางปัญญาประดิษฐ์)

2.5.4 สถาปัตยกรรมโครงข่ายประสาทเทียม

สถาปัตยกรรมโครงข่ายประสาทเทียม สามารถแบ่งเป็น 4 ประเภทใหญ่ ๆ คือ

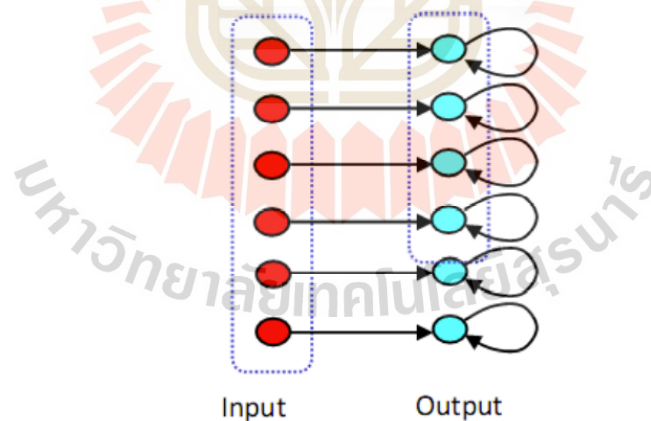
(1) Feed-Forward Network คือรูปแบบข้อมูลที่ประมวลผลในวงจรเครือข่ายโดยส่งไปในทิศทางเดียวกันจากโหนด Input และนำส่งต่อเรื่อยมาจนถึงโหนด Output ซึ่งการดำเนินการแบบนี้จะไม่มี การส่งย้อนกลับของ หรือแม้แต่โหนดในชั้น (Layer) เดียวกันก็ไม่มี การเชื่อมต่อกัน ดังแสดงในรูปที่ 2.18



รูปที่ 2.18 สถาปัตยกรรมของ Feed-Forward Network

(ที่มา: ผู้ช่วยศาสตราจารย์ ดร. ชูพันธุ์ รัตนโกคา เอกสารความรู้เบื้องต้นทางปัญญาประดิษฐ์)

(2) Feed-Back Network คือข้อมูลจากภาพประมวลผลในเครือข่าย จะทำการป้อนกลับกลับเข้าไปยังเครือข่ายหลายๆครั้ง เพื่อให้ได้ผลลัพธ์คำตอบ (บางที่เรียกว่า Recurrent Network) โครงข่ายมีลักษณะดังแสดงในรูปที่ 2.19



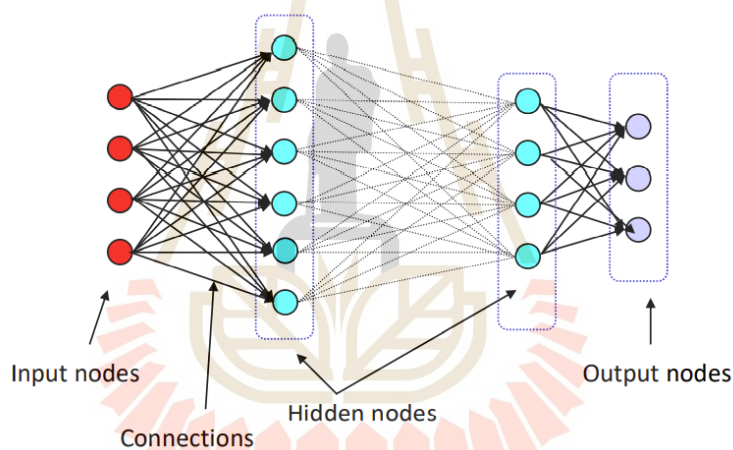
รูปที่ 2.19 สถาปัตยกรรมของ Feed-Back Network

(ที่มา: ผู้ช่วยศาสตราจารย์ ดร. ชูพันธุ์ รัตนโกคา เอกสารความรู้เบื้องต้นทางปัญญาประดิษฐ์)

(3) Network Layer เป็นพื้นฐานที่สำคัญของโครงข่ายประสาทเทียม ซึ่งโดยหลักพื้นฐานประกอบ 3 ส่วน ได้แก่ ชั้นของ Input Units เชื่อมโยงกับ Hidden Units และเชื่อมโยง Output Units ตามลำดับ

- Input Units รับข้อมูลเพื่อส่งต่อเข้าไปในเลเยอร์ถัดไป
- Hidden Units จะถูกกำหนด โดยการทำงานของ Input Units และค่าน้ำหนักบนความสัมพันธ์ระหว่าง Input Units และ Hidden Units
- Output Units จะขึ้นอยู่กับการทำงานของ Hidden Units และค่าน้ำหนักระหว่าง Hidden Units และ Output Units

ประเภทของเครือข่ายนี้เป็นที่น่าสนใจ เพราะเราสามารถกำหนดการแทนค่าให้แก่ Input Units ได้อย่างอิสระ ค่าน้ำหนักระหว่าง Input Units และ Hidden Units จะถูกกำหนดเมื่อ Hidden Unit กำลังทำงานเพราะฉะนั้นเวลาที่แก้ไขค่าน้ำหนัก Hidden Units จะสามารถเลือกได้ว่าอะไรคือค่าที่เราแทนเข้ามา ดังแสดงในรูปที่ 2.20



รูปที่ 2.20 สถาปัตยกรรมของ Network Layer

(ที่มา: ผู้ช่วยศาสตราจารย์ ดร. ชูพันธุ์ รัตนโกศา เอกสารความรู้เบื้องต้นทางปัญญาประดิษฐ์)

(4) Architecture of Layer สามารถจำแนกสถาปัตยกรรมของชั้น (Layer) ออกเป็น 2 ประเภท

- Single-Layer Perceptron เครือข่ายประสาทที่ประกอบด้วยชั้นเพียงชั้นเดียว จำนวนโหนด Input ขึ้นอยู่กับจำนวน Components ของ Input Data และ Activation Function ขึ้นอยู่กับลักษณะข้อมูลของ Output ดังแสดงตามด้านล่าง เช่น ถ้า Output ที่ต้องการเป็น “ใช่” หรือ “ไม่ใช่” เราจะต้องใช้ Threshold Function

$$f(x) = \begin{cases} 1 & \text{if } x \geq T \\ 0 & \text{if } x < T \end{cases} \quad T = \text{Threshold level} \quad (2.12)$$

หรือถ้า Output เป็นค่าตัวเลขที่ต่อเนื่อง เราต้องใช้ Continuous Function เช่น Sigmoid Function

$$f(x) = \frac{1}{1 + e^{-\alpha x}} \quad (2.13)$$

- Multi-layer Perceptron เครือข่ายประสาทจะประกอบด้วยหลายชั้นโดยในแต่ละชั้นจะประกอบด้วย โหนด (Nodes) หรือเปรียบได้กับตัวเซลล์ประสาท (Neurons) ค่าน้ำหนักของเส้นที่เชื่อมต่อระหว่างโหนด ของแต่ละชั้น (เมทริก W), ค่า Bias Vector (b) และค่า Output Vector (a) โดย m เป็นตัวเลขบอกลำดับชั้นกำกับไว้ด้านบน เมื่อ p เป็น Input Vector การคำนวณค่าเอาต์พุตสำหรับเครือข่ายประสาทที่มี M ชั้นจะเป็นดังสมการ

$$\begin{aligned} \text{เมื่อ } & m = 0, 2, \dots, M-1 \\ & a^0 = p \\ & a = a^m \end{aligned} \quad (2.14)$$

และ f เป็น Transfer Function

$$a^{m+1} = f^{m+1}(W^{m+1}a^m + b^{m+1})$$

2.5.5 ประโยชน์ของโครงข่ายประสาทเทียม

- เกิดข้อผิดพลาดได้ยาก (Fault Tolerance) เนื่องจากโครงข่ายประสาทเทียมประกอบด้วยโครงข่ายที่ใช้ประมวลผลมีจำนวนมากมายหลากหลายโครงข่าย หากเกิดความผิดพลาดจากโครงข่ายใดโครงข่ายหนึ่งจะไม่ทำให้ระบบเกิดข้อผิดพลาด

- ความสามารถในการหาเหตุผล (Generalization) เมื่อระบบโครงข่ายประสาทเทียมได้รับข้อมูลนำเข้าที่ไม่ครบถ้วนหรือไม่เพียงพอต่อการ หาข้อสรุป หรือได้รับข้อเท็จจริงที่ไม่เคยได้รับมาก่อน ระบบจะสามารถลำดับการเชื่อมโยงข้อเท็จจริงจน สามารถให้ข้อสรุปและเหตุผลได้

- ความสามารถในการปรับเปลี่ยน (Adaptability) โครงข่ายประสาทเทียมสามารถเรียนรู้สภาพแวดล้อมใหม่ได้ ดังนั้นเมื่อมีเหตุการณ์ใหม่ๆ เข้าสู่ระบบก็จะสามารถปรับเปลี่ยนหรือปรับปรุงองค์ความรู้ให้ทันสมัยตามเหตุการณ์ใหม่นั้น

- ความสามารถในการปรับเปลี่ยน (Adaptability) โครงข่ายประสาทเทียมสามารถเรียนรู้สภาพแวดล้อมใหม่ได้ ดังนั้น เมื่อมีเหตุการณ์ใหม่ ๆ เข้าสู่ระบบก็จะสามารถปรับเปลี่ยนหรือปรับปรุงองค์ความรู้ให้ทันสมัยตามเหตุการณ์ใหม่ นั้น

- ความสามารถในการพยากรณ์ (Forecasting Capability) โครงข่ายประสาทเทียมสามารถนำข้อมูลทางสถิติเดิมที่มีอยู่ในระบบ มาใช้คาดการณ์ หรือพยากรณ์ข้อมูลในอนาคตได้

2.5.6 การประยุกต์ใช้งานโครงข่ายประสาทเทียม

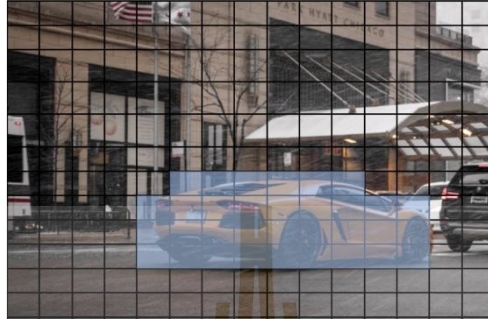
โครงข่ายงานระบบประสาท(Neural Network) มีทักษะความสามารถการเลียนแบบจำลองพฤติกรรมทางกายภาพของระบบที่มีความซับซ้อนจากข้อมูลที่ถูกป้อน ดังนั้น จึงเป็นตัวเลือกใหม่ที่ใช้ในการควบคุมและนำไปประยุกต์ในงาน ๆ ได้แก่

- การจดจำรูปแบบมีลักษณะแตกต่างไม่มีความคล้ายคลึง เช่น ลายมือ ลายเซ็น ตัวอักษร รูปภาพ
- การประมาณค่าฟังก์ชันหรือการประมาณความสัมพันธ์
- งานที่สิ่งแวดล้อมเปลี่ยนแปลงอยู่เสมอ(วงจรรข่ายนิวรอนสามารถปรับตัวเองได้)
- การจัดหมวดหมู่และแยกแยะสิ่งของ
- การทำนาย ได้แก่ สภาพอากาศจากข้อมูลสถิติหรือการขยับตัวของหุ้น
- งานประยุกต์ใช้ข่ายงานระบบประสาทควบคุมกระบวนการทางเคมีโดยวิธีพยากรณ์แบบจำลอง (Model Predictive Control)
- การประยุกต์ใช้ข่ายงานระบบประสาทแบบแพร่กระจายกลับในการทำนายพลังงานความร้อน ที่สะสมอยู่ในตัวอาคาร

2.6 อัลกอริทึมพื้นฐานของYolov5

YOLOเข้ามามีบทบาทในอุตสาหกรรมคอมพิวเตอร์และเป็นที่รู้จักในวงกว้างจากการตีพิมพ์และเผยแพร่โดยJoseph Redmon ในชื่อบทความคือ “You Only Look Once” ที่เป็นการตรวจแบบเรียลไทม์ หลังจากนั้น ได้ถูกนำเสนอได้รับความสนใจเป็นอย่างมากจากกลุ่มนักวิจัยComputer Vision โดยการประยุกต์ใช้โครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Networks (CNN)) โดยจะเป็นลักษณะการสร้างกล่อง (Bounding Boxes) ขึ้นมาในภาพที่เราป้อนข้อมูลเข้าไป

หลักการของYolo คือ การประยุกต์ใช้Grid Cellที่มีขนาด $S \times S$ ป้อนเข้าไปในภาพที่ถูก Input เข้าไป แสดงในรูปที่ 2.21 ถ้าตรงตำแหน่งตรงกลางของวัตถุตรงกับGrid Cell ตัวGrid Cell จะทำการตรวจจับวัตถุนั้น ส่วนGrid Cellอื่น ๆ ที่ไม่มีวัตถุตรงกันจะไม่แสดงตัวแปร ดังนั้นจะมีค่าเป็น 0

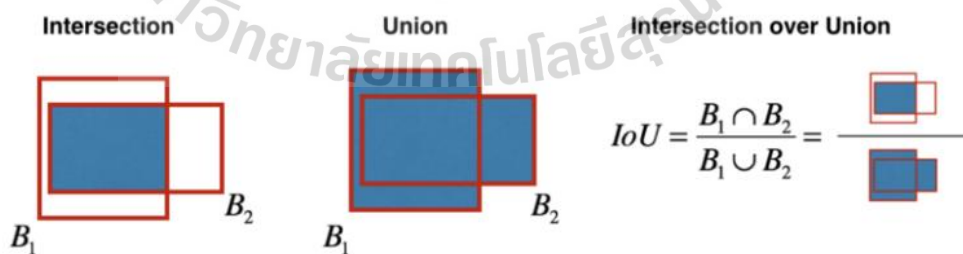


รูปที่ 2.21 Grid Cellในภาพเพื่อทำการค้นหาวัตถุ (15 x 12)

(ที่มา: <https://towardsdatascience.com/mask-detection-using-yolov5-ae40979227a6>)

เพื่อที่จะตรวจจับวัตถุ แต่Grid Cellจะทำนายค่าตัวแปร B bounding boxes กับค่าตัวแปร พารามิเตอร์ และค่า Confidence สำหรับ boxes นั้น ค่า Confidence จะมีผลต่อการแสดงตัวแปร ของวัตถุว่ามีหรือไม่มีใน bounding box นั้น ค่า Confidence แสดงตามสมการรูป 2.22 ด้านล่าง

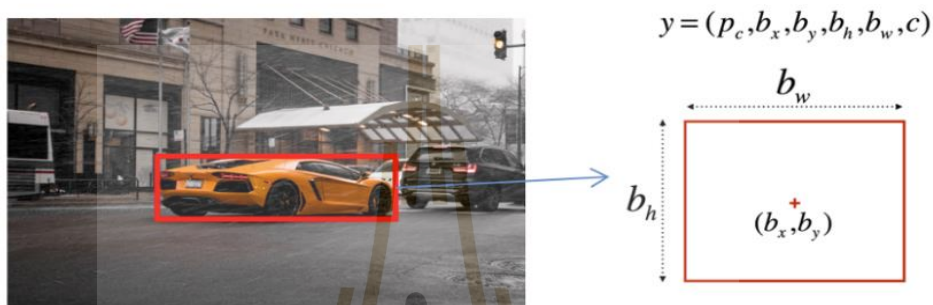
$$\text{confidence score} = p(\text{Object}) * IOU_{pred}^{\text{truth}} \quad (2.15)$$



รูปที่ 2.22 สมการ Intersection over Union

(ที่มา: <https://www.oreilly.com/library/viewhttps://www.oreilly.com/library/view/hands-on-convolutional-neural/9781789130331/a0267a8a-bd4a-452a-9e5a-8b276d7787a0.xhtml>)

ค่า $p(\text{Object})$ คือ ค่าความน่าจะเป็นที่มีวัตถุอยู่ใน cell และค่า IOU คือส่วนที่ intersection ผ่าน union ของ box ที่ได้ทำนายค่าออกมา ค่า $p(\text{Object})$ จะมีค่าอยู่ระหว่าง 0 ถึง 1 ถ้าค่า confidence มีค่าเท่ากับ 0 จะสรุปได้ว่าไม่มีวัตถุอยู่ใน grid cell นั้น นอกจากนี้ แต่ละ bounding box จะประกอบด้วยค่าตัวแปรพารามิเตอร์ 6 ตัว ($p_c, b_x, b_y, b_w, b_h, c$) ตัวแปรค่า b_x กับ b_y จะเป็นจุดศูนย์กลาง coordinate b_w คือ ความกว้าง และ b_h คือ ความสูงของ bounding box รวมถึงค่า confidence ที่รวมอยู่ด้วยคือค่า p_c และ c คือ คลาสของรูปภาพ ตามรูปที่ 2.23

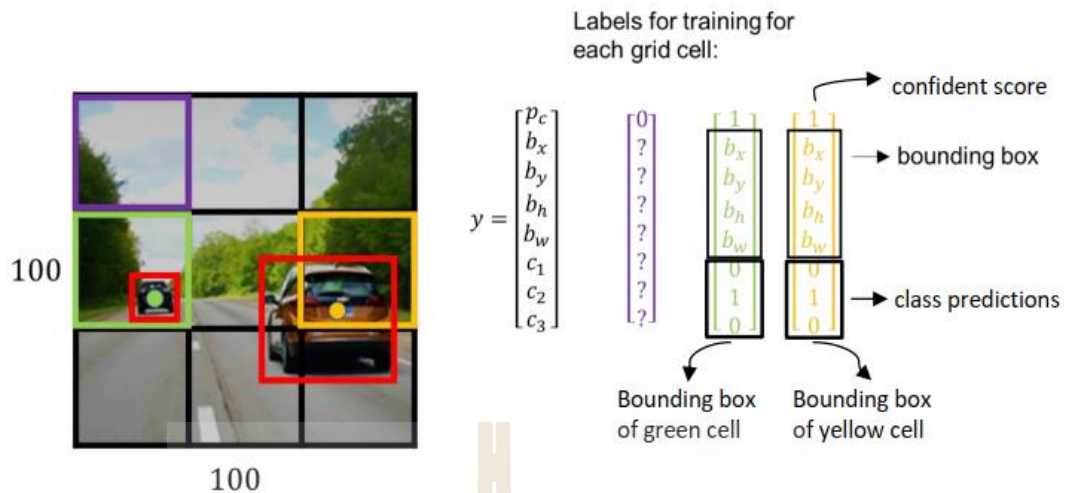


รูปที่ 2.23 พารามิเตอร์ของ bounding box

(ที่มา: <https://towardsdatascience.com/mask-detection-using-yolov5-ae40979227a6>)

ดังนั้น โมเดลมี grid cell ขนาด $S \times S$ ของภาพ โดยแต่ละ cell จะทำนาย B bounding boxes ซึ่งประกอบด้วย 5 พารามิเตอร์ และตัวแปรความน่าจะเป็นของคลาสนี้คือ c ทั้งหมดของ Yolo output ของพารามิเตอร์โมเดลจะเป็น $S \times S \times (5 * B + C)$ ยกตัวอย่างเช่น จากข้อมูลที่ถูกนำมาทดสอบ COCO dataset ประกอบด้วย 80 คลาส และแต่ละเซลล์ทำนาย 2 anchor สรุปพารามิเตอร์ output คือ $7 \times 7 \times (5 * 2 + 80)$

วัตถุประสงค์อัลกอริทึมของ Yolo คือการตรวจจับวัตถุด้วยการทำนายที่แม่นยำของ bounding box ที่ประกอบด้วยวัตถุและตำแหน่งของวัตถุที่อ้างอิงตามจุดพิกัด จากตัวอย่างรูปที่ 2.24 โมเดลมี 2 anchor อยู่ใน grid cell 3×3 แสดงเป็น vector ของกรอบสี่เหลี่ยม และสี่เหลี่ยมตามด้านล่าง



รูปที่ 2.24 โมเดล Yolo ขนาดgrid cell 3 x3 ทำนายวัตถุประกอบด้วย 3 คลาส
(ที่มา: <https://ai-pool.com/a/s/yolov3-and-yolov4-in-object-detection>)

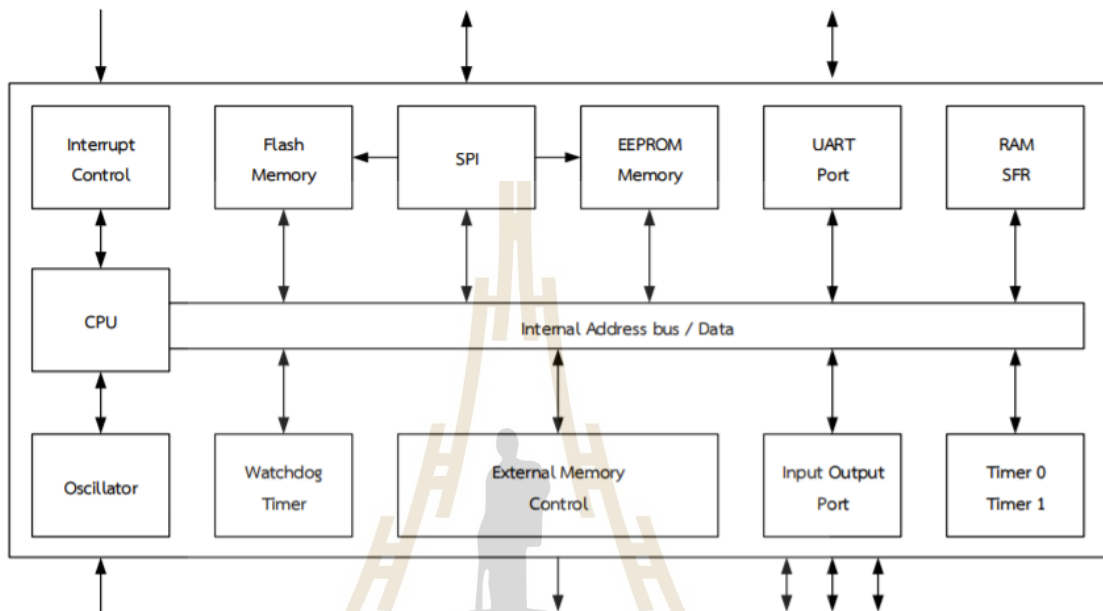
2.7 ไมโครคอนโทรลเลอร์เบื้องต้น

ปัจจุบันไมโครคอนโทรลเลอร์ถูกนำมาใช้อย่างกว้างขวางในงานด้านอิเล็กทรอนิกส์ เพราะใช้งานได้ง่าย สะดวก มีขนาดเล็ก ราคาถูก และที่สำคัญสามารถสั่งงานให้ทำงานตามเงื่อนไขใหม่ตามผู้ใช้งานโดยไม่จำเป็นต้องเปลี่ยนวงจรภายนอก มีบริษัทผู้ผลิตไอซีไมโครคอนโทรลเลอร์ออกมาหลายตระกูล เช่น Z80, PIC, MCS51, AVR, ARM7, ARM9, 68HC แต่ละตระกูลก็จะมีคุณสมบัติต่างกันออกไป

2.7.1 โครงสร้างไมโครคอนโทรลเลอร์

ไมโครคอนโทรลเลอร์ (Microcontroller) มาจากคำว่า ไมโคร (Micro) หมายถึงขนาดเล็ก ส่วนคำว่าคอนโทรลเลอร์ (Controller) หมายถึงตัวควบคุมหรืออุปกรณ์ควบคุม ไมโครคอนโทรลเลอร์ จึงหมายถึงอุปกรณ์ควบคุมขนาดเล็กที่สามารถควบคุมอุปกรณ์ภายนอกได้ แต่ในตัวอุปกรณ์ควบคุม ขนาดเล็กนี้ได้บรรจุความสามารถที่คล้ายคลึงกับระบบคอมพิวเตอร์ กล่าวคือ ไมโครคอนโทรลเลอร์ ได้รวมเอาซีพียู (CPU) หน่วยความจำ (Memory) อินพุตเอาต์พุตพอร์ต (I/O Port) รีจิสเตอร์ (Register) วงจรผลิตสัญญาณ (Oscillator) และวงจรอิเล็กทรอนิกส์ เช่น วงจรแปลงจากสัญญาณอนาล็อกเป็นดิจิทัลวงจรสร้างสัญญาณ PWM (มีเฉพาะบางเบอร์) ไมโครคอนโทรลเลอร์ จึงสามารถนำไปประยุกต์ใช้งานควบคุมต่าง ๆ ได้ดี

ไมโครคอนโทรลเลอร์แต่ละเบอร์แต่ละตระกูลนั้น มีโครงสร้างภายในแตกต่างกัน ขึ้นอยู่กับ บริษัทผู้ผลิตและจุดประสงค์ของการสร้างไมโครคอนโทรลเลอร์เบอร์นั้น ๆ พื้นฐานของโครงสร้าง ภายในของไมโครคอนโทรลเลอร์มีส่วนประกอบดังรูปที่ 2.25



รูปที่ 2.25 โครงสร้างของไมโครคอนโทรลเลอร์
(ที่มา: นายวิโรจน์ กิตติวรปริดา เอกสารไมโครคอนโทรลเลอร์)

โครงสร้างโดยทั่วไปของไมโครคอนโทรลเลอร์ นั้น สามารถแบ่งออกมาได้เป็น 5 ส่วนใหญ่ ๆ ดังต่อไปนี้

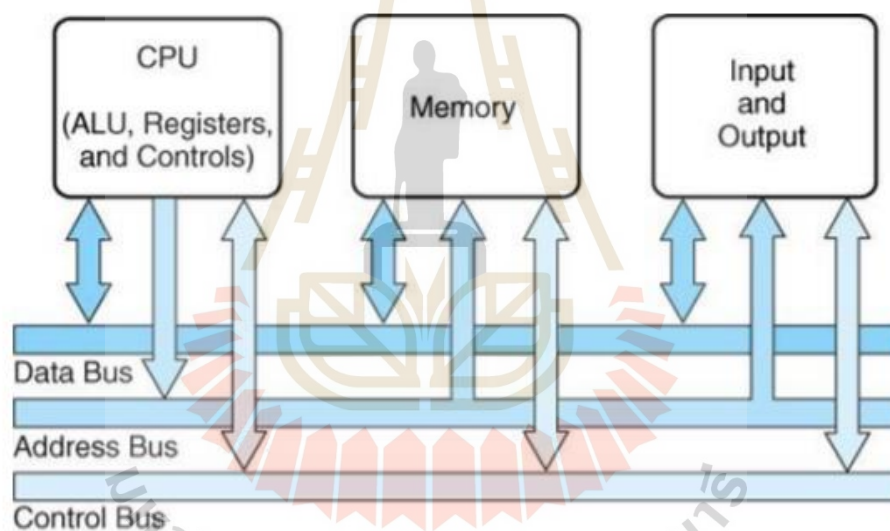
1) หน่วยประมวลผลกลางหรือซีพียู (CPU: Central Processing Unit) ทำหน้าที่เป็น ศูนย์กลางควบคุมการทำงานของระบบคอมพิวเตอร์ทั้งหมด โดยนำข้อมูลจากอุปกรณ์รับข้อมูลมา ทำงานประมวลผลข้อมูลตามคำสั่งของโปรแกรม และส่งผลลัพธ์ออกไปหน่วยแสดงผลโดยการทำงานของ ซีพียูมี 2 จังหวะ (1) เฟตซ์ (Fetch) คือ การอ่านคำสั่งโปรแกรมแล้วทำการถอดรหัสคำสั่งเป็น ภาษาเครื่อง (2) เอ็กซิคิวต์ (Executed) คือ การทำตามคำสั่งโปรแกรมที่ถอดรหัสแล้ว

2) หน่วยความจำ (Memory) สามารถแบ่งออกเป็น 2 ส่วน คือ (1) หน่วยความจำที่มีไว้สำหรับเก็บโปรแกรมหลัก (Program Memory) เปรียบเสมือน ฮาร์ดดิสก์ของเครื่องคอมพิวเตอร์ตั้งโต๊ะ คือข้อมูลใด ๆ ที่ถูกเก็บไว้ในนี้จะไม่สูญหายไปแม้ไม่มีไฟเลี้ยง อีกส่วนหนึ่งคือ (2) หน่วยความจำข้อมูล (Data Memory) ใช้เป็นเหมือนกระดานชนวนในการคำนวณของ ซีพียู และเป็นที่พักข้อมูลชั่วคราวขณะทำงาน แต่หากไม่มีไฟเลี้ยง ข้อมูลก็จะหายไปคล้ายกับหน่วย ความแรม (RAM) ในเครื่องคอมพิวเตอร์ทั่ว ๆ ไป แต่สำหรับไมโครคอนโทรลเลอร์สมัยใหม่ หน่วยความจำข้อมูลจะมีทั้งที่เป็น

หน่วยความจำแรม ซึ่งข้อมูลจะหายไปเมื่อไม่มีไฟเลี้ยง และเป็นอีอีพรอม (EEPROM: Erasable Electrically Read-Only Memory) ซึ่งสามารถเก็บข้อมูล ได้แม้ไม่มีไฟเลี้ยง

3) ส่วนติดต่อกับอุปกรณ์ภายนอก หรือพอร์ต (Port) มี 2 ลักษณะคือ พอร์ตอินพุต (Input Port) และพอร์ตส่งสัญญาณหรือพอร์ตเอาต์พุต (Output Port) ส่วนนี้จะใช้ในการเชื่อมต่อ กับ อุปกรณ์ภายนอก ถือว่าเป็นส่วนที่สำคัญมาก ใช้ร่วมกันระหว่างพอร์ตอินพุตเพื่อรับสัญญาณ อาจจะใช้การกดสวิตช์เพื่อนำไปประมวลผลและส่งไปพอร์ตเอาต์พุต เพื่อแสดงผลเช่น การติดสว่าง ของ หลอดไฟ เป็นต้น

4) ช่องทางเดินของสัญญาณ หรือบัส (BUS) คือเส้นทางการแลกเปลี่ยนสัญญาณข้อมูล ระหว่าง ซีพียู หน่วยความจำและพอร์ต เป็นลักษณะของสายสัญญาณจำนวนมากอยู่ในตัว ไมโครคอนโทรลเลอร์ โดยแบ่งเป็นบัสข้อมูล (Data Bus), บัสแอดเดรส (Address Bus) และบัสควบคุม (Control Bus)



รูปที่ 2.26 การสื่อสารของไมโครคอนโทรลเลอร์

(ที่มา: นายวิโรจน์ กิตติวรปรีดา เอกสารไมโครคอนโทรลเลอร์)

- บัสตำแหน่ง (Address Bus) จะเป็นกลุ่มสายสัญญาณที่ใช้สำหรับกำหนดตำแหน่งของ ข้อมูลในหน่วยความจำหรือระบุตำแหน่งที่อยู่ของอุปกรณ์ภายนอกต่าง ๆ โดยบัสตำแหน่งจะเป็น เส้นทางที่ใช้ส่งข้อมูลออกจากไมโครโพรเซสเซอร์เพียงทิศทางเดียว โดยใช้เพื่อระบุตำแหน่งของ อุปกรณ์ที่ต้องการจะติดต่อกับ เท่านั้น

- บัสข้อมูล (Data Bus) จะเป็นกลุ่มสายสัญญาณที่ใช้ในการรับส่งข้อมูลหรือรหัสคำสั่งต่าง ๆ ระหว่างไมโครโพรเซสเซอร์กับหน่วยความจำโปรแกรมหน่วยความจำข้อมูล อุปกรณ์อินพุตและ

อุปกรณ์เอาต์พุตทุกตัวที่ต่อพ่วงอยู่กับระบบ โดยบัสข้อมูลนี้จะเป็นแบบสองทิศทางสามารถรับและส่งข้อมูลทั้งไปและกลับได้ในสายสัญญาณเดียวกัน โดยบัสข้อมูลเปรียบเสมือนช่องทางเดิน รถถ้ามีมาก จะทำให้การติดต่อทำได้รวดเร็วขึ้น เช่น บัสข้อมูลแบบ 8 บิตจะทำงานได้เร็วกว่า บัสข้อมูลแบบ 4 บิต 3

- บัสควบคุม (Control Bus) จะเป็นกลุ่มสายสัญญาณที่ส่งสัญญาณควบคุมการทำงานต่าง ๆ ของระบบ โดยไมโครโพรเซสเซอร์สามารถส่งสัญญาณไปควบคุมหน่วยความจำและอุปกรณ์อินพุต/เอาต์พุตภายนอกเพื่ออ่านหรือเขียนข้อมูลได้ และอุปกรณ์ภายนอกสามารถส่งสัญญาณมาควบคุมการทำงานของไมโครโพรเซสเซอร์ได้เช่น การรีเซ็ต

5) วงจรกำเนิดสัญญาณนาฬิกา นับเป็นส่วนประกอบที่สำคัญมากอีกส่วนหนึ่ง เนื่องจากการทำงานที่เกิดขึ้นในตัวไมโครคอนโทรลเลอร์ จะขึ้นอยู่กับกำหนัดจังหวะ หากสัญญาณนาฬิกาที่มีความถี่สูง จังหวะการทำงานก็จะสามารถทำได้ถี่ขึ้นส่งผลให้ไมโครคอนโทรลเลอร์ตัวนั้น มีความเร็วในการประมวลผลสูงตามไปด้วย

2.7.2 ไมโครคอนโทรลเลอร์และไมโครโพรเซสเซอร์

จากการพัฒนาเทคโนโลยีในการผลิตไอซีรวมได้มีความก้าวหน้าขึ้นอย่างมากมีการนำทรานซิสเตอร์หลาย ๆ ตัวมาสร้างรวมไว้ด้วยกันและพัฒนาจนเป็นทรานซิสเตอร์ล้าน ๆ ตัวมาไว้ในวงจรรวมแบบ LSI (Large Scale integrated Circuit) ได้ และมีการพัฒนาต่อไปโดยสามารถรวมเอาวงจรรวมที่ใช้เป็นหน่วยประมวลผลกลางของระบบคอมพิวเตอร์ มาบรรจุอยู่ในไอซีโดยเรียกไอซีนี้ว่า ไมโครโพรเซสเซอร์(Microprocessor) มาต่อหน้าไมโครโพรเซสเซอร์ร่วมกับหน่วยความจำ และหน่วย อินพุต/เอาต์พุต จะทำให้ได้เป็นระบบคอมพิวเตอร์ขนาดเล็กขึ้นมาเรียกว่า ไมโครคอมพิวเตอร์ (Microcomputer) คือคอมพิวเตอร์ที่ใช้ไมโครโพรเซสเซอร์เป็นหน่วยประมวลผลกลาง ไมโครคอนโทรลเลอร์ (Microcontroller) คือ อุปกรณ์ควบคุมขนาดเล็กที่มีความสามารถ ในการประมวลผลและตัดสินใจต่าง ๆ สามารถเขียนโปรแกรมควบคุมได้อย่างอิสระ ทำให้สามารถ นำไปใช้งานแทนวงจรรีเลย์ทรานซิสเตอร์ที่ซับซ้อนได้เป็นอย่างดี เหมาะกับการนำไปใช้ในงานควบคุม ต่าง ๆ ซึ่งภายในตัวไมโครคอนโทรลเลอร์จะประกอบด้วย หน่วยประมวลผลกลาง หน่วยความจำ อุปกรณ์อินพุต/เอาต์พุต และวงจรสัญญาณนาฬิกา เหมือนกับนำระบบคอมพิวเตอร์ขนาดเล็กบรรจุไว้ในไอซีตัวเดียว ไมโครโพรเซสเซอร์ไม่สามารถต่อใช้งานได้เพียงตัวเดียว จะต้องต่อกับหน่วยความจำและอุปกรณ์อินพุต/เอาต์พุตจึงจะสามารถนำไปใช้งานได้ แต่ตัวไมโครคอนโทรลเลอร์สามารถนำไปใช้ใน งานควบคุมได้ทันทีเพราะภายในได้รวมเอาหน่วยความจำอุปกรณ์อินพุต/เอาต์พุต และวงจรรวมที่จำเป็นบางส่วนเอาไว้ภายในตัวไอซีเรียบร้อยแล้ว

บทที่ 3

วิธีการดำเนินงานวิจัย

3.1 กล่าวนำ

การดำเนินงานวิจัยทำการศึกษวิเคราะห์สาเหตุการเกิดรอยตำหนิวัตถุ เก็บรวบรวมข้อมูลภาพตัวอย่างทั้ง 4 ประเภท รวมทั้งนำหลักการทฤษฎีที่เกี่ยวข้องมาประยุกต์ใช้กับอัลกอริทึม YOLOv5 (You Look Only Once version 5) การเรียนรู้เชิงลึกโครงข่ายประสาทเทียมแบบสังวัตนาการที่มีการเพิ่มอัลกอริทึม Bounding Box เพื่อให้โมเดลเหมาะสมสำหรับการตรวจจ็บรอยตำหนิชิ้นงานแบบเรียลไทม์

3.2 ซอฟต์แวร์และอุปกรณ์ที่ใช้

3.2.1 ระบบปฏิบัติการและภาษาคอมพิวเตอร์

3.2.1.1 ระบบปฏิบัติการ Ubuntu เวอร์ชัน 18.04

3.2.1.2 โปรแกรมArduino Integrated Development Environment

3.2.1.3 โปรแกรม Logitech Webcam Software

3.2.1.4 โปรแกรม Label IMG

3.2.1.5 ภาษาโปรแกรมคอมพิวเตอร์ Python เวอร์ชัน 3.6.7

3.2.1.6 ไลบรารีสำหรับอัลกอริทึมYOLO (You Look Only Once)

3.2.2 อุปกรณ์

3.2.2.1 บอร์ด NVIDIA Jetson Nano การประมวลผลGPU 128 Core NVIDIA Maxwell และ CPU 64-bit Quad-core ARM A57 หน่วยความจำ RAM 4 กิกะไบต์

3.2.2.2 เครื่องคอมพิวเตอร์ AMD Ryzen 5 4600H, หน่วยความจำ 16 กิกะไบต์ ระบบปฏิบัติการ 64 บิต Windows 10

3.2.2.3 บอร์ดNode MCU ชิฟ ESP-12F/N

3.2.2.4 สายพานลำเลียงวัตถุ ขนาด 20 เซ็นติเมตร x 120 เซ็นติเมตร x สูง 75 เซ็นติเมตร ขนาดกำลังมอเตอร์ 60 วัตต์ โครงสร้างสแตนเลส

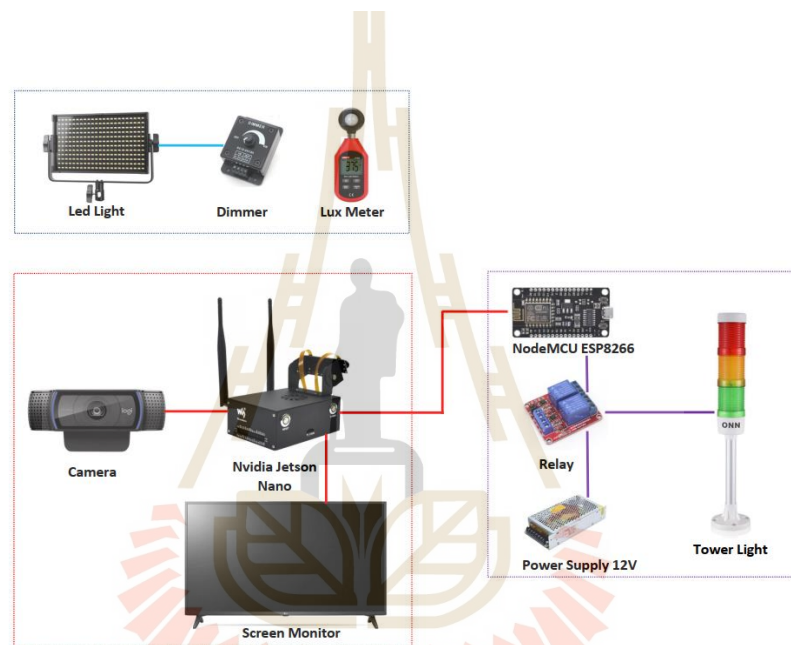
3.2.2.5 อุปกรณ์ปรับควบคุมความเร็วมอเตอร์ แรงดันไฟฟ้า 220 โวลต์ +/- 10% ความถี่ 60 เฮิรตซ์

3.2.2.6 อุปกรณ์Dimmerปรับกระแสของแสงสว่างไฟฟ้าLED แรงดันกระแสไฟฟ้าตรง 12 โวลต์ ถึง 24 โวลต์ กระแส 30 แอมป์

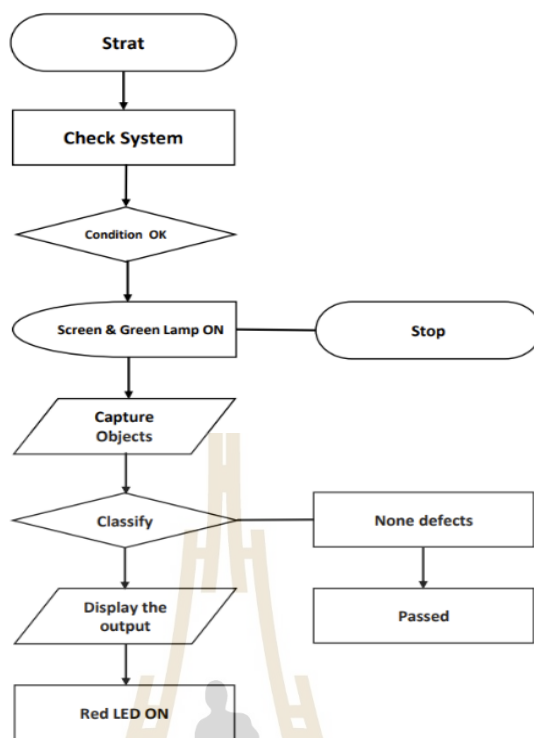
3.2.2.7 สวิตชิงเพาเวอร์ซัพพลาย แรงดันไฟฟ้ากระแสสลับ 12 โวลต์ แปลงเป็นกระแสไฟฟ้าตรง 12 โวลต์ 5 แอมป์ 60 วัตต์

3.2.2.8 เครื่องวัดความสว่าง UT383

3.2.2.9 หน้าจอทีวี ขนาด 24 นิ้ว ความละเอียด HD 1366 x 768 พิกเซล โดยสรุปเป็นแผนภาพของอุปกรณ์ตามรูปที่ 3.1 และอธิบายระบบการทำงานของอุปกรณ์ตามรูปที่ 3.2



รูปที่ 3.1 อุปกรณ์ต้นแบบ



รูปที่ 3.2 ระบบการทำงานของอุปกรณ์

3.2.3 การติดตั้งอุปกรณ์บนระบบสายพานลำเลียงวัตถุ

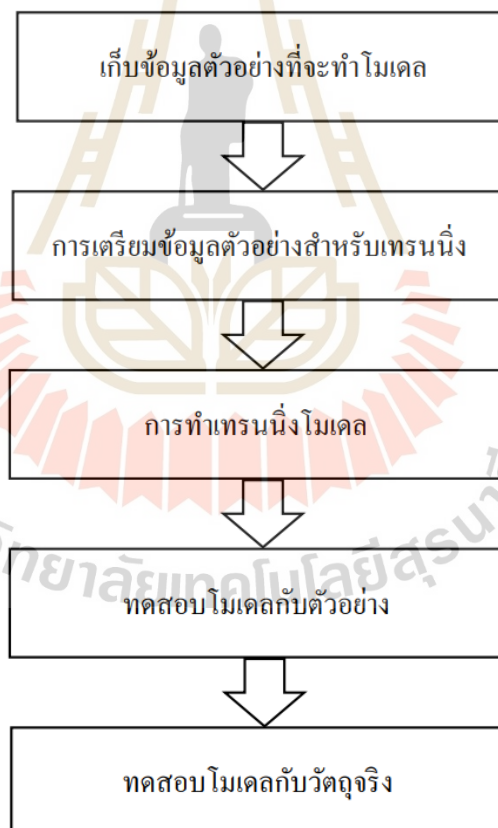
กล้อง Logitech ถูกติดตั้งระยะห่างจากชิ้นงานที่ตรวจสอบ 25 เซ็นติเมตร ด้านในติดตั้งหลอดไฟ LED ทั้ง 2 ตำแหน่งซ้ายขวาด้านบน ใช้แหล่งจ่ายไฟแปลงไฟจากกระแสไฟฟ้าสลับ 220 โวลต์เป็นกระแสไฟฟ้ากระแสตรง 12 – 24 โวลต์ ให้กับ Dimmer เพื่อปรับให้ค่าความสว่างคงที่ 1,800 ลักซ์ สร้างโครงสร้างแผ่นพลาสติกเพื่อป้องกันแสงสะท้อน มีขนาด กว้าง 45 เซ็นติเมตร x ยาว 60 เซ็นติเมตร x สูง 35 ระบบสายพานติดตั้งอุปกรณ์ควบคุมความเร็วมอเตอร์ตรงตำแหน่งด้านล่างขวาของสายพาน ตามรูปที่ 3.3



รูปที่ 3.3 การติดตั้งอุปกรณ์

3.3 ขั้นตอนการดำเนินการวิจัย

ขั้นตอนการดำเนินการวิจัยมีขั้นตอนดังต่อไปนี้



3.3.1 ขั้นตอนการเก็บข้อมูลตัวอย่างที่ทำโมเดล

การเก็บข้อมูลทำการวิเคราะห์ปัญหาปัจจัยที่ทำให้รอยบนผิววัตถุแม่แรงรถยนต์ โดยศึกษาจากกระบวนการพ่นสีและโครงสร้างของเครื่องจักรซึ่งเป็นระบบการอบวัตถุด้วยความร้อน โดยมีรายละเอียดตามด้านล่างต่อไปนี้

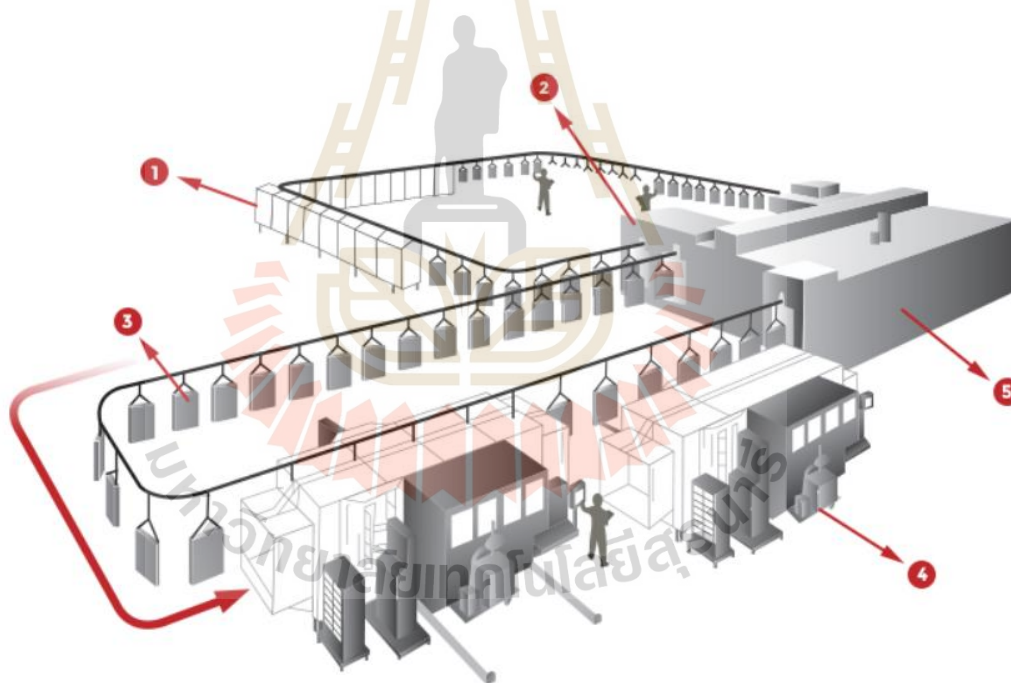
ขั้นตอนที่ 1 คือ ขั้นตอนการทำความสะอาดพื้นผิววัตถุ

ขั้นตอนที่ 2 คือ ขั้นตอนการทำความสะอาดพื้นผิววัตถุด้วยแรงดันลม

ขั้นตอนที่ 3 คือ ขั้นตอนการลำเลียงวัตถุ

ขั้นตอนที่ 4 คือ ขั้นตอนการพ่นสีวัตถุ

ขั้นตอนที่ 5 คือ ขั้นตอนการอบวัตถุด้วยความร้อน

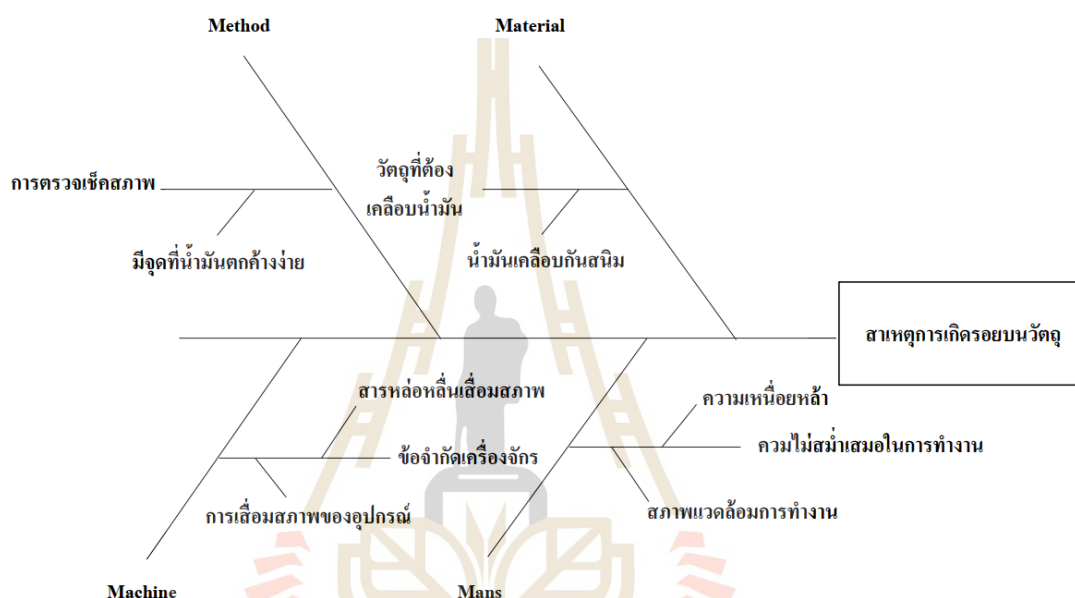


รูปที่ 3.4 กระบวนการพ่นสีชิ้นงาน

(ที่มา: <http://www.powton.co.kr/en/application/layout.asp>)

ในขั้นตอนที่ 1 และ 2 เป็นกระบวนการทำความสะอาดพื้นผิววัตถุแม่แรงรถยนต์ เพื่อชำระล้างคราบน้ำมัน และในขั้นตอนที่ 3 ลำเลียงวัตถุไปพ่นสีฝุ่น เมื่อพ่นสีที่วัตถุเสร็จแล้ว ในขั้นตอนที่ 4 ระบบสายพานจะ

นำชิ้นงานเข้าเตาอบเพื่อให้สีฝุ่นเกาะติดผิววัตถุด้วยหลักการประจุโปรตอนอิเล็กตรอน และในขั้นตอนที่ 5 อบวัตถุด้วยความร้อนประมาณ 180 – 200 องศา เพื่อให้สีฝุ่นละลายติดผิว จากปัญหาที่พบในกระบวนการตรวจสอบคุณภาพชิ้นงาน มี 4 ปัญหาหลัก ๆ ได้แก่ สีลอกเนื่องจากคราบน้ำมันตกค้างบนพื้นผิววัตถุ, การเกิดรอยขีดข่วนและการกระแทกที่ผิววัตถุจากการกระทบกับส่วนต่าง ๆ ของเครื่องจักรและรอยร้าวบนผิววัตถุจากกระบวนการผลิตวัตถุดิบ สามารถอธิบายสภาพปัญหาโดยใช้เครื่องมือ 7 Tools แผนภาพก้างปลาตามรูปที่ 3.5



รูปที่ 3.5 แผนภูมิก้างปลา 7 Tools

การเก็บข้อมูลตัวอย่างเพื่อทำโมเดล มีข้อมูลทั้งหมด 1,060 ภาพ ดังต่อไปนี้

- 1) รอยขีดข่วน 265 ภาพ ขนาดเฉลี่ย กว้าง 3 มิลลิเมตร x ยาว 15 มิลลิเมตร
- 2) รอยสีลอก 265 ภาพ ขนาดเฉลี่ย กว้าง 10 มิลลิเมตร x ยาว 20 มิลลิเมตร
- 3) รอยกระแทก 265 ภาพ ขนาดเฉลี่ย กว้าง 1 มิลลิเมตร x ยาว 1.5 มิลลิเมตร
- 4) รอยวัตถุดิบร้าว 265 ภาพ ขนาดเฉลี่ย กว้าง 0.5 มิลลิเมตร x ยาว 15 มิลลิเมตร

3.3.2 ขั้นตอนการเตรียมข้อมูลตัวอย่างสำหรับเทรนนิ่ง

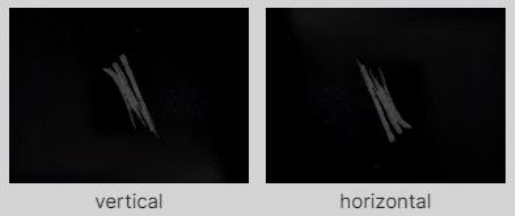

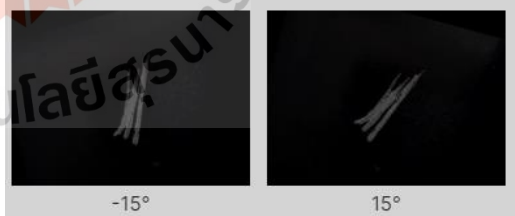
3.3.2.1 การเตรียมข้อมูลภาพสำหรับเทรนนิ่งโมเดล เป็นขั้นตอนที่สำคัญเพื่อให้ได้โมเดลที่มีความแม่นยำสูงและป้องกันการเกิดปัญหา Error ขณะทำการเทรนนิ่งข้อมูล ข้อมูลภาพทั้งหมด เป็นไฟล์นามสกุล JPG. มีขนาดความกว้าง 640 พิกเซลและขนาดความสูง 480 พิกเซล ทุกรูปจะมีขนาดเท่ากันทั้งหมด เพื่อให้ได้ภาพที่มีลักษณะที่เหมาะสมสำหรับการประมวลผลของอัลกอริทึมจะใช้หลักการปรับค่าความสว่างของภาพและการทำ Contrast ซึ่งเป็นกระบวนการปรับภาพระหว่างส่วนที่สว่างและส่วนที่มืด โดยปรับค่า Contrast ให้สูงขึ้นเพื่อทำภาพตรงรอยดำหิ้งขึ้นงานดูคม ดูชัดและตัดกัน ตามรูปที่ 3.6



รูปที่ 3.6 ผลลัพธ์หลังจากปรับความสว่างและทำ Contrast

3.3.2.2 การทำ Augmentation process เป็นวิธีการดัดแปลงข้อมูลต้นฉบับให้เป็นข้อมูลใหม่ที่มีความใกล้เคียงกับข้อมูลเดิม เพื่อเพิ่มปริมาณของข้อมูลที่นำมาใช้ในการเทรนนิ่งอัลกอริทึมของโมเดล มีกระบวนการทั้งหมด 3 รูปแบบ ตามตารางที่ 3.1

ตารางที่ 3.1 การทำAugmentation Pattern

Augmentation Pattern	ผลลัพธ์
การพลิก (flip) <ul style="list-style-type: none"> - แนวตั้ง (Vertical) - แนวนอน (Horizontal) 	
การหมุน 90° <ul style="list-style-type: none"> - ตามเข็มนาฬิกา (Clockwise) - ทวนเข็มนาฬิกา (Counter-Clockwise) 	
การหมุน 15° <ul style="list-style-type: none"> - หมุน -15° - หมุน 15° 	

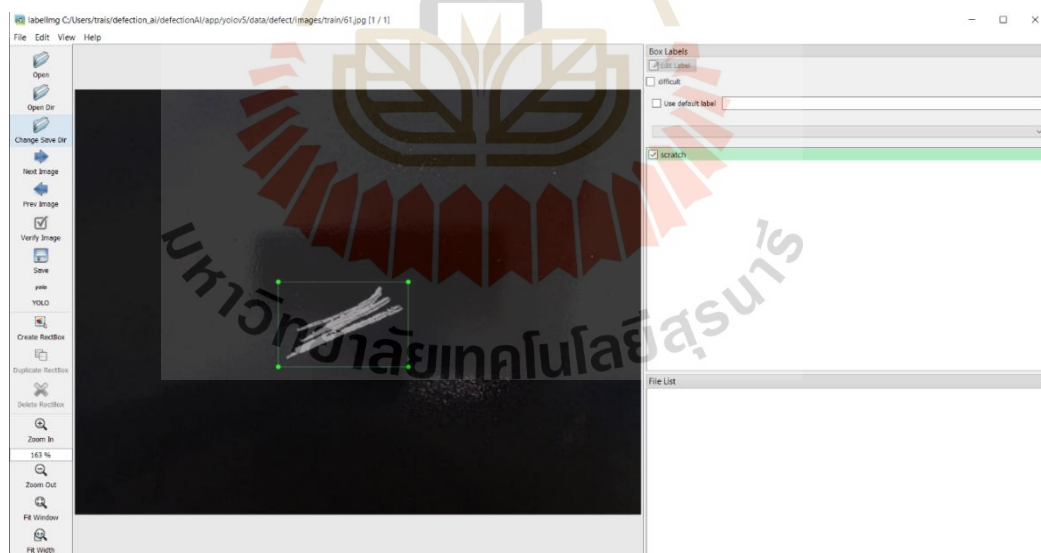
3.3.2.3 การแบ่งข้อมูลสำหรับทำ Train Set, Validation Set และ Test Set เพื่อให้โมเดลมีประสิทธิภาพสามารถตรวจสอบรอยตำหนิกับวัตถุจริงได้จะใช้หลักการงานวิจัยของ L.Shao,A.Mahajan,T.Schreck และ D.J.Lehmann หัวข้องานวิจัย “Interactive Regression Lens for Exploring Scatter Plots.” เพื่อป้องกันการเกิดปัญหา Overfitting และ Under fitting โดยใช้วิธี Cross Validation แบ่งข้อมูลออกเป็นส่วน ๆ ส่วนแรกแบ่งให้กับ Test Set 20 เปอร์เซ็นต์

เพื่อกำหนดความแม่นยำของโมเดล และแบ่งให้กับ Validation Set 20 เปอร์เซ็นต์ เพื่อทดสอบโมเดลไม่ให้เกิดปัญหา Overfitting ส่วนที่เหลือทั้งหมดแบ่งให้กับ Train Set ตามตารางที่ 3.2

ตารางที่ 3.2 การแบ่งข้อมูลสำหรับเทรนนิ่งโมเดล

No.	Classes	Train Set	Test Set	Validation Set
1	รอยขีดข่วน	165	55	45
2	รอยสีลอก	165	55	45
3	รอยกระแทก	165	55	45
4	รอยวัตถุติดบร้าว	165	55	45
	รวม	660	220	180

3.3.2.4 การทำLabel เพื่อกำหนดให้โมเดลเรียนรู้รอยตำหนิวัตถุที่ต้องการให้แยกแยะแต่ละประเภท วิธีการเป็นการสร้าง Bounding Box เป็นกล่องสี่เหลี่ยมครอบคลุมที่รอยตำหนิ ในงานวิจัยนี้จะใช้เครื่องมือที่มีชื่อว่า Label IMG ตามรูปที่ 3.7



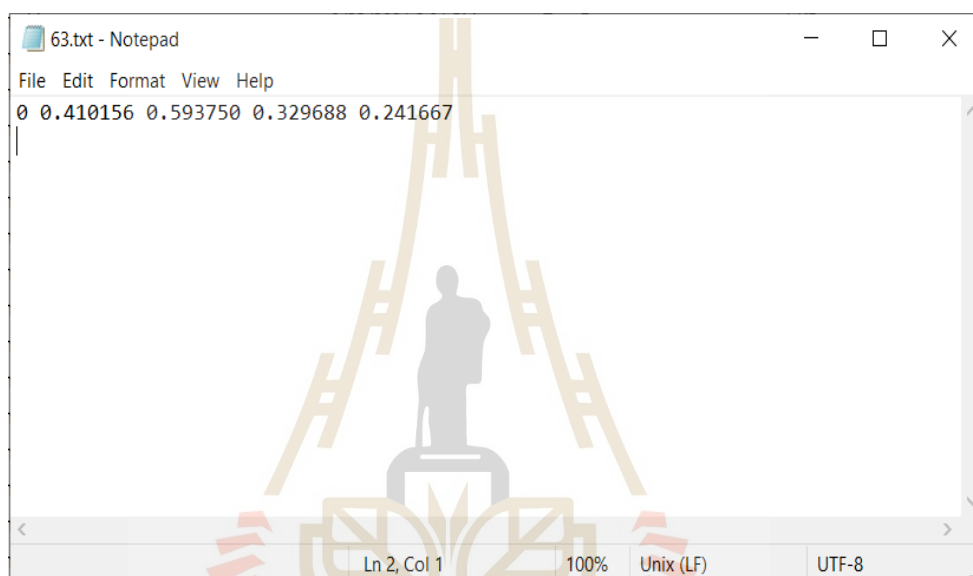
รูปที่ 3.7 การใช้โปรแกรม Label IMG สร้าง Bounding Box

ผลลัพธ์ที่ได้จากการทำ Bounding box จะได้ไฟล์ในรูปแบบของ text file ตามรูปที่ 3.8 มีข้อมูลตัวแปร 5 ตัวแปรดังต่อไปนี้

ค่าตัวแปรที่ 1 ประเภทของ Object Class รอยบนวัตถุ

ค่าตัวแปรที่ 2, 3 คือตำแหน่งตรงกลางของรูปอยู่ในแกน X,Y

ค่าตัวแปรที่ 4, 5 คือค่าความกว้างและความสูงของ Bounding box

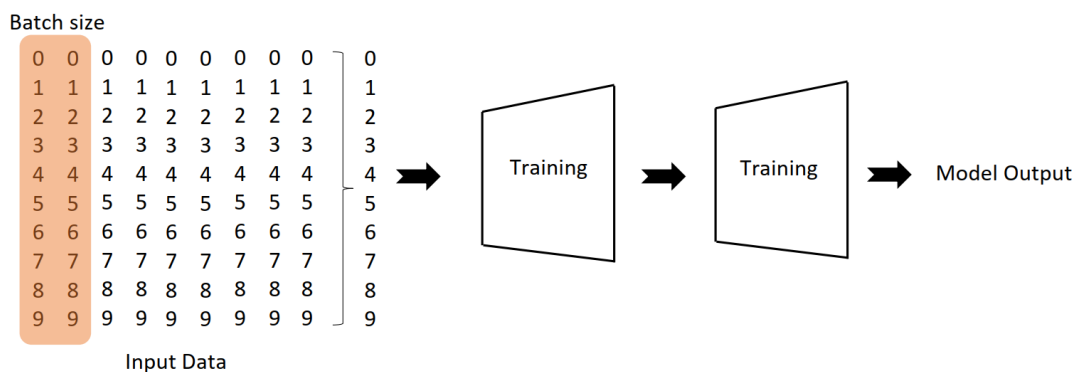


```

63.txt - Notepad
File Edit Format View Help
0 0.410156 0.593750 0.329688 0.241667
Ln 2, Col 1 100% Unix (LF) UTF-8
  
```

รูปที่ 3.8 ค่าตัวแปรหลังจากทำ Bounding box

3.3.2.5 การป้อนข้อมูลเพื่อทำโมเดล จากข้อมูลที่ได้เตรียมพร้อมเรียบร้อยแล้วจะมีขนาดข้อมูลปริมาณเยอะ จึงไม่สามารถป้อนเข้าไปในโมเดลเพื่อทำการประมวลผลตามอัลกอริทึมได้ภายในครั้งเดียว ดังนั้น จะมีกระบวนการที่เรียกว่า Batch Size การแบ่งข้อมูลออกเป็นทีละส่วน อธิบายเป็นหลักการง่าย ๆ ตามรูปที่ 3.9



รูปที่ 3.9 การ Batch Size ข้อมูลเพื่อทำการเทรนนิ่งโมเดล

สำหรับอัลกอริทึมโมเดลของ Yolo งานวิจัยฉบับนี้เวอร์ชัน 5 โดยทำการกำหนดค่าพารามิเตอร์ให้เหมาะสมเพื่อป้องกันการเกิดปัญหาในระหว่างทำการเทรนนิ่ง กำหนดค่าพารามิเตอร์ขนาดของภาพที่ป้อนเข้าโมเดลเป็น 640 พิกเซล ตามขนาดของภาพที่ได้เก็บข้อมูลก่อนหน้านี้ พารามิเตอร์ Batch Size กำหนดเป็น 20 เพื่อป้องกันการใช้ทรัพยากรของหน่วยประมวลผลคอมพิวเตอร์มากเกินไป ข้อมูลจะถูกป้อนเข้าไปในโมเดลวนจนครบ 53 รอบ (53 รอบ = 1 epochs) กำหนดพารามิเตอร์ epochs เท่ากับ 100

ตารางที่ 3.3 การตั้งค่าพารามิเตอร์ของโมเดล

การตั้งค่า	รายละเอียด
Model	Yolov5s
Dataset	1,060 ภาพ
Image size	640
Batch size	20
Epochs	100 * 3

การเทรนนิ่งข้อมูลของอัลกอริทึม จะอ้างอิงค่า weight ของโมเดล YOLOv5s เป็นตัวเริ่มต้น ภายในไฟล์มีตัวแปรพารามิเตอร์ต่าง ๆ เพื่อปรับค่า weight ไฟล์มีข้อมูล 14,449 กิโลไบต์ นามสกุลไฟล์เป็นรูปแบบ .PT (Panther) มีจำนวน 8 คอลัมน์ 924,698 แถว เมื่อทำการเทรนนิ่งโมเดลจนครบตามค่าพารามิเตอร์ จะได้ค่า weight ของโมเดลต้นแบบ มีข้อมูล 14,056 กิโลไบต์ 8 คอลัมน์ 899,567 แถว ตามรูปที่ 3.10

924689	0800	0000	0000	00d1	9e67	5502	0000	0002	899558	0000	d19e	6755	0200	0000	0200	0000	0f00
924690	0000	000f	0000	0000	0000	0000	0000	0000	899559	0000	0000	0000	0000	0000	0000	d044	db00
924691	0090	57e1	0061	7263	6869	7665	2f76	6572	899560	6172	6368	6976	652f	7665	7273	696f	6e50
924692	7369	6f6e	504b	0606	2c00	0000	0000	0000	899561	4b06	062c	0000	0000	0000	001e	032d	0000
924693	1e03	2d00	0000	0000	0000	0000	7101	0000	899562	0000	0000	0000	0072	0100	0000	0000	0072
924694	0000	0000	7101	0000	0000	0000	2269	0000	899563	0100	0000	0000	002d	5900	0000	0000	0052
924695	0000	0000	1258	e100	0000	0000	504b	0607	899564	45db	0000	0000	0050	4b06	0700	0000	007f
924696	0000	0000	34c1	e100	0000	0000	0100	0000	899565	9edb	0000	0000	0001	0000	0050	4b05	0600
924697	504b	0506	0000	0000	7101	7101	2269	0000	899566	0000	0072	0172	012d	5900	0052	45db	0000
924698	1258	e100	0000						899567	00							

Weight ของ YOLOv5s

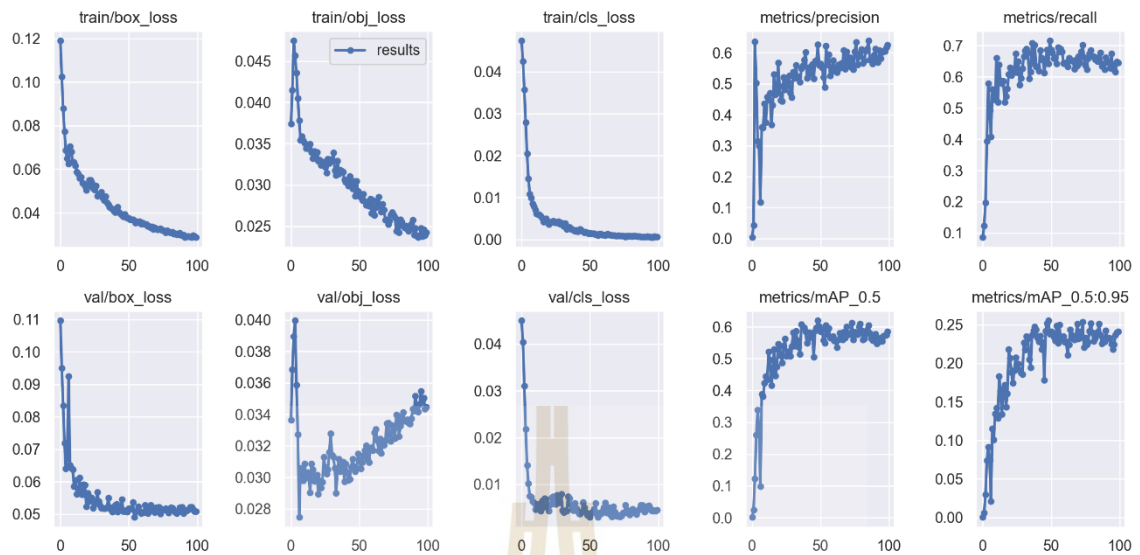
Weight ของ โมเดลต้นแบบ

รูปที่ 3.10 ตัวแปร weight ของโมเดล

เพื่อให้ได้ค่าความแม่นยำของโมเดลเพิ่มขึ้นจะนำค่า weight ของการเทรนนิ่งแต่ละรอบที่มีโครงสร้างของพารามิเตอร์ปรับจูนจากครั้งก่อนมาใช้ กระบวนการนี้เรียกว่า transfer learning จะทำให้การเทรนนิ่งในครั้งถัด ๆ ไปเร็วขึ้น ผลลัพธ์ค่าความแม่นยำของโมเดลแสดงตามตารางที่ 3.4 และในรูปที่ 3.11 แสดงข้อมูลการเทรนนิ่งของแต่ละช่วงที่ทำครบแต่ละ 1 epoch

ตารางที่ 3.4 ค่าความแม่นยำจากการทำ transfer learning

Transfer learning	ผลลัพธ์ค่าความแม่นยำ % (Precision)
ครั้งที่ 1 อ้างอิง weight ของ YOLOv5s	93.8 %
ครั้งที่ 2 ใช้ weight ครั้งที่ 1	94.8 %
ครั้งที่ 3 ใช้ weight ครั้งที่ 2	95.6 %



รูปที่ 3.11 กราฟการเทรนนิ่งข้อมูล 100 epoches

3.4 สรุป

โมเดลต้นแบบสำหรับตรวจสอบและแยกแยะรอยตำหนิ ทำการวิเคราะห์เก็บข้อมูลในกระบวนการฟอสซิลขึ้นงาน หลังจากนั้น เตรียมข้อมูลและกำหนดค่าพารามิเตอร์ให้เหมาะสมสำหรับการเทรนนิ่ง ใช้เวลาเทรนนิ่งโมเดลทั้งหมด 96 ชั่วโมง 12 นาที และใช้วิธีการ transfer learning นำค่า weight จากการเทรนนิ่งก่อนหน้ามาเทรนนิ่งโมเดลซ้ำเพื่อเพิ่มค่าความแม่นยำตามเกณฑ์ที่กำหนด ส่วนการวิเคราะห์ผลประเมินประสิทธิภาพของอัลกอริทึมถูกอธิบายไว้ในบทต่อไป

บทที่ 4

ผลการศึกษาและวิเคราะห์ผล

4.1 กล่าวนำ

ในการวัดประสิทธิภาพของโมเดลใช้วิธีการคำนวณแบบ Confusion Matrix ซึ่งเป็นเครื่องมือสำคัญในการประเมินผลลัพธ์การทำนายของโมเดลที่สร้างขึ้น หลักการคือวัดผลลัพธ์ที่โมเดลทำนายกับผลลัพธ์จริง และทำการทดสอบความถูกต้องการแยกแยะประเภทรอยตำหนิกับตัวอย่าง

4.2 การทดสอบประสิทธิภาพโมเดลแบบ Confusion Matrix

การวัดประสิทธิภาพโมเดลแบบ Confusion Matrix คือการวัดความสามารถของโมเดลในการแยกแยะประเภท โดยแบ่งเป็น 4 ค่า ตามรูปที่ 4 ดังต่อไปนี้

True Positive (TP) คือ สิ่งที่โมเดลทำนายว่า จริง และมีค่าเป็น จริง ค่าทำนายถูกต้องเชิงบวก

True Negative (TN) คือ สิ่งที่โมเดลทำนายว่า ไม่จริง และมีค่า ไม่จริง ค่าทำนายถูกต้องเชิงลบ

False Positive (FP) คือ สิ่งที่โมเดลทำนายว่า จริง แต่มีค่า ไม่จริง ค่าทำนายผิดพลาดเชิงบวก

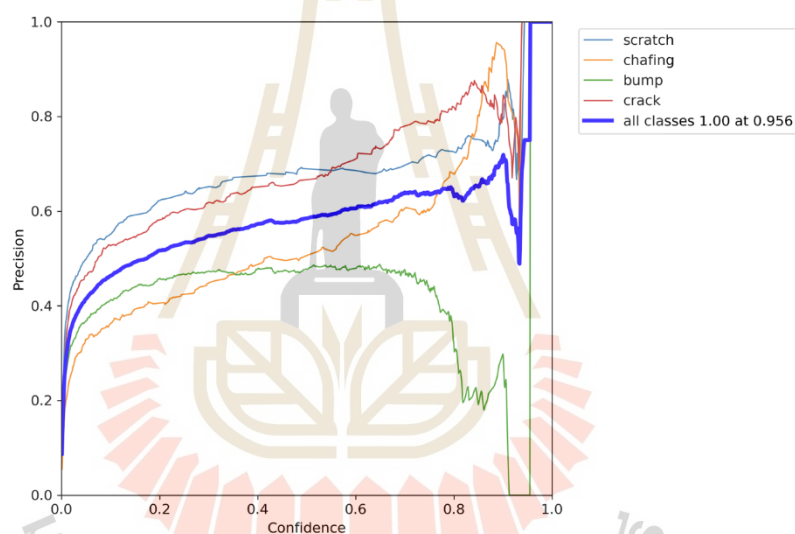
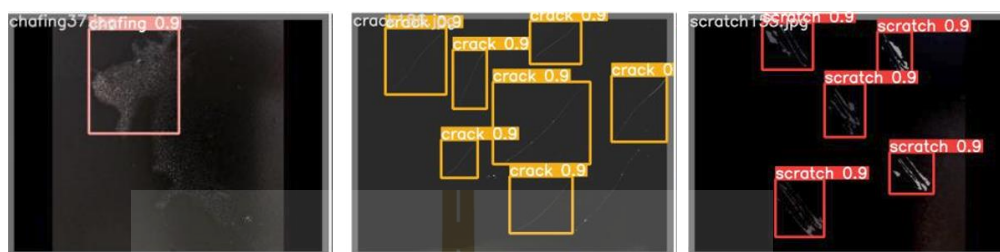
False Negative (FN) คือ สิ่งที่โมเดลทำนายว่า ไม่จริง แต่มีค่า จริง ค่าทำนายผิดพลาดเชิงลบ

ใช้ตัวแปร 4 ค่า TP, TN, FP และ FN เพื่อทำการวัดประสิทธิภาพความแม่นยำ (Precision) โดยพิจารณาแยกที่ละคลาสของโมเดล

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

$$Precision = \frac{TP}{TP+FP} \quad (4.1)$$

จากผลลัพธ์ของการสร้างโมเดลแบบจำลองของทั้ง 4 คลาสได้แก่ รอยขีดข่วน (scratch) รอยสีลอก (chafing) รอยกระแทก (bump) และรอยผิวดูแตก (crack) โดยมีค่าความแม่นยำ 95.6 เปอร์เซ็นต์ต่อค่าความเชื่อมั่น (confidence) ตามรูปที่ 4.1



รูปที่ 4.1 กราฟค่าความแม่นยำของผลลัพธ์การสร้างโมเดล

จากกราฟมีเส้นสีเขียวคือคลาสของรอยกระแทก (bump) ที่มีช่วงความแม่นยำต่ำกว่าคลาสอื่น เนื่องจากรอยกระแทกมีลักษณะคล้าย ๆ กันมีขนาดเล็กและความหลากหลายของคุณลักษณะค่อนข้างน้อย การแก้ไขจะดำเนินการปรับฝึกซ้ำโดยใช้ค่าคุณลักษณะจากครั้งก่อนมาเทรนนิ่งและเพิ่มรอบการทำซ้ำมากกว่าเดิม

4.3 การทดสอบประสิทธิภาพโมเดลการแยกประเภทรอยต่าง ๆ

ทดสอบการทำงานของโมเดลกับตัวอย่างประเภทรอยต่าง ๆ ที่ได้เตรียมไว้ก่อนหน้านี้ โดยมีทั้งหมด 4 ประเภท ได้แก่ รอยขีดข่วน, รอยกระแทก, รอยสีลอกและรอยร้าว รวมทั้งหมด 400

ตัวอย่าง จากตารางที่ 4.1 แสดงให้เห็นว่าโมเดลจำแนกประเภทรอยต่าง ๆ โดยมีค่าความถูกต้อง 96%

ตารางที่ 4.1 การทดสอบประสิทธิภาพความถูกต้องการแยกประเภทรอยต่างๆบนวัตถุ

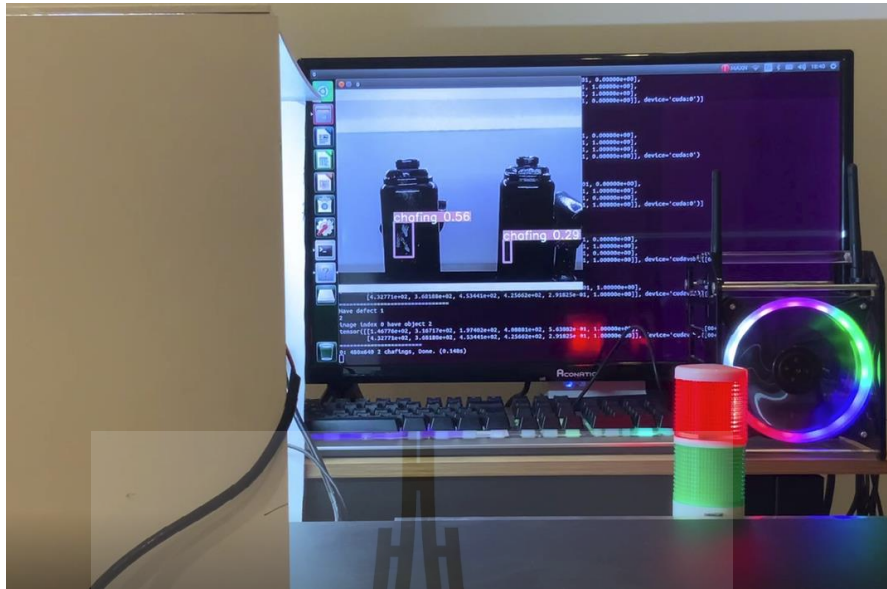
No.	ประเภทรอย (classes)	จำนวนตัวอย่างภาพ	ผลลัพธ์	ความถูกต้อง (%)
1	รอยขีดข่วน	100	98	98%
2	รอยกระแทก	100	94	94%
3	รอยสีลอก	100	97	97%
4	รอยร้าว	100	95	95%
	ค่าเฉลี่ยทั้ง 4 คลาส			96%

4.4 การทดสอบประสิทธิภาพโมเดลกับวัตถุจริงแบบเรียลไทม์

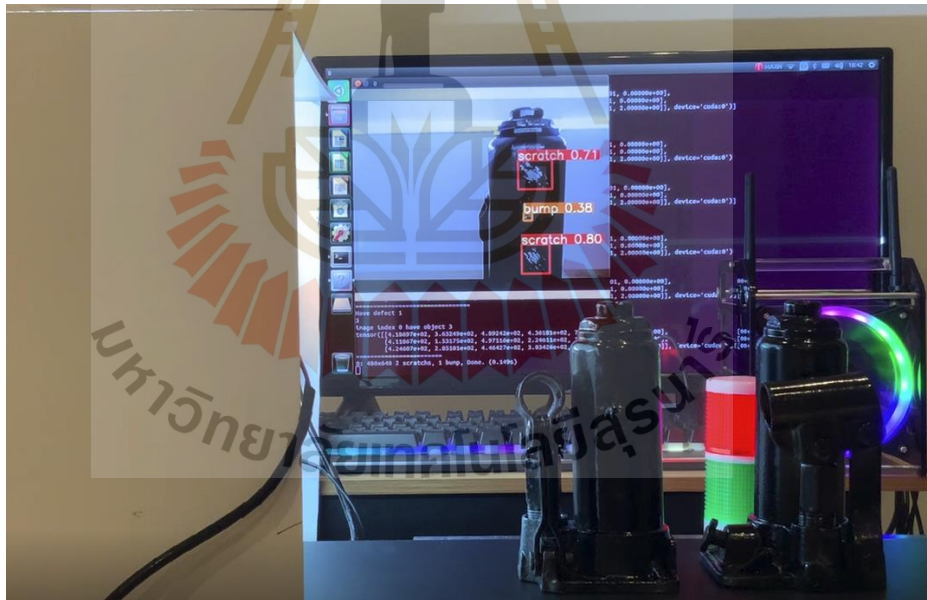
การทดสอบประสิทธิภาพโมเดลกับวัตถุจริง เพื่อทดสอบการตรวจสอบและแยกประเภทรอยตำหนิวัตถุโดยไม่เกิดปัญหา Underfitting และ Overfitting วิธีการทดสอบดำเนินการโดยนำชิ้นงานแม่แรงรถยนต์วางตรงจุดที่กำหนด โดยจำลองสภาพแวดล้อมเสมือนไลน์กระบวนการตรวจสอบคุณภาพในกระบวนการผลิตจริง ซึ่งจะมีการผลิตชิ้นงานที่มีขนาดแตกต่างกัน 3 รุ่น โดยแต่ละชิ้นงานจะมีรอยตำหนิที่ไม่เหมือนกัน ได้แก่

- 1) แม่แรงรถยนต์รับน้ำหนัก 2 ตัน ขนาด $9.5 \times 7.5 \times 16$ เซ็นติเมตร รอยสีลอก
- 2) แม่แรงรถยนต์รับน้ำหนัก 3 ตัน ขนาด $10.5 \times 9 \times 20$ เซ็นติเมตร รอยกระแทกและรอยขีดข่วน
- 3) แม่แรงรถยนต์รับน้ำหนัก 5 ตัน ขนาด $11.5 \times 9.5 \times 20$ เซ็นติเมตร รอยสีลอกและรอยร้าว

ผลการทดสอบโมเดลต้นแบบ สามารถตรวจสอบและแยกแยะประเภทได้อย่างถูกต้อง และสามารถระบุรอยตำหนิได้มากกว่า 1 ตำแหน่ง ตามรูปที่ 4.2, 4.3 และ 4.4 ตามลำดับ



รูปที่ 4.2 การตรวจสอบรอยสีลอกกับแม่แรงรถยนต์ขนาด 2 ตัน



รูปที่ 4.3 การตรวจสอบรอยขีดข่วนและรอยกระแทกกับแม่แรงรถยนต์ขนาด 3 ตัน

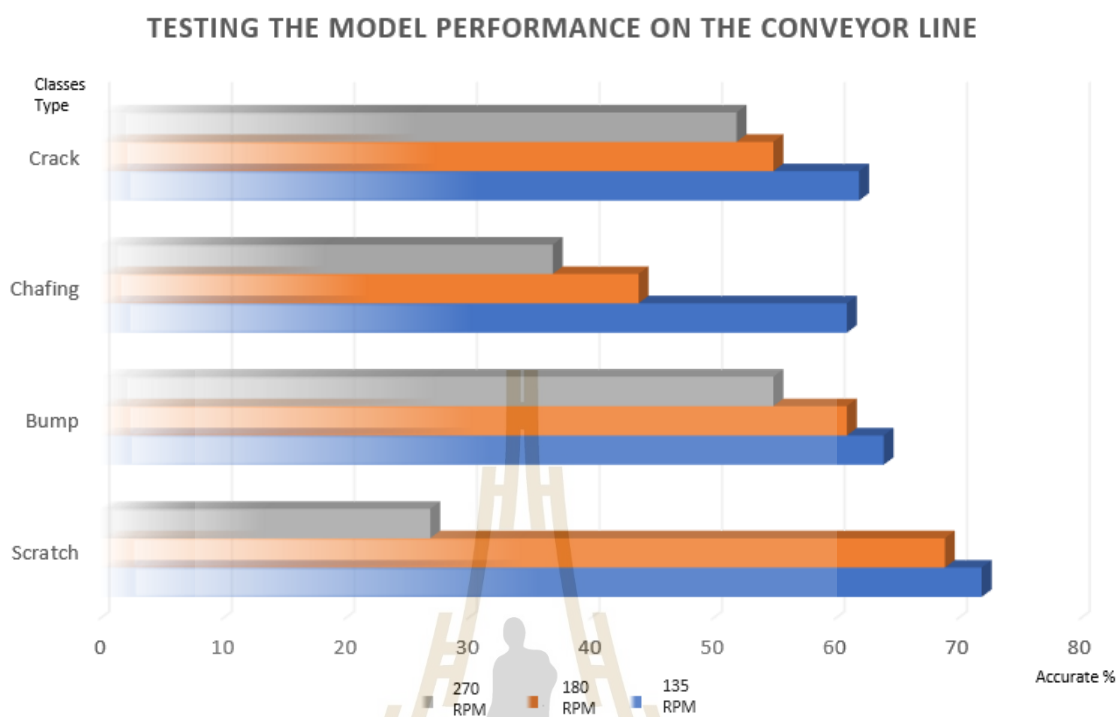


รูปที่ 4.4 การตรวจสอบรอยแตกและรอยวัสดุกับแม่แรงรถยนต์ขนาด 5 ตัน

4.5 การทดสอบประสิทธิภาพโมเดลกับข้อจำกัดของอุปกรณ์และอัลกอริทึม

การทดสอบข้อจำกัดของอุปกรณ์กล้องที่มีค่าเฟรมเรท 30 เฟรม/วินาที และอัลกอริทึมที่ใช้คุณลักษณะของรอยตำหนิแต่ละส่วนมาเชื่อมโยงงานกัน วิธีการทดสอบนำวัตถุวางบนสายพานลำเลียงปรับค่าความถี่ของตัวควบคุมความเร็วมอเตอร์เริ่มต้นย่านความถี่ 50 Hz เพื่อปรับมอเตอร์ให้หมุนสายพานความเร็ว 3 ระดับ คือ 135 รอบต่อนาที 180 รอบต่อนาที 270 รอบต่อนาที และเก็บผลลัพธ์ตามตารางที่ 4.2 ผลจากการทดสอบโมเดล ประสิทธิภาพความแม่นยำการตรวจสอบรอยตำหนิลดลงตามลำดับเมื่อวัตถุเคลื่อนที่เร็วขึ้น เนื่องจาก อัตราเฟรมเรทของกล้องมีผลโดยตรงต่ออัลกอริทึม

ตารางที่ 4.2 การทดสอบประสิทธิภาพโมเดลกับข้อจำกัดของอุปกรณ์และอัลกอริทึม



4.6 สรุป

จากผลการศึกษาและวิเคราะห์ประสิทธิภาพโมเดลต้นแบบ โดยใช้หลักการทฤษฎี Confusion Matrix มีความแม่นยำของโมเดล 95.6 เปอร์เซ็นต์ ทดสอบความถูกต้องในแยกแยะรอยตำหนิวัตถุ กับตัวอย่างที่เตรียมไว้ 400 ตัวอย่าง มีความถูกต้อง 96% ดำเนินการทดสอบกับวัตถุจริงแบบเรียลไทม์โดยจำลองลักษณะการปฏิบัติงานเสมือนกระบวนการตรวจสอบคุณภาพชิ้นงานจริงในกระบวนการผลิต ผลลัพธ์โมเดลต้นแบบสามารถตรวจสอบและแยกแยะประเภทได้อย่างถูกต้อง และระบุรอยตำหนิได้มากกว่า 1 ตำแหน่ง ซึ่งทั้ง 3 ผลลัพธ์นี้ บรรลุตามวัตถุประสงค์ของงานวิจัยที่กำหนดไว้ นอกจากนี้ เพื่อเป็นการศึกษาข้อจำกัดของอุปกรณ์โมเดลต้นแบบเพื่อนำไปพัฒนาในงานด้านอื่น ๆ ทางผู้วิจัยได้ทำการทดสอบการตรวจสอบและแยกแยะรอยตำหนิขณะวัตถุกำลังเคลื่อนที่ตามความเร็วสายพาน ผลลัพธ์คือความแม่นยำลดลงเรื่อย ๆ แปรผันตามความเร็ววัตถุที่เพิ่มขึ้น

บทที่ 5

สรุปผลการวิจัย

5.1 สรุปผลการวิจัย

งานวิจัยฉบับนี้มีวัตถุประสงค์เพื่อศึกษาอัลกอริทึมการประมวลผลรอยตำหนิวัตถุ นำมาใช้เพิ่มประสิทธิภาพกระบวนการตรวจสอบคุณภาพในอุตสาหกรรมยานยนต์ เพื่อพัฒนาอุปกรณ์ประมวลผลภาพประยุกต์ใช้ร่วมกับปัญญาประดิษฐ์การเรียนรู้เชิงลึก โดยสามารถระบุรอยตำหนิวัตถุแบบเรียลไทม์อัตโนมัติ และเพื่อพัฒนาอุปกรณ์ที่สามารถแยกแยะรอยตำหนิวัตถุแทนการมองเห็นมนุษย์จากผลการศึกษาและวิเคราะห์การออกแบบและพัฒนาโมดูลการมองเห็นแบบอัจฉริยะสำหรับประยุกต์ใช้ในกระบวนการตรวจสอบคุณภาพชิ้นงานและแยกแยะตำหนิวัตถุด้วยเทคนิคทางปัญญาประดิษฐ์ บรรลุตามวัตถุประสงค์ที่กำหนดไว้ทั้ง 3 ข้อ โมเดลมีประสิทธิภาพค่าความแม่นยำ 95.6% เพอร์เซ็นต์ ค่าความถูกต้อง 96 เพอร์เซ็นต์ และสามารถตรวจสอบและแยกแยะรอยตำหนิแบบเรียลไทม์ทั้ง 3 รุ่น ได้แก่ แม่แรงรถยนต์ขนาด 2 ตัน แม่แรงรถยนต์ขนาด 3 ตัน และแม่แรงรถยนต์ขนาด 5 ตัน โดยจำลองขั้นตอนการตรวจสอบเสมือนกระบวนการตรวจสอบคุณภาพในกระบวนการผลิตชิ้นงาน นอกจากนี้อัลกอริทึมที่นำมาประยุกต์ใช้สามารถระบุตำแหน่งของรอยตำหนิได้มากกว่า 1 ตำแหน่ง

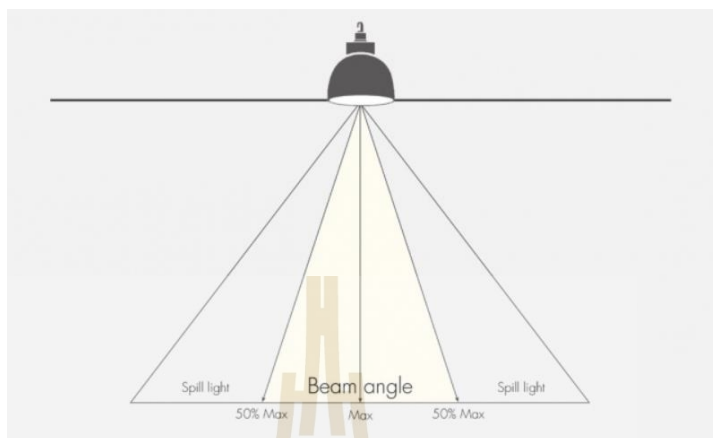
สรุปได้ว่า อุปกรณ์โมเดลต้นแบบของงานวิจัยฉบับนี้ สามารถนำไปพัฒนาและประยุกต์ใช้ในกระบวนการตรวจสอบคุณภาพในกระบวนการผลิตจริงได้ สามารถที่จะลดปัญหา Human Error ของมนุษย์ที่มีข้อจำกัดทางกายภาพของร่างกายและเพิ่มประสิทธิภาพการตรวจสอบรอยตำหนิในหลาย ๆ จุดของพื้นผิววัตถุ

5.2 ปัญหาที่พบในการทำวิจัย

5.2.1 ปัญหาแสงสว่างที่ไม่เหมาะสมทำให้มีแสงสะท้อนบนผิววัตถุ

การใช้แหล่งกำเนิดแสงที่ไม่เหมาะสมส่งผลต่อการประมวลผลของอัลกอริทึม ตอนช่วงแรกทดสอบการประมวลผลของอุปกรณ์ใช้แหล่งกำเนิดแสงจากสภาพแวดล้อมทั่วไป ตามรูปที่ 5.1 ซึ่งตรงจุดตรงกลางแสงจะตั้งฉาก ทำให้บางจุดค่าความสว่างมากเกินไปทำให้เกิดแสงสะท้อนที่วัตถุบางจุด

แสงสว่างน้อยเกินไป ค่าความสว่างจะค่อยๆลดลงตามลำดับ ถ้ารอยตำหนิอยู่ด้านข้างแม่แรงรถยนต์จะทำให้กล้องตรวจไม่พบ



รูปที่ 5.1 องศาของแสงแนวตั้งฉากกับวัตถุ

(ที่มา : <https://lightingequipmentsales.com/beam-angle.html>)

จากปัญหาดังกล่าว ดำเนินการแก้ไขโดยใช้วิธีการควบคุมสภาพแวดล้อมแสง โดยสร้างโครงสร้างกล่องที่มีขนาด กว้าง 45 เซ็นติเมตร x ยาว 60 เซ็นติเมตร x สูง 35 เซ็นติเมตร เพื่อป้องกันแสงสะท้อนจากภายนอก และติดตั้งหลอดไฟLED ที่มีขนาดความยาว 50 เซ็นติเมตร เพื่อให้แสงสว่างครอบคลุมพื้นที่ทั้งหมด และใช้ Dimmer ปรับค่าความสว่างที่เหมาะสม 1,800 ลักซ์

5.2.2 ข้อจำกัดของงานวิจัย

เนื่องจากงานวิจัยฉบับนี้ ใช้กล้องทั่วไปที่มีขนาดการรับข้อมูลภาพมีความละเอียดสูงสุดแค่ 1080 พิกเซล เฟรมเรทสูงสุด 30 เฟรมต่อวินาที และใช้อัลกอริทึมโครงข่ายประสาทเทียมมีรูปแบบการหาคุณลักษณะของภาพแล้วนำมาเชื่อมต่อกันเพื่อหาค่า weight ถ้านำโมเดลต้นแบบนี้ใช้ตรวจสอบกับวัตถุที่กำลังเคลื่อนที่ จะทำให้ค่าความแม่นยำและค่าความถูกต้องของโมเดลต้นแบบลดลงตามลำดับ

5.2 ข้อเสนอแนะ

การเก็บข้อมูลภาพตัวอย่าง ขั้นตอนนี้ควรเก็บข้อมูลภาพที่ชัดเจนและขนาดที่เหมาะสมเพื่อง่ายต่อการเตรียมข้อมูลสำหรับการเทรนนิ่งโมเดล เนื่องจาก ถ้าได้ภาพที่มีลักษณะไม่คมชัดและมีรอยตำหนิที่เล็กน้อยจะยากต่อการปรับความสว่างของภาพและการทำ Contrast

การควบคุมสภาพแวดล้อมและค่าความสว่างของแสง ควรกำหนดค่าความสว่างที่เหมาะสม และคงที่เพื่อให้กล้องสามารถเห็นรอยตำหนิได้ชัดเจน ถ้าแสงสว่างไม่เพียงพอจะทำให้รอยตำหนิดูมืดและถ้าสว่างมากเกินไปจะทำให้เกิดแสงสะท้อนที่ชิ้นงาน ทั้ง 2 ปัจจัยนี้มีผลต่ออัลกอริทึมโดยตรง

5.3 ข้อสังเกต

กระบวนการตรวจสอบคุณภาพชิ้นงานแม่แรงรถยนต์ มีหลายปัจจัยสาเหตุทำให้เกิดปัญหา รอยตำหนิบนผิววัตถุ เช่น รอยสีลอก โดยเกิดจากน้ำมันตกค้างที่ผิววัตถุติดจากกระบวนการกลึงและ เจาะรูชิ้นงาน หรือรอยกระแทกต่าง ๆ ที่เกิดจากปัญหาระบบสายพานลำเลียงชิ้นงานเกิดการกระชาก ขณะมอเตอร์หมุน ปัจจัยเหล่านี้มีการเกิดขึ้นเป็นครั้งคราวและไม่สามารถอาจคาดคะเนการเกิดได้ รวมถึงพนักงานทำการตรวจเช็คแต่ละวันมีจำนวนปริมาณมาก เพราะต้องปฏิบัติงานให้ทันตาม ความเร็วของเตาสายพาน ทำให้บางครั้งการตรวจสอบคุณภาพของชิ้นงานไม่สามารถได้ผล 100 เปอร์เซ็นต์ ดังนั้น ความสามารถของอุปกรณ์โมเดลต้นแบบนี้เปรียบเสมือนระบบป้องกันความ ผิดพลาดหรือที่เรียกว่า POKA-YOKE ซึ่งสามารถช่วยตรวจเช็ครอยตำหนิเพื่อป้องกันปัญหา Human Error ได้อย่างดี



รายการอ้างอิง

- รศ.ดร.อาทิตย์ ศรีแก้ว (2555). การออกแบบและพัฒนามอดูลการมองเห็นแบบชาญฉลาดสำหรับ
ประยุกต์ใช้งานด้านเกษตรกรรมด้วยเทคนิคทางปัญญาประดิษฐ์. **โครงการวิจัย รหัสSUT7-
711-55-12-70 สาขาวิศวกรรมไฟฟ้า สำนักวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยี
สุรนารี**
- รศ.ดร.สนั่น ศรีสุข. เอกสารการประมวลผลภาพขั้นสูง. ISBN:978-974-8242-80-4 คณะ
วิศวกรรมศาสตร์คอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนครพนม
- ผศ.ดร.ชูพันธุ์ รัตนโกศา (2559). เอกสารความรู้เบื้องต้นทางปัญญาประดิษฐ์. **ภาควิชาเทคโนโลยี
วิศวกรรม อิเล็กทรอนิกส์ วิทยาลัยเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระ
จอมเกล้าพระนครเหนือ**
- วิโรจน์ กิตติวรปรีดา. เอกสารประกอบการสอน วิชา ไมโครคอนโทรลเลอร์. **สาขาวิชาช่าง
อิเล็กทรอนิกส์ วิทยาลัยเทคนิคราชบุรี**
- ธวัช รัตนชัย (2555). การปรับปรุงกระบวนการตรวจสอบแผ่นแม่เหล็กเก็บข้อมูลที่นำกลับมาใช้ใหม่
ด้วยระบบการมองเห็นของคอมพิวเตอร์. **วิทยานิพนธ์ปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาแมคคาทรอนิกส์ มหาวิทยาลัยเทคโนโลยีสุรนารี**
- เอกชัย พรรณวัลย์ และ สุภัทรชัย สุดสวาท (2559). การตรวจสอบชิ้นงานอัตโนมัติผ่านระบบการ
ประมวลผลภาพในการผลิตชิ้นส่วนยานยนต์. **วารสารวิทยาศาสตร์ประยุกต์ มหาวิทยาลัย
พระจอมเกล้าพระนครเหนือ วิทยาเขตระยอง**
- Venkat Anli Adibhatla, Huan-Chuang Chih, Chi-Chang Hsu, Joseph Cheng, Maysam F.
Abbod and Jiann-Shing Shieh, (2021). Applying deep learning to detect
detection in printed circuit boards via a newest model of you-only-look-once,
Mathematical Biosciences and Engineering.
- YITING LI, HAISONG HUANG, QIPENG CHEN, (Student Member, IEEE), QINGSONG FAN and
HUAFENG QUAN, (2021). Research on a Product Quality Monitoring Method
Based on Multi Scale PP-YOLO. **IEEE Access.**
- Dr. S.V. Viraktamath, Madhuri Yavagal and Rachita Byahatti, (2021). Object Detection
and Classification using YOLOv3. **International Journal of Engineering
Research & Technology (IJERT) ISSN : 2778-0181.**

- Wentao Liu, Zhangyu Wang, Bin Zhou, Songyue Yang and Ziren Gong, (2021). Real-time Signal Light Detection based on Yolov5 for Railway. **Earth and Environmental Science**.
- OMAR MOURED (2020), Evaluation of deep learning based multiple object trackers, **MIDDLE EAST TECHNICAL UNIVERSITY**.
- THAILAND AUTOMOTIVE INSTITUTE. Statistics of Thai automotive industry [Online]. Available :<https://data.thaiauto.or.th/>.[Accessed:18-August-2020]
- Dr.Panadda (2014). Surface Engineering. **Material and Surface Technology Damage Analysis Laboratory National Metal and Materials Technology Center**
- Narathip (2014). Quality Control. Faculty of Technology, **UDON THANI RAJABHAT UNIVERSITY**.
- Akekachai Pannawan and Supattarachai Sudsawat. Automated part inspection by image processing system in vehicle part manufacturing. **ISSN 1512-7805 Printed in Thailand**.
- Kwanchanok. Artificial Intelligence. [Online].Available: <https://www.lib.ku.ac.th/2019/index.php/covid-19/1045-artificialintelligence> [Accessed:30-August-2020]
- Napparut Pattanasarn, Nattavut Sriwiboon, (2020). Image Processing for Classifying the Quality of the Chok-Anan Mango by Simulating the Human Vision using Deep Learning. Department of Informatics and Computer. **Faculty of Science and Health Technology. Kalasin University**.
- Ashwin Bhandare, Maithili Bhide, Pranav Gokhale, Rohan Chandavarkar, (2016). Applicatoin of Convolutional Neural Networks. Department of Computer Engineering, **Pune Institute of Computer Technology Savitribai Phule Pune University**.
- ODSC (2021). Overview of the Yolo Object Detection Algorithm. [Online]. Available: <https://medium.com/@ODSC/overview-of-the-yolo-object-detection-algorithm-7b52a745d3e0>[Accessed:28-September-2021].
- DAT (2021). CNN Yolo Algorithm. [Online]. Available: <http://datahacker.rs/object-detection-yolo-algorithm/>-Accessed:28-September-2021].

- Mark Harlan Goadrich, (2007). Learning Ensembles of First-Order Clauses That Optimize Precision-Recall Curves. **UNIVERSITY OF WISCONSIN-MADISON.**
- Stardet (2021). Confusion Matrix. [Online]. Available: <https://www.stardat.net/post/confusion-matrix>[Accessed:30-April-2021].
- 305393 Special Topic in Human Computer Interaction, (2021). How to adjust the Brightness and Contrast of the image by using OpenCV with Visual Studio. [Online]. Available: <https://305393advancedtopicsinhci.wordpress.com/2017/02/15/how-to-adjust-the-brightness-and-contrast-of-the-image-by-using-open-cv-with-visual-studio/>[Accessed:20-April-2021].
- Sitthitouch (2018). Generating an architectural model using object recognition to observe human movement in public space case study: 3-dimensional parametric model of bus station platforms. **Ref. code: 25605816030075HOB.**
- Grace Karmini (2021) Introduction to Yolo Algorithm for object detection. [Online]: Available: <https://www.section.io/engineering-education/introduction-to-yolo-algorithm-for-object-detection/>[Accessed:28-September-2021].
- L.Shao,A.Mahajan,T.Schreck, D.J.Lehmann (2017). Interactive Regression Lens for Exploring Scatter Plots. **Eurographics Conference on Visualization (EuroVis)**
- Mr.PL, (2021) Deep Learning EP.3 with Keras with Image Classification. [Online]. Available: <https://medium.com/mmp-li/deep-learning-ep-3-keras-with-image-classification-2d5d95ee5613>[Accessed:15-August-2021].



ภาคผนวก ก

คำสั่งติดตั้งเทอร์นิงโมเดล

ก.1 คำสั่งโค้ดดังต่อไปนี้ป้อนข้อมูลเข้าไปในโมเดลเพื่อทำการเทรนนิ่ง

```
python train.py --img 640 --batch 20 --epochs 100 --data ./data/defect/defection.yaml
--cfg ./models/yolov5s.yaml --weights ./weights/yolov5s.pt --cache
```

```

C:\Users\trais\OneDrive\งาน\วิชา\LearnOpenCV\cmd.py - Sublime Text (UNREGISTERED)
File Edit Selection Find View Goto Tools Project Preferences Help
trainingCodeMethod.py  defectKey.py  cmd.py  defection.yaml  yolov5s.yaml  train.py  t.py  Multithread.py  puttybit  PuttytersonLamp.txt  RunProgramonJetson.txt  Te
143 train.py error: unrecognized arguments: --cache images
144 (defection) PS C:\Users\trais\OneDrive\งาน\วิชา\LearnOpenCV> python train.py --img 640 --batch 10 --epochs 10 --data ./data/defect/defection.yaml --cfg ./models/yolov5s.yaml --weights ./weights/yolov5s.pt
145 train: weights=./weights/yolov5s.pt, cfg=./models/yolov5s.yaml, data=./data/defect/defection.yaml, hyp-data/hyps/hyp.scratch.yaml, epochs=10, batch_size=1, imgsz=640, rect=False, resume=False, nosave=False, noval=False
146 github: up to date with https://github.com/ultralytics/yolov5
147 YOLOv5 v5.0-372-g8baccad torch 1.9.0-cpu-cv2
148
149 hyperparameters: lr=0.01, lr_f=0.2, momentum=0.937, weight_decay=0.0005, warmup_epochs=1.0, warmup_momentum=0.8, warmup_bias_lr=0.1, box=0.05, cls=0.5, cls_pw=1.0, obj=1.0, obj_pw=1.0, iou_t=0.2, anchor_t=4.0, fl_gamma=0.0
150 Weights & Biases: run 'pip install wandb' to automatically track and visualize YOLOv5 runs. (https://wandb.ai)
151 TensorBoard: Start with 'tensorboard --logdir runs/train', view at http://localhost:6000/
152 Downloading https://github.com/ultralytics/yolov5/releases/download/v5.0/yolov5s.pt to weights/yolov5s.pt...
153 100%
154
155 Overriding model.yaml nc=80 with nc=2
156
157 from n      params module                                arguments
158 0          -1 1      3520  models.common.Focus                                [1, 32, 1]
159 1          -1 1     18560 models.common.Conv                                [256, 64, 3, 2]
160 2          -1 1     18816 models.common.C3                                  [64, 64, 1]
161 3          -1 1     73984 models.common.Conv                                [64, 128, 3, 2]
162 4          -1 1     36928 models.common.C3                                  [128, 128, 1]
163 5          -1 1    295424 models.common.Conv                                [128, 256, 3, 2]
164 6          -1 1    825352 models.common.C3                                  [256, 256, 1]
165 7          -1 1   1180072 models.common.Conv                                [256, 512, 3, 2]
166 8          -1 1   6568904 models.common.SPP                                [512, 512, [5, 9, 13]]
167 9          -1 1   1822720 models.common.C3                                  [512, 512, 1, False]
168 10         -1 1   131504 models.common.Conv                                [512, 256, 1, 1]
169 11         -1 1      0 torch.nn.modules.upsampling.Upsample            [None, 2, 'nearest']
170 12         -1 1      0 models.common.Concat                            [1]
171 13         [-1, 6] 1      0 models.common.C3                                  [512, 256, 1, False]
172 14         -1 1    33824 models.common.Conv                                [256, 128, 1, 1]
173 15         -1 1      0 torch.nn.modules.upsampling.Upsample            [None, 2, 'nearest']
174 16         [-1, 4] 1      0 models.common.Concat                            [1]
175 17         -1 1    90880 models.common.C3                                  [256, 128, 1, False]
176 18         -1 1    44712 models.common.Conv                                [128, 128, 3, 2]
177 19         [-1, 14] 1      0 models.common.Concat                            [1]
178 20         -1 1   256448 models.common.C3                                  [256, 256, 1, False]
179 21         -1 1   598136 models.common.Conv                                [256, 256, 3, 2]
180 22         [-1, 10] 1      0 models.common.Concat                            [1]
181 23         -1 1   1122720 models.common.C3                                  [512, 512, 1, False]
182 24         [17, 20, 23] 1   18679 models.yolo.detect                                [24, [10, 13, 16, 30, 33, 23], [30, 61, 62, 45, 59, 119], [116, 90, 156, 498, 373, 326]], [128, 256, 512]]
183 Model Summary: 283 layers, 7866239 parameters, 7866239 gradients, 16.4 GFLOPs
184
185 Transferred 354/362 items from weights/yolov5s.pt
186 Scaled weight_decay = 0.0005
187 optimizer: SGD with parameter groups 99 weight, 62 weight (no decay), 62 bias
188 train: Scanning data dir: ./data/defect/images and labels... 6 found, 0 missing, 0 empty, 0 corrupted: 100%
189 train: New cache created: data/defect11labels11train.cache
190 val: Scanning data dir: ./data/defect/images and labels... 2 found, 0 missing, 0 empty, 0 corrupted: 100%
191 val: New cache created: data/defect11labels11val.cache
192 Plotting labels...
193
194 autoanchor: Analyzing anchors... anchor
195 s/target = 5.19, Best Possible Recall (BPR) = 1.0000
196 Image sizes 640 train, 640 val
197 tiling 0 data loader time:
  
```

รูป ก.1 โค้ดดังต่อไปนี้ป้อนข้อมูลเข้าโมเดล

มหาวิทยาลัยเทคโนโลยีสุรนารี

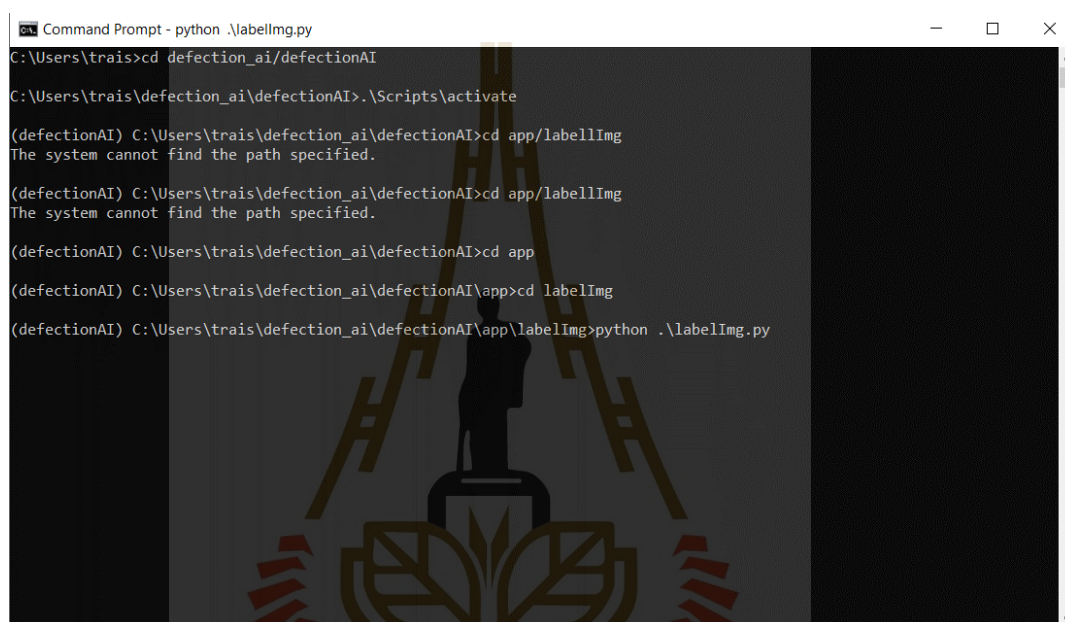


ภาคผนวก ข

การใช้โปรแกรม Label IMG

ข.1 การสร้างเรียกใช้โปรแกรมผ่านทาง CMD

ในขั้นตอนนี้ตามที่อธิบายในบทที่ 3 วิธีดำเนินงานวิจัย หลัก ๆ ขั้นตอนนี้ใช้สำหรับเทรนนิ่งโมเดลในการจำแนกประเภทสิ่งที่เราต้องการให้โมเดลเรียนรู้ ใช้เครื่องมือ Label IMG และลงไลบรารี รวมถึงการสร้างEnvironmentเฉพาะเพื่อป้องกันการลงไลบรารีทับซ้อนกับไลบรารีหลักของเครื่องคอมพิวเตอร์ จากรูป ข.1 เราเรียกใช้Environmentที่ได้สร้างก่อนหน้าซึ่งชื่อว่า defectionAI และเรียกใช้โปรแกรม Label IMG

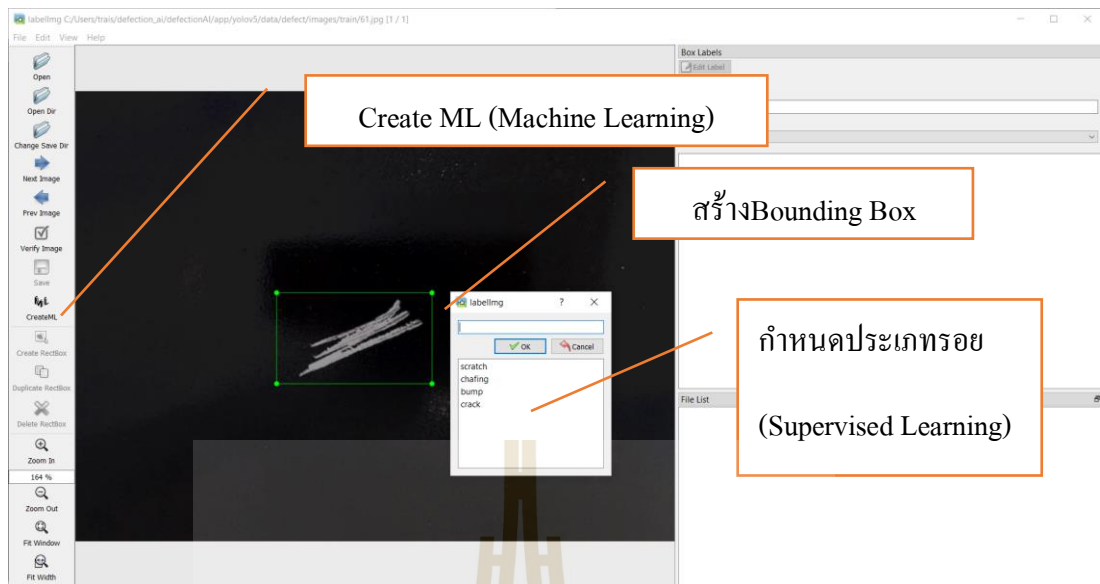


```

Command Prompt - python .\labelImg.py
C:\Users\trais>cd defection_ai\defectionAI
C:\Users\trais\defection_ai\defectionAI>. \Scripts\activate
(defectionAI) C:\Users\trais\defection_ai\defectionAI>cd app\labelImg
The system cannot find the path specified.
(defectionAI) C:\Users\trais\defection_ai\defectionAI>cd app\labelImg
The system cannot find the path specified.
(defectionAI) C:\Users\trais\defection_ai\defectionAI>cd app
(defectionAI) C:\Users\trais\defection_ai\defectionAI\app>cd labelImg
(defectionAI) C:\Users\trais\defection_ai\defectionAI\app\labelImg>python .\labelImg.py
  
```

รูปที่ ข.1 การเรียกใช้โปรแกรม Label IMG

วิธีการใช้โปรแกรม Label IMG จะต้องกำหนดประเภทไฟล์ ในงานวิจัยฉบับนี้เลือกไฟล์นามสกุล .txt เพื่อบันทึกข้อมูลกรอบBounding Box และกำหนดประเภทของรอยบนวัตถุ ตามรูปที่ ข.2

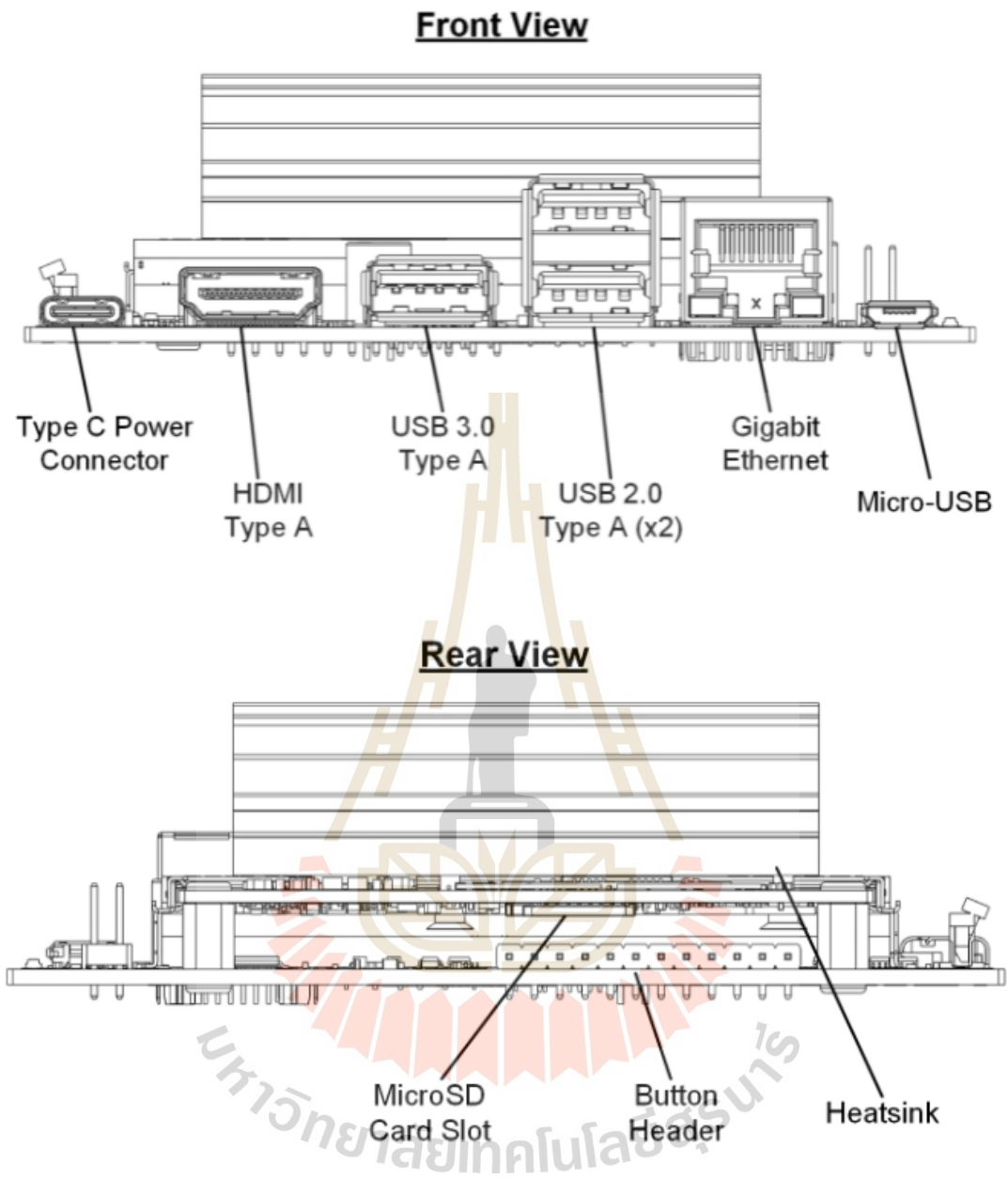


รูป ข.2 การใช้โปรแกรม Label IMG

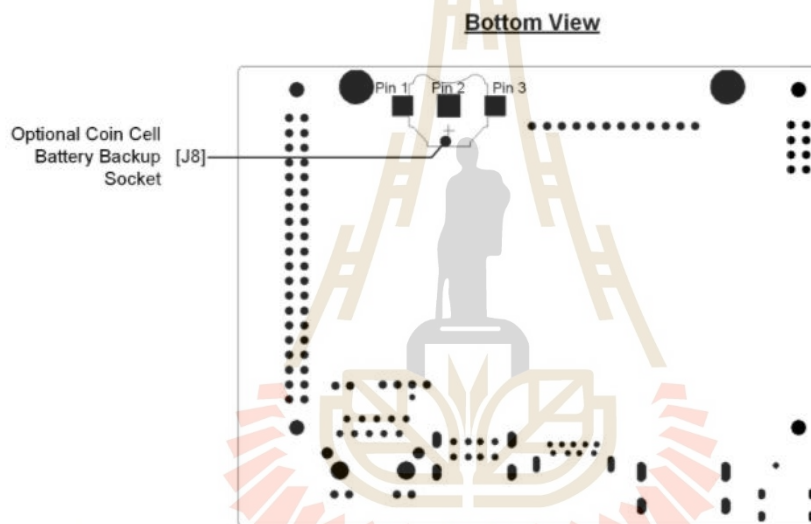
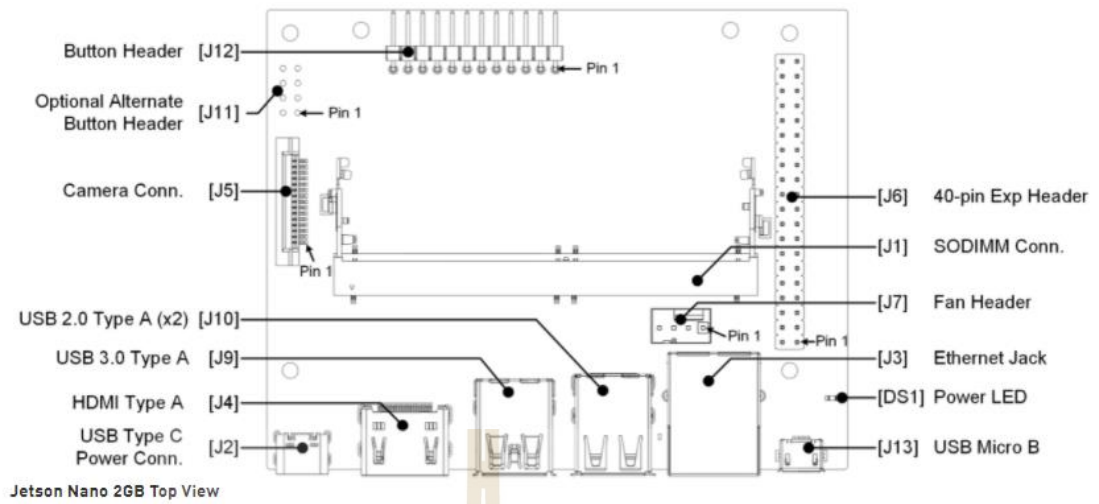


ภาคผนวก ค

โครงสร้างอุปกรณ์ NVIDIA Jetson Nano



ค.1 โครงสร้างอุปกรณ์ NVIDIA Jetson Nano ด้านหน้าและหลัง



ค.2 โครงสร้างอุปกรณ์ NVIDIA Jetson Nano ด้านบนและด้านล่าง



ภาคผนวก ง

บทความวิชาการที่ได้รับการตีพิมพ์เผยแพร่ในระหว่างการศึกษา

มหาวิทยาลัยเทคโนโลยีสุรนารี

รายชื่อบทความที่ได้รับการตีพิมพ์เผยแพร่ในระหว่างการศึกษา

Traisirin Saengseangfa and Suradet Tantrairatn, (2021). **DESIGN AND DEVELOPMENT OF A VISION SYSTEM FOR IDENTIFYING A WORKPIECE QUALITY AND DETECTS USING OBJECT DETECTION.** 25TH International Computer Science and Engineering Conference. 18th-20th November, 2021 at Phayao University, Thailand.



Design and development of a vision system for identifying a workpiece quality and defects using object detection

Traisirin Saengseangfa and Suradet Tantrairatn

School of Mechatronics Engineering, Institute of Engineering

Soranee University of Technology, 111 university Avenue, Muang District, Nakhon Ratchasima, Thailand

Email: traisirin@gmail.com, suradety@sut.ac.th,

Abstract— In automotive parts manufacture, the quality of the products is the most important issue for a manufacturer. This research presents the design and development of an intelligent vision module classifying the defect on object surfaces using image processing, artificial intelligence, convolutional neural network algorithm as if using human vision. The developed prototype was a neural network-supported computing system for classifying the defects on the surface of black car jacks of 10x10x19 cm (W x L x H) in an environment with specific brightness. The test result was the average precision of 95.6%, showing that it could detect defects on the four types of surfaces previously specified in the Deep Learning process. It could also detect more than one defects on the same object on the belt conveyor system in real time without underfitting or overfitting problem.

Keywords—Image Processing, Artificial Intelligence, Convolution Neural Network, Deep Learning

I. INTRODUCTION

Thailand depends largely on exports and domestic use in automotive industry [1], aerospace industry, shipyard industry, electrical appliance and electronics industry, etc. In those industry, product quality control process is, without doubt, required. For example, the spray-painting process, as well as many other coating methods [2] depending on the types of the manufactured, is widely performed to enhance the durability of workpieces. Then, there comes the next process, quality control, to inspect conditions of the outer part of the workpieces, whether there are scratches, chafing, bumps and cracks or not.

Quality control [3] inspecting products according to the specified features is an essential process every industry has to go through before products are delivered to consumers. To improve this process's efficiency, technologies to process images is regarded to be useful as, currently, there are algorithms developed in accordance with the development of the computer industry.

In addition, Artificial intelligence (AI) is considered to be suitable to perform analysis of objects located at any places as instructed in place of human vision. Hence, in this research, a prototype able to detect the defects on objects has been developed. It was equipped with NVIDIA Jetson Nano device supporting neural networks to classify the detected defects into four types according to previously input data; it was also equipped with a microcontroller to send signals indicating errors detected on the surface of objects.

II. THEORY AND METHOD

A. Algorithm

• Image Processing

Image processing [4] is photograph manipulation with computers to improve or change the quality of images so that they will be more suitable for further analysis and other purposes. Images are, firstly, retrieved (Image Acquisition), compressed (Image Compression) and enhanced (Image Enhancement) by adjusting the brightness and the color sharpness of each pixel to make them suitable for Deep Learning process. It is commonly used in the process of inspection and quality control in every industry to help reducing operation time and increase efficiency as well as reliability of inspection which is useful for industries requiring high accuracy.

• Artificial Intelligence

Artificial Intelligence [5] is a computing system modeled upon neural networks in human brains. It can learn and optimize processing upon the increasing amount of data through a self-learning process consisting of memorizing, thinking, analyzing, learning and linking complex information at fast speed. It collects large amount of data at high speed and processes them repeatedly through intelligent processing steps.

• Deep Learning

Deep learning [6] is a computer technology developed to be able to imitate human behaviors. It has a computational-thinking process called a neural network (NN) as it works like humans' neural networks. The specific characteristic of deep learning is that, when applied to image processing to classify the defects on target objects, it can simulate conditions and find answers by piling hidden layers, multiple layers of neural networks, to more than two layers to achieve learning and simulation. Therefore, it can be concluded that the more layers of the neural networks there are in processing, the more accurate deep learning can become.

• Convolutional Neural Network

Convolutional Neural Network is one of the deep learning methods that simulates human vision. It is able to differentiate the features of objects such as colors, borders, color contrast and keep them within themselves. As a result, it can recognize what those features are and classify them. Convolutional neural network uses aforementioned hidden layers [7] to

identify features and iterates until the answers with accuracy according to the classification come out. Convolutional neural network analyzes the relationship between the extracted features in each layer and the outcomes. How it works is divided into 3 parts as follows. (Fig. 1)

- 1) Inputs are the parts where data or objects are input; it functions as human vision.
- 2) Hidden layers are the parts where processing is performed into layers to do training and classification; it functions as human brains.
- 3) Outputs are the parts that show the results of feature classification resulting from using multiple hidden layers analyzing each image until the answer is output.

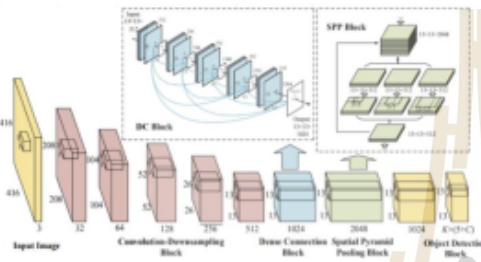


Fig. 1. Convolutional Neural Network

B. Application of YoloV5 algorithm

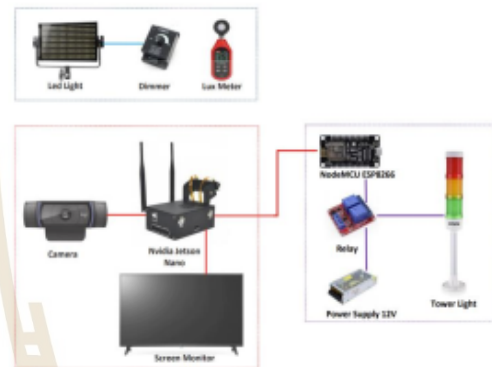
Concerning the YOLO (You Only Look Once) real-time object detection algorithm, which is one of the most effective object detection algorithms that also encompasses many of the most innovative ideas coming out of the computer vision research community. Object detection is a critical capability of autonomous vehicle technology. It is an area of computer vision that is exploding and working so much better than just a few years ago. At the end of this article, we will see a couple of recent updates to YOLO by the original researchers of this important technique. YOLO uses a totally different approach. YOLO is a clever convolutional neural network (CNN) for doing object detection in real-time. The algorithm applies a single neural network to the full image, and then divides the image into regions and predicts bounding boxes and probabilities for each region. These bounding boxes are weighted by the predicted probabilities [8]. Therefore, we have decided to use Yolo version 5 to apply with this thesis.

YOLO version 5 algorithm is chosen to apply this thesis because of the following reasons:

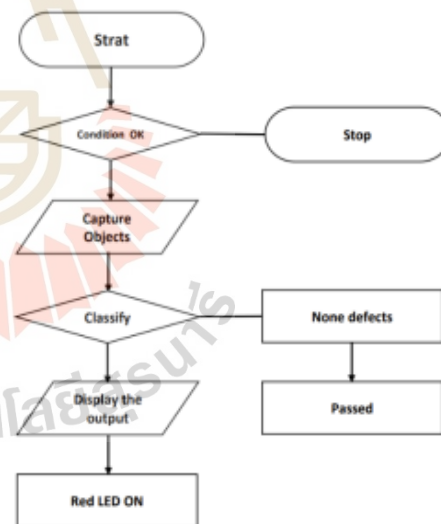
- Speed: This algorithm improves the speed of detection because it can predict objects in real-time.
- High accuracy: YOLO is a predictive technique that provides accurate results with minimal background errors.
- Learning capabilities: The algorithm has excellent learning capabilities that enable it to learn the representations of objects and apply them in object detection [9].

C. Processing system and microcontroller

The prototype was equipped with the Nvidia Jetson Nano computer board with a 128-core GPU and Linux-based system for convolutional neural network algorithms. It received data from a 30-fps Logitech Full-HD camera for processing. A general monitor connected with the computer board with a HDMI cable was used to display the system operation. In addition, LED lights were installed along with an adjustable power supply device for optimal light output. When the main system detected a problem on the surface of an object, it would send commands in serial communication at a speed of 115,200 through a digital pin for the relay to supply current to the coil. As a result, the red light bulb would light up as shown in Fig. 2.



(a) Processing system of the prototype



(b) Working process

Fig. 2. The prototype system and working process

D. Installation of equipment

The Logitech camera was placed 25 centimeters far from the target objects. Above its left and right sides were LED lights whose brightness was fixed at 1,800 lux because of a dimmer supplied with power to convert 220v AC to 12-24v DC. A solid plastic structure of 45x60x35 centimeter (W x L x H) was placed to avoid reflection of light. There was a belt conveyor system with a 20x120x75-cm belt (W x L x H); it was equipped with a 60W motor using a speed controller as shown in Fig 3.



Fig. 3. Installation of equipment

E. Measuring Performance

The prototype's performance efficiency was measured according to Confusion Matrix [10]. The results were classified into four values as follows (Fig 4).

- 1) True Positive (TP): it refers to the number of predictions where the classifier correctly predicts the positive class as positive.
- 2) True Negative (TN): it refers to the number of predictions where the classifier correctly predicts the negative class as negative.
- 3) False Positive (FP): it refers to the number of predictions where the classifier incorrectly predicts the negative class as positive.
- 4) False Negative (FN): it refers to the number of predictions where the classifier incorrectly predicts the positive class as negative.

The above 4 values, TP, TN, FP, and FN, are used to measure the efficiency with Precision, the ratio of correct positive predictions to the total predicted positives.

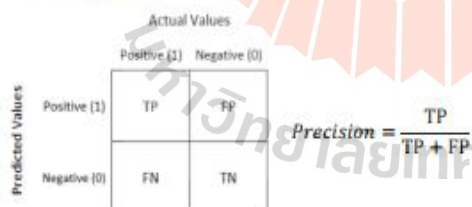


Fig. 4. Confusion Matrix size 2 x 2 matrix [11]

E. Data Preparation

The data used to train the model deep learning had to be prepared first. They were categorized according to the types previously defined to make labels. The dataset consisted of a total of 1,060 images divided into 4 categories below:

- 1) Scratch with average size of 3x15 millimeters (W x L)
- 2) Chafing with average size 10x20 millimeters (W x L)
- 3) Bump with average size 1x1.5 millimeters (W x L)
- 4) Crack with average size 0.5x15 millimeters (W x L)

Every image went through a cleaning process (Clean Data) and pre-processing process to ensure their accuracy and suitability accordingly as shown in Fig. 5.

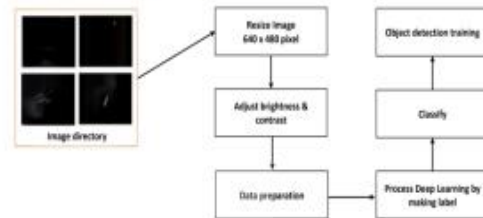


Fig. 5. Data Preparation

Moreover, every image had to be resized to 640 x 490 pixels to prevent processing failure due to difference in pixel sizes. Brightness and contrast of each also had to be adjusted [12] according to the equation below:

$$g(x, y) = af(x, y) + \beta$$

After that, a target location was set so that the system would learn what was there, leading to add metadata [13] to the images as well. Because this process affected the efficiency in identification and detection directly, a bounding box had to be created to determine the object range, neither bigger nor smaller than the target, maximizing accuracy and minimizing deviation as shown in Fig. 6.

- 1) p_c was probability for detected objects
- 2) b_w and b_h were the width and height of the bounding box
- 3) b_x, b_y was the center point of an object
- 4) C was a defined class.

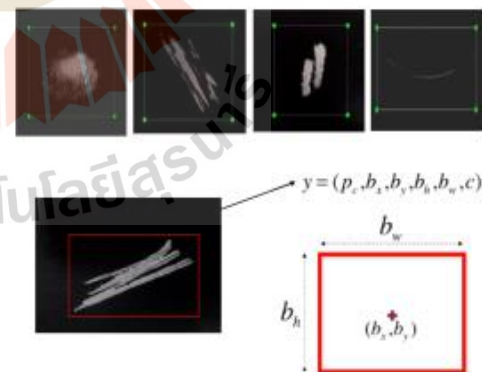


Fig. 6. Adding the Metadata (bounding box regression) [14]

Bounding box labels in each class were created for training. There were 1,060 label files in total, one label for one image as shown in TABLE I

TABLE I. AMOUNT LABELS EACH CLASS

No.	Classes	training	test	validation
1	scratch	165	55	45
2	chafing	165	55	45
3	bump	165	55	45
4	crack	165	55	45
	Total	660	220	180

In this process, the data had to be divided correctly to avoid underfitting, wrong outputs of the model, and overfitting, wrong outputs in actual uses despite being correct and precise during training. The latter greatly affects the accuracy of the model when the model's performance is tested with actual objects and the problems occur as shown in Fig. 7.

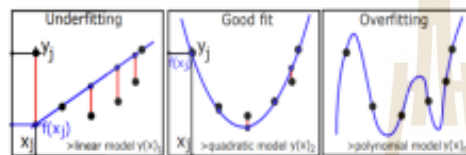


Fig. 7. Underfitting VS. Overfitting [15]

Among the dataset of 1,060 images in total, 660 images were used in the training for the model to learn the classes of different defect objects and 220 images to test the performance of the model. 180 images among those for training were used to test the performance after tuning had been finished.

Next, the already prepared dataset was input into the model for them to be processed. Unable to be input at once, they all were divided into small groups which was called a batch size. Suitable variables were specified to avoid the excessive usage of the CPU and RAM.



Fig. 8. Input the data to the model [16]

For Yolo's algorithm model, this research used Yolo version 5. The images' parameters were configured to have the 640-pixel size, the same as the images', to avoid problems during training. The batch sizes' parameters were set to 20, preventing excessive usage of CPU. The epoch parameters

were set to 100; by training with 53 iterations to get one epoch, the model carried out 5,300 iterations in total as in Table II.

TABLE II. AMOUNT LABELS EACH CLASS

Setting	Parameter
Model	YoloV5S
Dataset	1,060
Image size	640
Batch size	20
Epochs	100 * 3

III. EXPERIMENT RESULTS

A. The result of developing the model

Python and some libraries such as Keras, Tensorflow, Pytorch, Matplotlib, OpenCV, Numpy, Pandas, Torchvision, Scipy, and Scikit-learn used for deep learning process were applied to develop the model detecting the defects on objects and classify classes. More details on the installation of the libraries are provided in [17]. The result of the classification of the four classes, which were scratches, chafing, bumps and cracks, resulted in a relatively high precision of 95.6 percent against confidence as shown in Fig. 9.

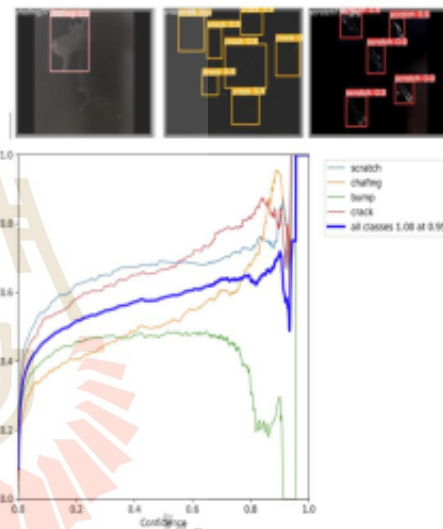


Fig. 9. Precision result from creating the model

According to the graph, the green line representing the Bump class had the lower precision range than the other classes because the bump had the same feature. To fix this, a previous feature was used for the next training and the number of iterations was increased.

B. Testing the model performance with the prepared samples

The model's performance was tested with previously prepared four types of samples which are scratches, chafing,

bumps, and cracks with 100 samples each, meaning 400 in total, to see if it could classify them correctly. According to TABLE III showing the test result, the success rate of classification using convolutional neural network algorithm was 96%.

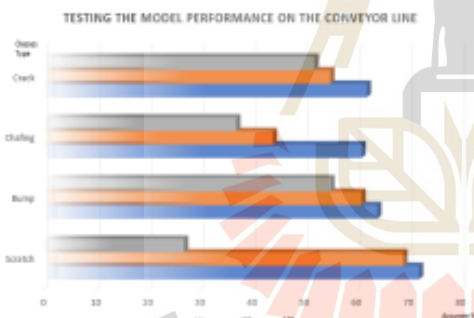
TABLE III. THE RESULT OF TESTING THE MODEL TO CLASSIFY EXAMPLES

No.	Class Type	Amount of the tested sample	Result	Accuracy (%)
1	scratch	100	98	98%
2	bump	100	94	94%
3	chafing	100	97	97%
4	crack	100	95	95%
	average			96%

C. Implementing the model in real-time testing on the conveyor system

The model's performance was tested on the conveyor system and actual car jacks in real time. Two variables, classified features and training data from the computer, were input into the NVIDIA Jetson Nano device with the installed libraries. Frequency of the speed controller of the conveyor's motor was adjusted to 50 Hz, controlling it to spin the belt at 3 speeds: 135 rpm, 180 rpm, and 270 rpm as shown in TABLE IV. The result shows that it could detect each defect on the objects which were running on the conveyor belt in real time, classifying them according to features of the defects correctly without underfitting or overfitting. Nevertheless, when the objects' moving speed increased, the camera's frame rate was affected, causing the accuracy to decrease accordingly.

TABLE IV. TESTING THE MODEL PERFORMANCE ON THE CONVEYOR LINE IN REAL-TIME



IV. EXPERIMENT RESULTS

This research to develop the prototype has integrated image processing algorithms, artificial intelligence performing deep learning and convolutional neural network algorithms to examine, analyze and classify the defects on objects. Firstly, dataset was collected from 4 types of samples, scratches; chafing; bumps; cracks. Then, convolutional neural network algorithms were applied to develop the ability of feature classification of the collected dataset which needed preparation using the principle from the related researches in order to avoid underfitting and overfitting when they were tested with actual objects.

the first test resulted in the average precision of all classes of 95.6%. The second one tested on various samples shows that it could classify the defects on the objects 96% correctly. Lastly, when the processing performance efficiency of the model in real time was tested on the belt conveyor system with a maximum speed of 270 rpm and on car jacks, it could detect and classify the defects on the objects accordingly. As a result, the three tests shows that this research on the design and development of an intelligent vision system for identifying a workpiece quality and defects using artificial intelligence can be an example of how technologies can be applied to encourage development and progress in various industries.

ACKNOWLEDGEMENT

The research would like to thank the King Mongkut's University of Technology North Bangkok and the Suranaree University of Technology for supporting the knowledge to successfully complete this research.

REFERENCES

- [1] THAILAND AUTOMOTIVE INSTITUTE, "Statistics of Thai automotive industry.", [Online]. Available: <https://data.thaiatao.or.th/> [Accessed:18-August-2020]
- [2] Dr.Panadda, "Surface Engineering.", Material and Surface Technology Damage Analysis Laboratory
- [3] Narathip, "Quality Control.", Faculty of Technology, UDON THANI RAJABHAT UNIVERSITY, 2014
- [4] Akekachai Pannawan and Supattarachai Sudawat, "Automated part inspection by image processing system in vehicle part manufacturing", ISSN 1512-7805 Printed in Thailand.
- [5] Kwanchanok, "AI:Artificial Intelligence.", [Online]. Available: <https://www.lib.ku.ac.th/2019/index.php?evid=1971045-artificialintelligence> [Accessed:30-August-2020]
- [6] Nappant Pattanasam and Nattavut Sriwiboon, "Image Processing for Classifying the Quality of the Chok-Anan Mango by Simulating the Human Vision using Deep Learning.", Department of Informatics and Computer, Faculty of Science and Health Technology, Kalasin University, 2020.
- [7] Ashwin Bhandare, Maithili Bhide, Pranav Gokhale, Rohan Chandavarkar, "Application of Convolutional Neural Networks.", Department of Computer Engineering, Pune Institute of Computer Technology Savitribai Phule Pune University, Pune, 2016.
- [8] ODSC, "Overview of the Yolo Object Detection Algorithm.", [Online]. Available: <https://medium.com/@ODSC/overview-of-the-yolo-object-detection-algorithm-7b52a74543e0> [Accessed:28-September-2021]
- [9] DATA, "CNN Yolo Algorithm.", [Online]. Available: <http://datahackers.co/object-detection-yolo-algorithm/> [Accessed:28-September-2021]
- [10] Mark Harlan Goodrich, "Learning Ensembles of First-Order Clauses That Optimize Precision-Recall Curves.", UNIVERSITY OF WISCONSIN-MADISON, 2007.
- [11] Studet, "Confusion Matrix.", [Online]. Available: <https://www.stardat.net/post/confusion-matrix/> [Accessed:30-April-2021]
- [12] 305393 Special Topic in Human Computer Interaction, "How to adjust the Brightness and Contrast of the image by using OpenCV with Visual Studio.", [Online]. Available: <https://305393advancedtopicsinhci.wordpress.com/2017/02/15/how-to-adjust-the-brightness-and-contrast-of-the-image-by-using-open-cv-with-visual-studio/> [Accessed:20-April-2021]
- [13] Sithitouch, "Generating an architectural model using object recognition to observe human movement in public space case study: 3-dimensional parametric model of bus station platforms.", Ref. code: 25605816030075HOB, 2018
- [14] Grace Kamini, "Introduction to Yolo Algorithm for object detection.", [Online]. Available: <https://www.section.in/engineering-education/introduction-to-yolo-algorithm-for-object-detection/> [Accessed:28-September-2021]

3-dimensional parametric model of bus station platforms.", Ref. code: 25605816030075HOB,2018

- [14] Grace Karmini, "Introduction to Yolo Algorithm for object detection." [Online]. Available: <https://www.section.io/engineering-education/introduction-to-yolo-algorithm-for-object-detection/> [Accessed: 26-September-2021]
- [15] L. Shao, A. Mahajan, T. Schreck, and D.J. Lehmann, "Interactive Regression Lens for Exploring Scatter Plots.", Eurographics Conference on Visualization (EuroVis) 2017
- [16] Mr.PI, "Deep Learning EP.3 with Keras with Image Classification." [Online]. Available: <https://medium.com/ai-ml/deep-learning-ep-3-keras-with-image-classification-2d5d95ee5613> [Accessed: 15-August-2021]
- [17] AyushExcel, "Install Library." [Online]. Available: <https://github.com/aianalytics/vokov5> [Accessed: 10-May-2021]



SURADET TANTRAIKARN is a lecturer at the School of Mechanical Engineering Suranaree University of Technology, Thailand. He received his Bachelor's degree and Master's degree in Aerospace Engineering in 2005 and 2007 from the Kasetsart University and a Ph.D. in Automatic Control and System Engineering in 2016 from University of Sheffield, UK. His research interests include Flight Dynamics Modelling, Flight Simulation, Aircraft System Identification or Estimation and Selection, Aircraft State Estimation, Flight Control, Adaptive Control, Fault Tolerance Flight Control, Aircraft Fault Detection and Isolation, Aircraft State Monitoring Intelligent Control, and Autonomous UAS.



TRAISSIRIN SAENGSEANGFA received his Bachelor's degree in Industrial Engineering Technology in 2018 from Kasem Bundit University, Thailand. His final project focused on creating a vision system for identifying a workpiece quality and defects using Artificial Intelligence.



ประวัติผู้เขียน

นายไตรศิริรินทร์ แสงเสียงฟ้า เกิดเมื่อวันอาทิตย์ที่ 28 มิถุนายน พ.ศ. 2535 จังหวัดกรุงเทพมหานคร เริ่มการศึกษา โรงเรียนวัดคฤหบดี จังหวัดกรุงเทพมหานคร โรงเรียนวัดลาดหลุมแก้ว จังหวัดปทุมธานี ตามลำดับ และจบการศึกษาระดับมัธยมศึกษาตอนต้น โรงเรียนวัดจุกเขยและตอนปลาย โรงเรียนพุทธรังษีพิบูล จังหวัดฉะเชิงเทรา จากนั้น ศึกษาในหลักสูตรประกาศนียบัตรวิชาชีพชั้นสูง สาขา อิเล็กทรอนิกส์อุตสาหกรรม วิทยาลัยเทคโนโลยีภาคตะวันออก (อี.เทค) และเข้าศึกษาต่อในระดับปริญญาตรี สาขาเทคโนโลยีวิศวกรรมอุตสาหกรรม คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษมบัณฑิต และสาขาวิชาภาษาญี่ปุ่น ศิลปศาสตรบัณฑิต มหาวิทยาลัยรามคำแหง ปัจจุบันทำงานให้กับบริษัท Mitsubishi Electric Factory Automation Thailand ตำแหน่ง Business Development

ระหว่างการศึกษาในระดับปริญญาโท ได้นำเสนอผลงานทางวิชาการ เรื่อง DESIGN AND DEVELOPMENT OF A VISION SYSTEM FOR IDENTIFYING A WORKPIECE QUALITY AND DETECTS USING OBJECT DETECTION. ในการประชุมนานาชาติ 25TH International Computer Science and Engineering Conference. 18th-20th November, 2021 ณ มหาวิทยาลัยพะเยา

มหาวิทยาลัยเทคโนโลยีสุรนารี