

การตรวจจับหน้าตัดของปลายท่อนซุงจากภาพถ่ายด้านท้ายรถบรรทุกไม้ยูคาลิปตัส



บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ ที่ส่งผ่านทางบัณฑิตวิทยาลัย

The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository (CUIR)
are the thesis authors' files submitted through the University Graduate School.

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2560

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

LOG END CROSS-SECTION DETECTION IN IMAGES TAKEN FROM REAR END OF
EUCALYPTUS TIMBER TRUCKS

Mr. Noppawat Samdangdech



A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in Computer Science
Department of Computer Engineering
Faculty of Engineering
Chulalongkorn University
Academic Year 2017
Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์

การตรวจจับหน้าตัดของปลายท่อขนส่งจากภาพถ่ายด้าน
ท้ายรถบรรทุกไม้อูคาลิปตัส

โดย

นายณพัทธ์ สำแดงเดช

สาขาวิชา

วิทยาศาสตร์คอมพิวเตอร์

อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

ผู้ช่วยศาสตราจารย์ ดร. สีบสกุล พิภพมงคล

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้บัณฑิตวิทยาลัย
หนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

.....คนบดีคณะวิศวกรรมศาสตร์
(รองศาสตราจารย์ ดร. สุพจน์ เตชวรสินสกุล)

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ
(รองศาสตราจารย์ นงลักษณ์ โควาวิสารัช)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ผู้ช่วยศาสตราจารย์ ดร. สีบสกุล พิภพมงคล)

.....กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร. ธนารัตน์ ชลิตาพงศ์)

.....กรรมการภายนอกมหาวิทยาลัย
(ดร. เดชานุกิต กัตัญญทวิฑิพย์)

นพวิษฐ์ สำแดงเดช : การตรวจจับหน้าตัดของปลายท่อนซุงจากภาพถ่ายด้านท้ายรถบรรทุก
ไม้ยูคาลิปตัส (LOG END CROSS-SECTION DETECTION IN IMAGES TAKEN FROM
REAR END OF EUCALYPTUS TIMBER TRUCKS) อ.ที่ปริกษาวิทยานิพนธ์หลัก: ผศ. ดร.
สีบสกุล พิภพมงคล, 104 หน้า.

งานวิจัยนี้เสนอวิธีการตรวจจับหน้าตัดของปลายท่อนซุงที่อยู่ในภาพถ่ายท้ายรถบรรทุกไม้ยูคาลิปตัสโดยใช้การเรียนรู้ด้วยเครื่องและเทคนิคการประมวลผลภาพดิจิทัล โดยใช้ภาพถ่ายจากด้านท้ายรถบรรทุกไม้ยูคาลิปตัส ณ ทางเข้าโรงงานผลิตกระดาษในประเทศไทย โดยแบ่งเป็น 3 ขั้นตอนสำคัญ ขั้นตอนแรกคือการตรวจหารถบรรทุกไม้ยูคาลิปตัส เริ่มจากการตรวจจับวัตถุด้วยโครงสร้างซึ่งเก็ลชอติเทคเตอร์ ถูกนำไปใช้ในการหาพื้นที่ของรถบรรทุกไม้โดยระบุพิกัดของขอบเขตที่ตรวจพบรถบรรทุกไม้และกำจัดวัตถุที่ไม่เกี่ยวข้องในภาพ ขั้นตอนที่สองคือการแยกส่วนหน้าตัดของปลายท่อนซุง ซึ่งถูกนำไปใช้เพื่อสกัดจุดภาพเฉพาะส่วนที่เป็นหน้าตัดของปลายท่อนซุงเท่านั้นและแบ่งส่วนปลายท่อนซุงแต่ละท่อนออกจากภาพพื้นหลัง โดยในขั้นตอนนี้ดำเนินการด้วยวิธีการจำแนกทางความหมายตามโครงข่ายคอนโวลูชันแบบทั่วถึง แต่เนื่องจากภาพอาจมีหน้าตัดของปลายท่อนซุงบางส่วนที่ติดกับหน้าตัดของปลายท่อนซุงอื่น รวมถึงอาจมีส่วนที่ไม่ใช่หน้าตัดของปลายท่อนซุงปรากฏขึ้นในภาพผลลัพธ์ ดังนั้นในขั้นตอนสุดท้ายนี้จึงมีวัตถุประสงค์เพื่อแยกปลายท่อนซุงที่เชื่อมติดกันและเพื่อนับหน้าตัดของปลายท่อนซุงโดยใช้เทคนิคการประมวลผลภาพ วิธีการที่เสนอนี้ได้รับการทดสอบด้วยชุดข้อมูลไม้ยูคาลิปตัสบนรถบรรทุกจำนวน 300 ภาพและมีความถูกต้องเฉลี่ยร้อยละ 94.45 ในการแยกส่วนหน้าตัดของปลายท่อนซุงและมีค่าลบเท็จ (False negative) เฉลี่ยร้อยละ 2.71 และค่าลบจริง (False positive) เฉลี่ยร้อยละ 2.84

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ภาควิชา วิศวกรรมคอมพิวเตอร์

ลายมือชื่อนิสิต

สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์

ลายมือชื่อ อ.ที่ปรึกษาหลัก

ปีการศึกษา 2560

5770204221 : MAJOR COMPUTER SCIENCE

KEYWORDS: LOG DETECTION / LOG COUNTING / EUCALYPTUS LOG / IMAGE PROCESSING
/ IMAGE SEGMENTATION / MACHINE VISION / MACHINE LEARNING

NOPPAWAT SAMDANGDECH: LOG END CROSS-SECTION DETECTION IN IMAGES
TAKEN FROM REAR END OF EUCALYPTUS TIMBER TRUCKS. ADVISOR: ASST.
PROF. SUEBSKUL PHIPHOBMONGKOL, Ph.D., 104 pp.

This research presents a method to detect log-end cut area in images by using machine learning and digital image processing techniques. The proposed method was applied to images taken from rear end of eucalyptus timber trucks at an entrance of a paper mill in Thailand. The method consists of three parts. The first part is the eucalyptus truck detection. Single Shot Detector was used to determine the area of a timber truck to identify the coordinates of timber truck area and eliminate undesired objects in the image. In the second part, log segmentation was applied to segment each log from the background using Fully Convolutional Network for semantic segmentation. Since the segmented images might have some log-end cut areas touching other logs and might contain some non-cut area, image processing techniques were then applied to separate and count each log-end cut area. The proposed method was tested with 300 images of a dataset of eucalyptus timber trucks and achieved an average accuracy of 94.45% in log segmentation with 2.71% of false negative and 2.84% of false positive.

Department: Computer Engineering Student's Signature

Field of Study: Computer Science Advisor's Signature

Academic Year: 2017

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้เสร็จลุล่วงได้ ด้วยความช่วยเหลืออย่างดียิ่งของอาจารย์ที่ปรึกษา ขอขอบพระคุณ ผศ. ดร.สีปสกุล พิภพมงคล ที่ให้ความอนุเคราะห์สละเวลาให้คำแนะนำ ปรึกษา และเสนอแนะข้อคิดเห็นในแนวทางวิจัย และได้ให้ความช่วยเหลือดูแลในด้านต่าง ๆ เป็นอย่างดี

ขอขอบคุณรองศาสตราจารย์นงลักษณ์ โควาวิสารัช ผศ. ดร.ธนรัตน์ ชลิตาพงศ์ และ ดร.เดชานุชิต กตัญญูทวีทิพย์ สำหรับการให้ความอนุเคราะห์มาเป็นประธาน และกรรมการสอบ วิทยานิพนธ์ตามลำดับ อาจารย์ทุกท่านให้คำแนะนำ ข้อคิดเห็น และแนวทางในการพัฒนางานวิจัย นี้ ตลอดจนถึงการตรวจสอบ และแก้ไขข้อบกพร่องของวิทยานิพนธ์ฉบับนี้

ขอขอบคุณบริษัทสยามเซลลูโลส จำกัด และนายณัฐวุฒิ แสนสุข วิศวกร ส่วนประกัน คุณภาพผลิตภัณฑ์ บริษัทผลิตภัณฑ์กระดาษไทย จำกัด สำหรับการอำนวยความสะดวกและ ประสานงานในการขออนุญาตเข้าไปบันทึกภาพถ่ายรถบรรทุกไม้ยูคาลิปตัส ณ โรงงานผลิต กระดาษ เพื่อนำมาใช้ในการทดลองของวิทยานิพนธ์ฉบับนี้

ขอบคุณพี่ ๆ น้อง ๆ และเพื่อน ๆ ในห้องปฏิบัติการ Computer Graphics & Computer Imaging (CGCI) ทุกคนที่ให้คำแนะนำ และคอยช่วยเหลือในการพัฒนางานวิจัย

ท้ายสุดนี้ผู้วิจัยขอขอบพระคุณครอบครัวสำแดงเดช บิดา มารดา และพี่สาว ที่ให้การ อุปการะเลี้ยงดู เป็นกำลังใจในทุก ๆ ด้าน รวมทั้งส่งเสริมให้ศึกษาอย่างต่อเนื่องกับผู้วิจัยเสมอมา

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

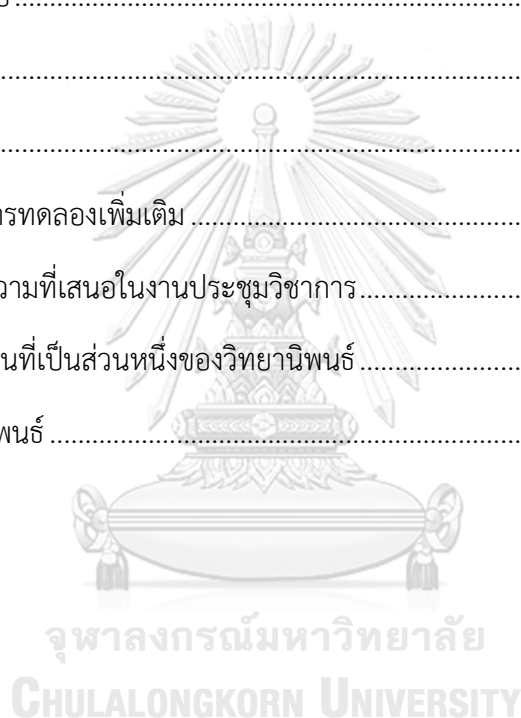
สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญรูป	ฎ
สารบัญตาราง.....	ณ
บทที่ 1 บทนำ	1
1.1 ที่มาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของการวิจัย	4
1.3 ขอบเขตของการวิจัย.....	4
1.4 ประโยชน์ที่ได้รับ.....	4
1.5 วิธีดำเนินการวิจัย.....	5
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง	6
2.1 ลักษณะท่อนซุง.....	6
2.2 แบบจำลองสี (Color model).....	7
2.2.1 แบบจำลองสีอาร์จีบี (RGB color model).....	7
2.2.2 แบบจำลองสีเอชเอสวี (HSV color model).....	8
2.3 ภาพระดับสีเทา (Grayscale image).....	9
2.3.1 การแปลงภาพสีจากแบบจำลองอาร์จีบีเป็นภาพระดับสีเทา (RGB to grayscale image conversion)	9
2.4 ภาพไบนารี (Binary image).....	10
2.4.1 การแปลงภาพระดับเทาเป็นภาพไบนารีด้วยวิธีโอทสุ (Grayscale to binary image conversion by Otsu's thresholding algorithm)	10

2.5	สี่เหลี่ยมในการปิดล้อมวัตถุ (Bounding box)	10
2.6	การติดป้ายส่วนประกอบที่เชื่อมติดกัน (Connected component labeling)	11
2.7	การเปลี่ยนแปลงลักษณะโครงร่างของภาพ (Morphological processing).....	11
2.8	การเรียนรู้ของเครื่อง (Machine learning)	13
2.8.1	การเรียนรู้แบบมีผู้สอน (Supervised learning)	13
2.8.2	การเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning)	13
2.8.3	การเรียนรู้แบบกึ่งมีผู้สอน (Semi-supervised learning)	14
2.8.4	การเรียนรู้แบบเสริมกำลัง (Reinforcement learning).....	14
2.8.5	การถ่ายโยงการเรียนรู้ (Transfer learning).....	15
2.9	การเรียนรู้เชิงลึก (Deep learning)	15
2.9.1	การถดถอยเชิงเส้น (Linear regression).....	17
2.9.2	การถดถอยโลจิสติก (Logistic regression).....	18
2.9.3	ฟังก์ชันการกระตุ้น (Activation function).....	19
2.9.4	ค่าน้ำหนัก (Weight).....	19
2.9.5	ค่าเบี่ยงเบน (Bias).....	20
2.9.6	โครงข่ายประสาทเทียม (Artificial Neural Network : ANN).....	20
2.10	นิรอรลเน็ตเวิร์กคอนโวลูชัน (Convolutional Neural Network : CNN or ConvNet)...	22
2.11	การรู้จำภาพ (Image Recognition).....	26
2.11.1	การจำแนกประเภทข้อมูลภาพ และการจัดกลุ่มภาพ (Image classification and image clustering).....	27
2.11.2	การหาตำแหน่ง (Localization).....	27
2.11.3	การตรวจจับวัตถุ (Object Detection).....	27
2.11.4	การจำแนกทางความหมาย (Semantic Segmentation).....	30

2.12 การเพิ่มคุณค่าของข้อมูล (Data augmentation)	34
2.13 การวัดประสิทธิภาพของการจำแนกประเภท (Performance Evaluation of Classification)	35
2.14 งานวิจัยที่เกี่ยวข้อง	36
2.14.1 ประเภทที่ใช้การประมวลผลจากแสงเลเซอร์.....	36
2.14.2 ประเภทที่ใช้ภาพถ่าย	37
บทที่ 3 แนวคิดและวิธีดำเนินงาน.....	43
3.1 การตรวจจับรถบรรทุกไม้ (Timber truck detection)	46
3.2 การแยกส่วนหน้าตัดของปลายท่อนซุง (Log segmentation)	49
3.3 การนับปลายท่อนซุง (Log counting).....	50
3.1.1 การสร้างภาพไบนารี.....	50
3.1.2 การเปลี่ยนแปลงลักษณะโครงร่างของภาพ	51
3.1.3 การตัดป้ายส่วนประกอบที่เชื่อมติดกัน.....	51
บทที่ 4 การทดลองและผลของการทดลอง.....	53
4.1 ภาพที่ใช้ในการทดลอง.....	53
4.2 สภาพแวดล้อมด้านซอฟต์แวร์และฮาร์ดแวร์.....	53
4.3 การตรวจจับรถบรรทุกไม้ยุคาลิปต์ด้วยโครงสร้าง SSD.....	54
4.3.1 การเตรียมข้อมูล	55
4.3.2 การสอน	60
4.3.3 การทดสอบ	62
4.4 การแยกส่วนหน้าตัดของปลายท่อนซุงด้วยโครงข่าย VGG16	65
4.4.1 การเตรียมข้อมูล	65
4.4.2 การสอน	69

4.4.3 การทดสอบ	71
4.5 การหาค่าแกนกลางที่เหมาะสมสำหรับการเปลี่ยนแปลงลักษณะโครงสร้างของภาพ	73
4.6 ผลการทดลอง	75
4.7 การวิเคราะห์ผลการทดลอง	76
บทที่ 5 สรุปการวิจัยและข้อเสนอแนะ	78
5.1 สรุปผลการวิจัย	78
5.2 ข้อเสนอแนะ	79
รายการอ้างอิง	80
ภาคผนวก ก ผลการทดลองเพิ่มเติม	85
ภาคผนวก ข บทความที่เสนอในงานประชุมวิชาการ	96
ภาคผนวก ค ผลงานที่เป็นส่วนหนึ่งของวิทยานิพนธ์	103
ประวัติผู้เขียนวิทยานิพนธ์	104



สารบัญรูป

รูปที่ 1-1 ตัวอย่างภาพถ่ายรถบรรทุกไม้ยูคาลิปตัส (ก) หน้าตัดของปลายท่อนซุงด้านท้ายรถบรรทุก (ข) แม่แบบกรอบสี่เหลี่ยมมาตรฐานสำหรับใช้ประมาณขนาดเส้นผ่านศูนย์กลางหน้าตัดของปลายไม้	2
รูปที่ 1-2 เปรียบเทียบระหว่างหน้าตัดของปลายท่อนซุงไม้สนและไม้ยูคาลิปตัส (ก) หน้าตัดของปลายท่อนซุงไม้สนจากงานวิจัยของ Rahman และคณะ [2] (ข) หน้าตัดของปลายท่อนซุงไม้ยูคาลิปตัสจากโรงงานผลิตกระดาษ.....	3
รูปที่ 2-1 โครงสร้างหลักของท่อนซุงไม้ยูคาลิปตัส	6
รูปที่ 2-2 หน้าตัดของปลายท่อนซุง (ก) หน้าตัดของปลายท่อนซุงไม้สน (ข) หน้าตัดของปลายท่อนซุงไม้ยูคาลิปตัส	7
รูปที่ 2-3 แบบจำลองสี RGB บนแกนสามมิติ [7]	8
รูปที่ 2-4 กรวยแสดงแบบจำลองสี HSV บนแกนสามมิติ [7]	8
รูปที่ 2-5 ตัวอย่างภาพระดับสีเทาในขนาด 8 บิต	9
รูปที่ 2-6 การแปลงภาพสีจากอาร์จีบีเป็นภาพไบนารี (ก) ภาพต้นฉบับ (ข) ภาพระดับสีเทา (ค) ภาพไบนารี [9]	10
รูปที่ 2-7 การตีกรอบสี่เหลี่ยมปิดล้อมรูปทรงเรขาคณิตในภาพ (ก) ภาพต้นฉบับ (ข) ภาพผลลัพธ์ที่ตีกรอบปิดล้อมวัตถุ.....	11
รูปที่ 2-8 การตัดป้ายส่วนประกอบที่เชื่อมติดกัน [10]	11
รูปที่ 2-9 การเปลี่ยนแปลงลักษณะโครงร่างของภาพทั้ง 4 ชนิด [11].....	12
รูปที่ 2-10 ตัวอย่างของปัญหาในชีวิตจริงของผลลัพธ์จากการเรียนรู้แบบมีผู้สอนและไม่มีผู้สอน (ก) การจำแนกประเภทผู้ป่วยว่ามีโรคหรือไม่ (ข) การทำนายราคาบ้าน (ค) การจัดกลุ่มพฤติกรรม การซื้อของลูกค้า	14
รูปที่ 2-11 การประยุกต์ความรู้ที่เรียนรู้ก่อนหน้าไปยังแบบจำลองใหม่	15
รูปที่ 2-12 ตัวอย่างโครงสร้างเลเยอร์ของแบบจำลองการเรียนรู้เชิงลึก.....	16

รูปที่ 2-13 เปรียบเทียบความแตกต่างในการสกัดคุณลักษณะของข้อมูล ระหว่างการเรียนรู้ของเครื่องและการเรียนรู้เชิงลึก	17
รูปที่ 2-14 ตัวอย่างแสดงการทำงานของการทำงานของกรดถอยเชิงเส้น ซึ่งประมาณการราคาบ้านที่ดีที่สุด	18
รูปที่ 2-15 ตัวอย่างแสดงการทำงานของการทำงานของกรดถอยโลจิสติก ซึ่งคำนวณหานักเรียนที่ผ่านเกณฑ์คะแนนและผลการทดสอบ	18
รูปที่ 2-16 ภาพตัวอย่างการทำงานของโครงข่ายประสาทเทียม.....	20
รูปที่ 2-17 ประเภทของโครงข่ายประสาทเทียม (ก) แบบป้อนไปข้างหน้า (ข) แบบย้อนกลับ.....	22
รูปที่ 2-18 ตัวอย่างโครงสร้างของนิวรอลเน็ตเวิร์กคอนโวลูชัน โดยตัวกรองถูกนำไปใช้ในการสอนที่ความละเอียดต่างกันและผลลัพธ์ของแต่ละภาพจากชั้นคอนโวลูชันจะถูกนำไปใช้เป็นข้อมูลนำเข้าของชั้นถัดไป [12].....	23
รูปที่ 2-19 กราฟแสดงสมการฟังก์ชันเรกติไฟด์เชิงเส้น	25
รูปที่ 2-20 ตัวอย่างการทำแอฟเวอเรจพูลลิ่งและแมกซ์พูลลิ่ง [13].....	26
รูปที่ 2-21 สถาปัตยกรรมของโครงสร้าง SSD โดยมีภาพนำเข้าขนาด 300 x 300 จุดภาพ	29
รูปที่ 2-22 ชั้นของคอนโวลูชันแบบ Depthwise separable [17].....	30
รูปที่ 2-23 ตัวอย่างการจำแนกทางความหมายของภาพ (ก) ภาพนำเข้า (ข) ภาพผลลัพธ์ของการจำแนกทางความหมาย [19].....	31
รูปที่ 2-24 โครงสร้างทำนายความหนาแน่นของโครงข่ายคอนโวลูชันแบบทั่วถึงแบบปลายทางถึงปลายทาง [18]	32
รูปที่ 2-25 สถาปัตยกรรมของโครงข่ายแบบ VGG [23].....	34
รูปที่ 2-26 ตัวอย่างการเพิ่มคุณค่าของข้อมูลด้วยวิธีการแปลงข้อมูลภาพ	35
รูปที่ 2-27 ตัวอย่างภาพถ่ายของรูปแบบต่าง ๆ ของรถบรรทุกไม้ยูคาลิปตัสในประเทศไทย.....	42
รูปที่ 3-1 ตัวอย่างภาพถ่ายด้านท้ายรถบรรทุกจากโรงงานผลิตกระดาษ ณ บริษัทสยามเซลลูโลส จำกัด.....	43
รูปที่ 3-2 ตัวอย่างภาพถ่ายด้านข้างรถกึ่งพ่วงไม้ยูคาลิปตัส	44
รูปที่ 3-3 แผนภาพแสดงภาพรวมของระบบการตรวจจับหน้าตัดของปลายท่อนซุง	45

รูปที่ 3-4 แผนภาพแสดงกระบวนการตรวจจับรถบรรทุกไม้ยูคาลิปตัส	46
รูปที่ 3-5 ภาพตัวอย่างผลลัพธ์ของขั้นตอนการตรวจจับรถบรรทุกไม้ (ก) ภาพนำเข้าปกติ (ข) ภาพนำเข้าที่มีรถบรรทุกไม้มากกว่า 1 คัน (ค) ผลลัพธ์ของการตรวจจับรถบรรทุกไม้ในภาพนำเข้าปกติ (ง) ผลลัพธ์ของการตรวจจับรถบรรทุกไม้ในภาพนำเข้าที่มีรถบรรทุกไม้มากกว่า 1 คัน (จ) และ (ฉ) ผลลัพธ์ภาพที่ถูกตัดภาพส่วนเฉพาะด้านท้ายรถบรรทุกไม้ยูคาลิปตัส	48
รูปที่ 3-6 ภาพผลลัพธ์ของการแยกส่วนหน้าตัดของปลายท่อนซุงด้วยโครงข่าย VGG16 (ก) ภาพนำเข้าปกติ และ (ข) ภาพนำเข้าที่มีรถบรรทุกไม้มากกว่า 1 คัน.....	49
รูปที่ 3-7 แผนภาพแสดงกระบวนการสร้างภาพไบนารี.....	50
รูปที่ 3-8 แผนภาพแสดงกระบวนการตัดป้ายส่วนประกอบที่เชื่อมติดกัน	51
รูปที่ 3-9 ภาพผลลัพธ์ของระบบการตรวจจับหน้าตัดของขั้นตอนการนับปลายท่อนซุง โดยที่ (ก) และ (ข) เป็นผลลัพธ์ของขั้นตอนการเปลี่ยนแปลงลักษณะโครงสร้างของภาพที่ประยุกต์บนภาพไบนารี (ค) และ (ง) เป็นผลลัพธ์ของการนับปลายท่อนซุง (จ) และ (ฉ) เป็นการซ้อนทับระหว่างภาพนำเข้าและภาพผลลัพธ์	52
รูปที่ 4-1 แผนภาพแสดงการทดลองการตรวจจับไม้ยูคาลิปตัสด้วยแบบจำลอง SSD MobileNet v1 COCO	55
รูปที่ 4-2 ชุดข้อมูลภาพนำเขารถบรรทุกไม้ยูคาลิปตัส	56
รูปที่ 4-3 การเตรียมข้อมูลด้วยการบรรยายภาพ (ก) เครื่องมือที่ใช้ในการบรรยายภาพในรูปแบบกรอบสี่เหลี่ยมในการปิดล้อมวัตถุ (ข) ไฟล์โครงสร้าง XML ของภาพที่สอดคล้องกันในรูปที่ 4-3 (ก).....	57
รูปที่ 4-4 โครงสร้างของไฟล์ Label map สำหรับการตรวจจับรถบรรทุกไม้ยูคาลิปตัส.....	58
รูปที่ 4-5 ตัวอย่างการแปลงไฟล์รูปแบบ XML เป็นไฟล์ CSV เพื่อนำไปสร้างไฟล์ TFRecord.....	58
รูปที่ 4-6 แผนภาพแสดงการบรรยายภาพและการแบ่งชุดข้อมูลสำหรับการสอน	59
รูปที่ 4-7 การเริ่มงานด้านการสอนและงานด้านการประเมินแบบจำลองซึ่งถูกประมวลผลในเวลาเดียวกัน (ก) งานด้านการสอน (ข) งานด้านการประเมิน.....	61

รูปที่ 4-8 ความแม่นยำและการสูญเสียทั้งหมดของแบบจำลอง (ก) กราฟแสดงความแม่นยำ (ข) กราฟแสดงการสูญเสียทั้งหมด (ค) การประเมินผลของภาพตัวอย่างในขณะที่ยังดำเนินการสอนแบบจำลอง.....	62
รูปที่ 4-9 ภาพผลลัพธ์ของวิธีการตรวจจําบรรทุกไม้ยูคาลิปตัสด้วยโครงสร้าง SSD (ก) ภาพทดสอบ (ข) ภาพผลลัพธ์จากการทดสอบ (ค) ภาพผลลัพธ์จากการตัดภาพส่วนเฉพาะด้านท้ายบรรทุกไม้ในภาพโดยอ้างอิงจากพิกัดของกรอบสี่เหลี่ยมในการปิดล้อมวัตถุ.....	64
รูปที่ 4-10 แผนภาพแสดงการทดลองการแยกส่วนหน้าตัดของปลายท่อนซุงด้วยแบบจำลอง VGG16.....	66
รูปที่ 4-11 การบรรยายภาพด้วยการติดป้ายชื่อกำกับข้อมูลประเภท ได้แก่ หน้าตัดของปลายท่อนซุงและที่ไม่ใช่หน้าตัดของปลายท่อนซุง (ก) ภาพนำเข้าในชุดข้อมูลสอน (ข) ภาพที่ได้รับการติดป้ายกำกับข้อมูลประเภท	67
รูปที่ 4-12 ตัวอย่างการเพิ่มคุณค่าของข้อมูลในชุดข้อมูลบรรทุกไม้ยูคาลิปตัส	68
รูปที่ 4-13 การเริ่มงานด้านการสอนและงานด้านการตรวจสอบแบบจำลองซึ่งถูกประมวลผลในเวลาเดียวกัน	70
รูปที่ 4-14 การสูญเสียทั้งหมดของแบบจำลอง (ก) กราฟแสดงการสูญเสียของการสอน (ข) กราฟแสดงการสูญเสียของการตรวจสอบ (ค) การประเมินผลของภาพตัวอย่างของการสอนและการตรวจสอบในขณะที่ยังดำเนินการสอนแบบจำลอง	71
รูปที่ 4-15 ภาพผลลัพธ์ของวิธีการแยกส่วนหน้าตัดของปลายท่อนซุงด้วยแบบจำลอง VGG16 (ก) ภาพทดสอบ (ข) ภาพผลลัพธ์ของกระบวนการแยกส่วนหน้าตัดของปลายท่อนซุง (ค) ภาพผลลัพธ์จากการแยกส่วนหน้าตัดของปลายท่อนซุงวางทับบนภาพทดสอบ	72
รูปที่ 4-16 ตัวอย่างกรณีของหน้าตัดของปลายท่อนซุงที่เชื่อมติดกัน (ก) ภาพทดสอบ (ข) ภาพผลลัพธ์การแยกส่วนหน้าตัดของปลายท่อนซุง (ค) ภาพผลลัพธ์ของกระบวนการโอเพนนิ่งในการเปลี่ยนแปลงลักษณะโครงร่างของภาพ	74
รูปที่ 4-17 การตรวจสอบวิธีที่เสนอเปรียบเทียบกับมาตรฐาน	75
รูปที่ 4-18 ตัวอย่างของภาพนำเข้าและภาพการแยกส่วนหน้าตัดของปลายท่อนซุงที่พบปัญหา (ก) และ (ข) ปลายท่อนซุงที่เชื่อมติดกัน (ค) และ (ง) การซ้อนทับของปลายท่อนซุง (จ) และ (ฉ)	

การแตกของหน้าตัดของปลายท่อนซุง (ช) และ (ซ) ทิศทางของปลายท่อนซุง (ณ) และ (ญ) วัตถุที่
ไม่พึงประสงค์ปกคลุมบนหน้าตัดของปลายท่อนซุง 77

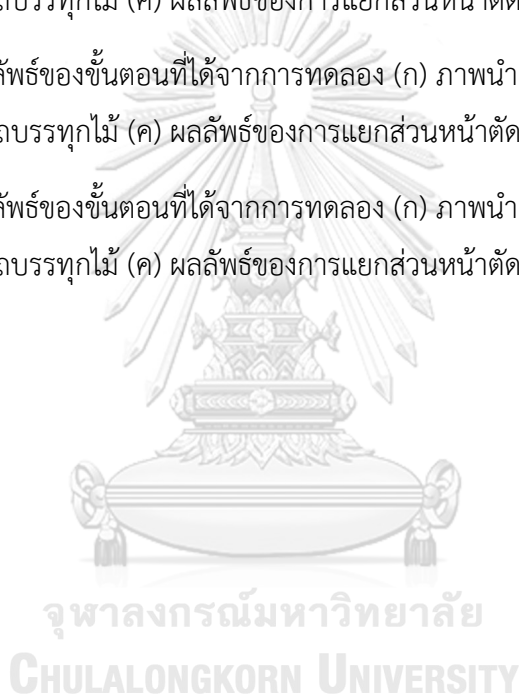
รูปที่ ก-1 ตัวอย่างผลลัพธ์ของขั้นตอนที่ได้จากการทดลอง (ก) ภาพนำเข้า (ข) ผลลัพธ์ของ
ขั้นตอนการตรวจจําบรทุกไม้ (ค) ผลลัพธ์ของการแยกส่วนหน้าตัดของปลายท่อนซุง 86

รูปที่ ก-2 ตัวอย่างผลลัพธ์ของขั้นตอนที่ได้จากการทดลอง (ก) ภาพนำเข้า (ข) ผลลัพธ์ของ
ขั้นตอนการตรวจจําบรทุกไม้ (ค) ผลลัพธ์ของการแยกส่วนหน้าตัดของปลายท่อนซุง 88

รูปที่ ก-3 ตัวอย่างผลลัพธ์ของขั้นตอนที่ได้จากการทดลอง (ก) ภาพนำเข้า (ข) ผลลัพธ์ของ
ขั้นตอนการตรวจจําบรทุกไม้ (ค) ผลลัพธ์ของการแยกส่วนหน้าตัดของปลายท่อนซุง 90

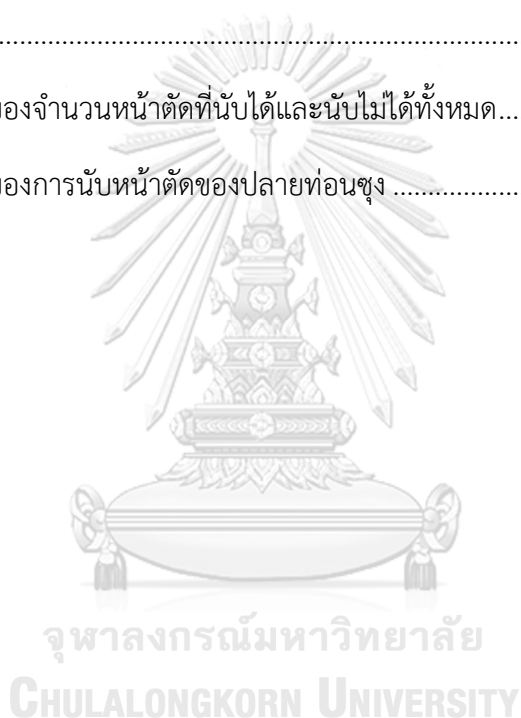
รูปที่ ก-4 ตัวอย่างผลลัพธ์ของขั้นตอนที่ได้จากการทดลอง (ก) ภาพนำเข้า (ข) ผลลัพธ์ของ
ขั้นตอนการตรวจจําบรทุกไม้ (ค) ผลลัพธ์ของการแยกส่วนหน้าตัดของปลายท่อนซุง 92

รูปที่ ก-5 ตัวอย่างผลลัพธ์ของขั้นตอนที่ได้จากการทดลอง (ก) ภาพนำเข้า (ข) ผลลัพธ์ของ
ขั้นตอนการตรวจจําบรทุกไม้ (ค) ผลลัพธ์ของการแยกส่วนหน้าตัดของปลายท่อนซุง 94



สารบัญตาราง

ตารางที่ 4-1 การตั้งค่าตัวแปรที่เปลี่ยนแปลงสำหรับการสอนแบบจำลอง	60
ตารางที่ 4-2 ผลลัพธ์ Precision Recall และ F-Measure จากการตรวจจับพื้นที่ที่สนใจของ รถบรรทุกไม้อยู่คาลิปดัส.....	63
ตารางที่ 4-3 การตั้งค่าตัวแปรที่เปลี่ยนแปลงสำหรับการสอนแบบจำลอง	68
ตารางที่ 4-4 ผลลัพธ์ Precision Recall และ F-Measure จากการแยกส่วนพื้นที่ที่สนใจหน้าตัด ของปลายท่อนซุง	72
ตารางที่ 4-5 ผลลัพธ์ของจำนวนหน้าตัดที่นับได้และนับไม่ได้ทั้งหมด.....	75
ตารางที่ 4-6 ผลลัพธ์ของการนับหน้าตัดของปลายท่อนซุง	76



บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของปัญหา

ในการนำไม้ไปใช้ด้านอุตสาหกรรมการผลิตกระดาษ ไม้ยูคาลิปตัสเป็นไม้ที่เหมาะสมในการนำไปแปรรูปเพื่อผลิตเยื่อกระดาษเนื่องจากเยื่อไม้ยูคาลิปตัสมีความฟูสูงและมีความทึบแสง รวมถึงไฟเบอร์มีความแข็งแรงและเป็นไม้ที่มีอัตราการเจริญเติบโตสูง สามารถตัดเพื่อนำไปใช้ได้ภายในระยะเวลา 4-5 ปี ด้วยเหตุผลนี้ไม้ยูคาลิปตัสจึงเหมาะกับการนำไปใช้ในอุตสาหกรรมการผลิตเยื่อกระดาษมากกว่าไม้พันธุ์อื่น ๆ ข้อมูลในการนำเยื่อไม้ไปผลิตกระดาษจากกรรมป่าไม้ กระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม ไม้ท่อนยูคาลิปตัส 4.5 ต้นสามารถนำไปผลิตเยื่อไม้เพื่อนำไปผลิตกระดาษได้ 1 ตัน [1] ดังนั้นในปัจจุบันจึงมีตลาดรองรับไม้ยูคาลิปตัสเป็นจำนวนมาก กระบวนการในการขนส่งไม้ยูคาลิปตัส ผู้ขายไม้จำเป็นต้องนำท่อนซุงไม้ยูคาลิปตัสบรรทุกทุกใส่บริเวณด้านท้ายรถบรรทุกเพื่อขนส่งไปยังโรงงานผลิตกระดาษ เมื่อรถบรรทุกเคลื่อนที่ถึง ผู้ตรวจสอบจะเริ่มกระบวนการตรวจรับสินค้าโดยทำการตรวจวัดท่อนซุงด้วยวิธีการประมาณขนาดเส้นผ่านศูนย์กลางของไม้ที่รับซื้อให้ใกล้เคียงขนาดตามจริง เนื่องจากจะต้องนำข้อมูลการวัดขนาดท่อนซุงไปใช้กำหนดราคา และกำหนดขนาดของไม้ในการขนส่งแต่ละครั้งได้

ปัจจุบันวิธีการวัดในประเทศไทยยังคงใช้วิธีการสังเกตด้วยสายตามนุษย์ โดยการนำแม่แบบกรอบสี่เหลี่ยมมาตรฐาน (Reference diameter template) นำไปครอบกับปลายไม้จำนวนหนึ่งบริเวณหลังรถบรรทุก หลังจากนั้นผู้ตรวจสอบเริ่มดำเนินการประมาณขนาดเส้นผ่านศูนย์กลางปลายไม้ด้วยแม่แบบมาตรฐานในพื้นที่ที่กำหนดดังตัวอย่างรูปที่ 1-1 แล้วนำไปเปรียบเทียบกับพื้นที่ทั้งหมดว่ามีขนาดเฉลี่ยของเส้นผ่านศูนย์กลางผ่านเกณฑ์มาตรฐานที่กำหนดไว้ เนื่องจากการนับด้วยสายตามนุษย์อาจทำให้ใช้ระยะเวลาในการนับนาน เกิดความคลาดเคลื่อนได้มาก และส่งผลต่อกระบวนการกำหนดราคา ซึ่งผู้ขายและผู้ซื้ออาจได้รับราคาที่สูงหรือต่ำกว่าความเป็นจริง ด้วยเหตุผลนี้เองจึงมีการนำเอาภาพถ่ายท่อนซุงจากด้านท้ายรถบรรทุกมาประยุกต์ใช้เทคโนโลยีการประมวลผลจากภาพถ่าย (Image processing) แทนระบบเดิม ซึ่งใช้การสังเกตจากสายตามนุษย์ เพื่อช่วยในการวัดให้ดียิ่งขึ้นกว่าเดิมและได้ความผิดพลาดน้อยลง



(ก)



(ข)

รูปที่ 1-1 ตัวอย่างภาพถ่ายรถบรรทุกไม้ยูคาลิปตัส (ก) หน้าตัดของปลายท่อนซุงด้านทำยรถบรรทุก (ข) แม่แบบกรอบสี่เหลี่ยมมาตรฐานสำหรับใช้ประมาณขนาดเส้นผ่านศูนย์กลางหน้าตัดของปลายไม้

งานวิจัยที่เกี่ยวข้องกับการตรวจจับและวัดขนาดปลายท่อนไม้ด้วยภาพถ่าย มีการนำเทคนิคการประมวลผลจากภาพถ่ายมาประยุกต์ใช้ในการนับจำนวนปลายท่อนซุงไม้สน งานวิจัยของ Rahman และคณะ [2] ได้นำภาพถ่ายท่อนซุงจากหลังรถบรรทุกมาประมวลผล มีการหาขอบภาพ การใช้ค่าสีในกระบวนการ เช่นเดียวกับงานวิจัยของ Shvarts [3] มีการใช้ค่าสี การเปลี่ยนแปลงลักษณะโครงสร้างของภาพ เป็นการทดลองใช้กับไม้สนที่มีสีหน้าตัดสม่ำเสมอ หากเปลี่ยนเป็นไม้ชนิดอื่นที่มีลวดลาย หรือสีต่างกันอาจต้องมีการปรับเปลี่ยนกระบวนการ หรือไม่สามารถทำงานได้ตามความต้องการ งานวิจัยของ Knyaz และ Maksimov [4] ได้ใช้เทคนิคหาคุณสมบัติของระบบกล้องการแยกลักษณะของภาพ เพื่อคำนวณหาปริมาณท่อนซุงจากภาพถ่าย แต่สามารถใช้ได้กับภาพถ่ายที่มีพื้นผิวหน้าตัดที่ไม่ถูกบังและไม้ทุกท่อนจะมีปลายท่อนยื่นออกมาใกล้เคียงกัน

อย่างไรก็ตามงานวิจัยของ Rahman และคณะ [2] และงานวิจัยของ Shvarts [3] ทดลองการตรวจจับหน้าตัดของปลายท่อนซุงได้เฉพาะที่มีค่าสีใกล้เคียงกัน และไม่สามารถนับจำนวนปลายท่อนซุงที่อยู่ด้านใน ซึ่งถูกสภาพแสงและเงาบดบัง ปัญหาที่สำคัญอย่างหนึ่งในการตรวจจับหน้าตัดของปลายท่อนซุง คือ ท่อนซุงที่อยู่ลึกเข้าไปอาจได้รับแสงหรือเงาน้อย หรือค่าสีของไม้เปลี่ยนจากแดดฝนทำให้กระบวนการดังกล่าวไม่สามารถตรวจจับหน้าตัดของปลายท่อนซุงได้



(ก)



(ข)

รูปที่ 1-2 เปรียบเทียบระหว่างหน้าตัดของปลายท่อนซุงไม้สนและไม้ยูคาลิปตัส (ก) หน้าตัดของปลายท่อนซุงไม้สนจากงานวิจัยของ Rahman และคณะ [2] (ข) หน้าตัดของปลายท่อนซุงไม้ยูคาลิปตัสจาก

โรงงานผลิตกระดาษ

สำหรับในงานวิจัยที่เสนอนี้มุ่งเน้นไปที่การตรวจจับหน้าตัดของปลายท่อนซุงไม้ยูคาลิปตัสซึ่งเป็นไม้เศรษฐกิจของประเทศไทย โดยปรับปรุงให้ความแม่นยำขึ้นเมื่อเทียบกับวิธีการคาดคะเนด้วยสายตามนุษย์ และตรวจนับไม้ได้ใกล้เคียงความจริงมากขึ้น เช่น ในบริเวณหน้าตัดของปลายไม้ที่อยู่ลึกเข้าไปจากแนวระนาบไม้ปกติจากด้านท้ายรถบรรทุก หากสามารถทำให้กระบวนการขั้นตอนการตรวจสอบปลายท่อนซุงถูกต้องกว่าปัจจุบัน จะช่วยนำไปสู่การคัดเลือกขนาดให้สอดคล้องกับความต้องการของผู้รับซื้อต่อไป

จากเหตุผลข้างต้น ผู้วิจัยพบว่าการตรวจจับหน้าตัดของปลายท่อนซุงไม้ยูคาลิปตัสให้ได้ความผิดพลาดน้อยลงนั้นเป็นปัญหาที่สำคัญ เนื่องจากปลายท่อนซุงไม้ยูคาลิปตัสมีลักษณะที่แตกต่างจากไม้สน บริเวณหน้าตัดมีสีที่ไม่สม่ำเสมอ ความชื้นของไม้ ซึ่งส่งผลต่อสีบริเวณหน้าตัดของปลายท่อนซุง รูปทรงของหน้าตัด และความยาวปลายท่อนไม้ใกล้เคียง ทำให้มีความตื่นตระหนกไม่เท่ากัน รวมถึงลักษณะของภาพถ่ายที่มีสิ่งรบกวนอยู่หลายปัจจัย เช่น สภาพแสงเงา ความละเอียดและมุมมองของภาพถ่าย หรือมีวัตถุอื่นบดบังปลายท่อนซุง ดังตัวอย่างรูปที่ 1-1 และรูปที่ 1-2 ด้วยปัจจัยเหล่านี้ทำให้การตรวจจับหน้าตัดของปลายท่อนซุงไม้ยูคาลิปตัสให้ได้ความผิดพลาดน้อยลงนั้น จึงมีความท้าทายเป็นอย่างยิ่ง และในข้อเสนอวิทยานิพนธ์ฉบับนี้ ผู้วิจัยต้องการพัฒนาวิธีการตรวจจับหน้าตัดของปลายท่อนซุงไม้ยูคาลิปตัสบนรถบรรทุกโดยใช้เทคนิคการประมวลผลภาพถ่าย และการเรียนรู้ของเครื่อง

(Machine Learning) ให้มีความแม่นยำมากขึ้นเมื่อเปรียบเทียบกับวิธีการคาดคะเนด้วยสายตามนุษย์ โดยมุ่งหวังว่าในอนาคตสามารถนำไปใช้ในโรงงานผลิตกระดาษได้จริง

1.2 วัตถุประสงค์ของการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อตรวจจับหน้าตัดของปลายท่อนซุงจากภาพถ่ายด้านท้ายรถบรรทุกไม้ยูคาลิปตัส

1.3 ขอบเขตของการวิจัย

การพัฒนาวิธีการตรวจจับหน้าตัดของปลายท่อนซุงจากภาพถ่ายด้านท้ายรถบรรทุกไม้ยูคาลิปตัส มีขอบเขตดังนี้

1.3.1 ทดสอบการตรวจจับหน้าตัดของปลายท่อนซุงจากภาพถ่ายด้านท้ายรถบรรทุกไม้ยูคาลิปตัสเท่านั้น

1.3.2 ภาพที่นำมาใช้เป็นภาพสี ถูกถ่ายในเวลาประมาณระหว่าง 8:00 – 17:00 น. ภายนอกอาคาร และเป็นภาพนิ่งเท่านั้น

1.3.3 ผลของการตรวจจับหน้าตัดของปลายท่อนซุงเปรียบเทียบกับการมองด้วยตามนุษย์

1.4 ประโยชน์ที่ได้รับ

1.4.1 สามารถนำการวิจัยไปใช้พัฒนา การตรวจจับหน้าตัดของปลายท่อนซุงจากภาพถ่ายด้านท้ายรถบรรทุกไม้ยูคาลิปตัส แทนระบบเก่าที่ใช้การคาดคะเนจากสายตามนุษย์

1.4.2 ลดความคลาดเคลื่อนเมื่อเทียบกับการคาดคะเนด้วยสายตามนุษย์ และช่วยประหยัดระยะเวลาในการตรวจจับหน้าตัดของปลายท่อนซุงในแต่ละครั้ง

1.4.3 ประยุกต์เพื่อใช้ตรวจจับไม้ชนิดอื่น และต่อยอดเพื่อใช้งานในสถานการณ์จริงต่อไป

1.5 วิธีดำเนินการวิจัย

งานวิจัยนี้ต้องการศึกษาค้นคว้าเพื่อหากระบวนการตรวจจับหน้าตัดของปลายท่อนของต้นหรือกิ่งยูคาลิปตัสจากภาพถ่ายหลังรถบรรทุกด้วยวิธีการประมวลผลภาพ วิธีการดำเนินงานวิจัยมีดังนี้

1.5.1 ศึกษาความรู้และทฤษฎีที่เกี่ยวข้อง

- 1) ดูสภาพการทำงานจริงของโรงงานที่เกี่ยวข้องอย่างน้อย 1 โรงงาน
- 2) ศึกษางานวิจัยที่เกี่ยวข้อง
- 3) ศึกษาความรู้เกี่ยวกับเทคนิคการประมวลผลจากภาพถ่าย
- 4) ศึกษาความรู้เกี่ยวกับการจำแนกวัตถุในภาพถ่าย
- 5) ศึกษาความรู้เกี่ยวกับการประเมินมูลค่าท่อนซุง

1.5.2 ออกแบบแนวคิดเพื่อแก้ปัญหา

- 1) ออกแบบแนวคิดในภาพรวม
- 2) ออกแบบแนวคิดในรายละเอียดส่วนต่าง ๆ
- 3) ทดสอบตามแนวคิดงานวิจัยที่เกี่ยวข้อง
- 4) วิเคราะห์ผลการทดสอบตามแนวคิดงานวิจัยที่เกี่ยวข้อง
- 5) ออกแบบแนวคิดโดยนำเอาแนวคิดงานวิจัยที่เกี่ยวข้องมาประยุกต์ใช้

1.5.3 ออกแบบวิธีการตรวจสอบแนวคิดที่นำเสนอ

1.5.4 ศึกษาและพัฒนาเครื่องมือเพื่อใช้ในการตรวจสอบแนวคิด

1.5.5 ตรวจสอบความถูกต้องของเครื่องมือและแนวคิดที่นำเสนอ

1.5.6 สรุปผลการทดลอง ตีพิมพ์งานวิจัย และจัดเล่มวิทยานิพนธ์

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ลักษณะท่อนซุง

ท่อนซุงเป็นส่วนประกอบหลักของต้นไม้ ซึ่งสามารถนำไปใช้ประโยชน์ในอุตสาหกรรม การแปรรูปไม้ และบ่งบอกคุณลักษณะของไม้ได้ เช่น ชนิด อายุ ความสูงต้นไม้โดยประมาณ เส้นรอบวง หรือเส้นผ่านศูนย์กลาง เป็นต้น [5] ท่อนซุงของไม้แต่ละชนิดจะมีลักษณะที่แตกต่างกันทั้งภายนอก และภายใน ถึงแม้ว่าจะเป็นไม้ชนิดเดียวกัน แต่ก็อาจมีลักษณะที่ไม่เหมือนกัน ขึ้นอยู่กับความชื้น ขนาด และสิ่งแวดล้อมในสภาวะต่าง ๆ สำหรับโครงสร้างท่อนซุงไม้ยูคาลิปตัสประกอบด้วยเปลือกไม้ และเนื้อไม้ ซึ่งในส่วนของเนื้อไม้แบ่งออกเป็น 2 ส่วน คือ กระทบและแก่น ดังที่แสดงในรูปที่ 2-1

- 1) กระทบ คือ ส่วนเนื้อไม้ด้านนอกซึ่งอยู่ติดกับเปลือกไม้ มีหน้าที่ในการลำเลียงน้ำ
- 2) แก่น คือ เนื้อไม้ด้านในสุดซึ่งถูกห่อหุ้มด้วยกระทบ



รูปที่ 2-1 โครงสร้างหลักของท่อนซุงไม้ยูคาลิปตัส

สำหรับงานวิจัยนี้สนใจในการตรวจจับหน้าตัดของปลายท่อนซุงไม้ยูคาลิปตัส เมื่อนำมาเปรียบเทียบกับปลายท่อนซุงไม้สน [2] จะพบว่าปลายท่อนซุงไม้ยูคาลิปตัสมีสีหน้าตัดที่ปรากฏ ลวดลายและวงปีปรากฏเด่นชัดกว่าไม้สนแสดงดังรูปที่ 2-2



(ก)

(ข)

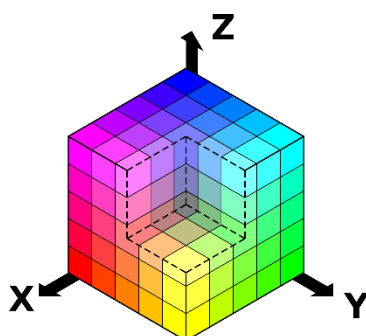
รูปที่ 2-2 หน้าตัดของปลายท่อนซุง (ก) หน้าตัดของปลายท่อนซุงไม้สน (ข) หน้าตัดของปลายท่อนซุง ไม้ยูคาลิปตัส

2.2 แบบจำลองสี (Color model)

แบบจำลองสี หรือระบบสีเป็นมาตรฐานในการอ้างอิงถึงสีใดสีหนึ่ง และระบุสีแต่ละสี ซึ่งพบเห็นได้ทั่วไปจากจอภาพ ภาพถ่าย และงานพิมพ์ โดยแบบจำลองมีคุณสมบัติของแม่สี และเกณฑ์การผสมสีที่แตกต่างกัน ให้เลือกใช้ตามความเหมาะสมกับงานนั้น ๆ ในงานวิจัยนี้สนใจแบบจำลองสีเพียง 2 แบบจำลองสีเท่านั้น ได้แก่ แบบจำลองสี RGB และแบบจำลองสี HSV [6] ซึ่งมีรายละเอียดดังนี้

2.2.1 แบบจำลองสีอาร์จีบี (RGB color model)

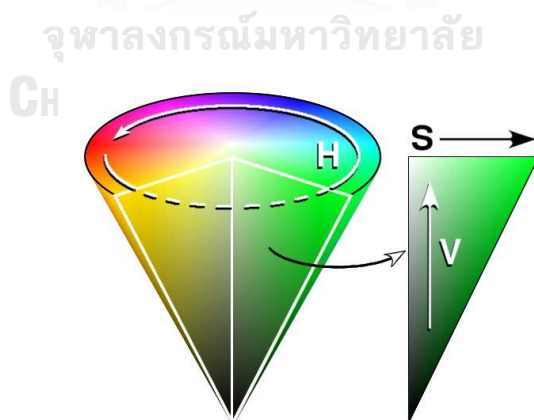
แบบจำลองสี RGB ประกอบขึ้นจากพื้นฐานสีจำนวน 3 สีคือ สีแดง (Red) สีเขียว (Green) และสีน้ำเงิน (Blue) โดยการสร้างสีที่ต้องการจะถูกระบุด้วยค่าสีทั้งสามสีนี้ผสมในอัตราส่วนที่แตกต่างกัน [6] ซึ่งสีที่ได้จากการผสมสีขึ้นอยู่กับระดับความเข้มของสีนั้น ถ้ามีความเข้มสูง เมื่อนำมาผสมกัน จะทำให้เกิดสีขาว แต่ถ้ามีความเข้มต่ำ จะทำให้เกิดสีดำ จึงเรียกแบบจำลองสี RGB นี้ว่าการผสมสีแบบบวก (Additive) ซึ่งเหมาะกับการแสดงผลบนอุปกรณ์ที่เกี่ยวข้องกับแสง เช่น จอภาพ กล้องดิจิทัล หรือภาพบนเว็บไซต์ เป็นต้น



รูปที่ 2-3 แบบจำลองสี RGB บนแกนสามมิติ [7]

2.2.2 แบบจำลองสีเอชเอสวี (HSV color model)

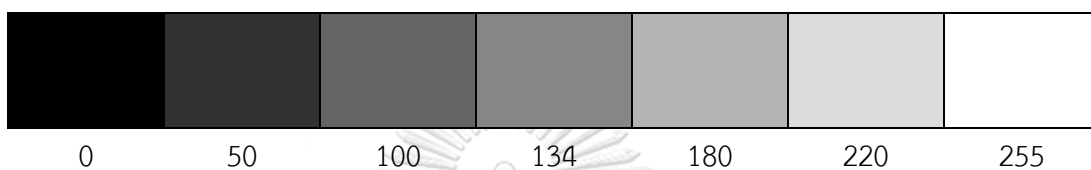
แบบจำลองสี HSV มีลักษณะเป็นพีระมิดฐานหกเหลี่ยม หรือทรงกรวยสี ซึ่งแปลงมาจากแบบจำลองสี RGB ที่ผ่านกระบวนการเปลี่ยนรูปที่ไม่เป็นเชิงเส้น (Non-linear transformation) เนื่องจากการระบุค่าสีของ RGB นั้นใช้เฉพาะแม่สีทั้งสามสีทำให้ไม่สะดวกในการใช้สีต่าง ๆ ดังนั้นในแบบจำลองสี HSV จึงใช้การระบุค่าสีแบบ Hue (ค่าสี) Saturation (ความอิ่มตัวสี) และ Value (ความสว่างสี) [6] ตัวอย่างเช่น เมื่อต้องการเปลี่ยนจากสีแดงอ่อนไปเป็นสีเขียวอ่อนที่มีระดับความอิ่มตัวเท่ากัน สำหรับแบบจำลองสี RGB ต้องเปลี่ยนค่าทั้งสามค่าเพื่อให้ได้สีเขียวอ่อนออกมา แต่ในแบบจำลองสี HSV จะทำการเปลี่ยนเฉพาะค่าสีเท่านั้น แสดงให้เห็นว่าแบบจำลองสี HSV เป็นแบบจำลองสีที่ทำให้ง่ายต่อการเลือกสี และการแบ่งแยกความแตกต่างของสี



รูปที่ 2-4 กรวยแสดงแบบจำลองสี HSV บนแกนสามมิติ [7]

2.3 ภาพระดับสีเทา (Grayscale image)

ภาพระดับสีเทาเป็นภาพที่แสดงถึงระดับเฉดของสีเทาเพียงสีเดียว โดยค่าแต่ละจุดในภาพคือค่าความเข้มของระดับเทา ซึ่งมักถูกจัดเก็บในขนาด 8 บิต และให้ค่าเฉดสีระดับเทาจากสีขาวไปจนถึงสีดำตั้งแต่ช่วง 0-255 ดังรูปที่ 2-5 โดยภาพระดับเทานิยมนำไปใช้ในกระบวนการต่าง ๆ ของการประมวลผลภาพถ่าย เพราะค่าความเข้มของระดับเทามีความสำคัญและสามารถแยกลักษณะเฉพาะของภาพได้



รูปที่ 2-5 ตัวอย่างภาพระดับสีเทาในขนาด 8 บิต

2.3.1 การแปลงภาพสีจากแบบจำลองอาร์จีบีเป็นภาพระดับสีเทา (RGB to grayscale image conversion)

ภาพในแบบจำลองอาร์จีบีมีค่าความเข้มของค่าสีพื้นฐานของสีแดง สีเขียว และสีน้ำเงิน โดยในแต่ละจุดภาพจะมีค่าสีพื้นฐานดังกล่าวตั้งแต่ 0-255 ซึ่งการแปลงภาพสีอาร์จีบีเป็นภาพระดับเทาทำให้แต่ละจุดภาพมีค่าความเข้มของสีเป็น 0-255 ดังรูปที่ 2-6 (ข) โดยใช้สมการดังต่อไปนี้

$$Y = 0.299 \times R + 0.587 \times G + 0.114 \times B$$

โดยที่

Y คือ ผลลัพธ์ค่าระดับเทาในแต่ละจุดภาพ

R คือ ค่าความเข้มของสีแดงในแต่ละจุดภาพ

G คือ ค่าความเข้มของสีเขียวในแต่ละจุดภาพ

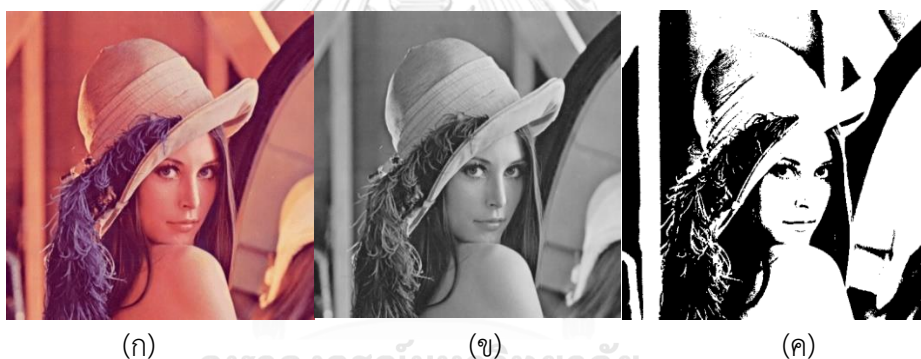
B คือ ค่าความเข้มของสีน้ำเงินในแต่ละจุดภาพ

2.4 ภาพไบนารี (Binary image)

ภาพไบนารีเป็นภาพขาวดำที่มีค่าความเข้มเท่ากับ 0 และ 1 (หรือ 0 และ 255) ในแต่ละจุดภาพเท่านั้น และมีขนาด 1 บิตต่อจุดภาพ ซึ่งภาพชนิดนี้เป็นหลักการพื้นฐานของกระบวนการประมวลผลภาพและถูกนำไปใช้อยู่บ่อยครั้ง เช่น การแยกส่วนภาพ (Image segmentation) การหาขอบภาพ (Edge detection) หรือการกำหนดค่าขีดแบ่ง (Threshold) เป็นต้น

2.4.1 การแปลงภาพระดับเทาเป็นภาพไบนารีด้วยวิธีโอทสุ (Grayscale to binary image conversion by Otsu's thresholding algorithm)

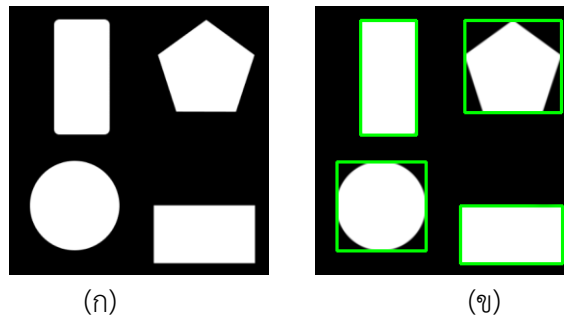
วิธี Otsu's thresholding [8] เป็นการแปลงภาพเป็นภาพไบนารีจากภาพระดับเทา โดยการคำนวณหาค่าขีดแบ่งแบบอัตโนมัติจากการวนซ้ำค่าขีดแบ่งที่น่าจะเป็นไปได้และกำหนดให้ภาพนำเข้าแบ่งออกเป็น 2 กลุ่มของจุดภาพ จากนั้นคำนวณหาค่าขีดแบ่งจากค่าความแปรปรวนระหว่างกลุ่มสูงสุดสำหรับแยก 2 กลุ่มของจุดภาพออกจากกัน ดังผลลัพธ์รูปที่ 2-6



รูปที่ 2-6 การแปลงภาพสีจากอาร์จีบีเป็นภาพไบนารี (ก) ภาพต้นฉบับ (ข) ภาพระดับสีเทา (ค) ภาพไบนารี [9]

2.5 สี่เหลี่ยมในการปิดล้อมวัตถุ (Bounding box)

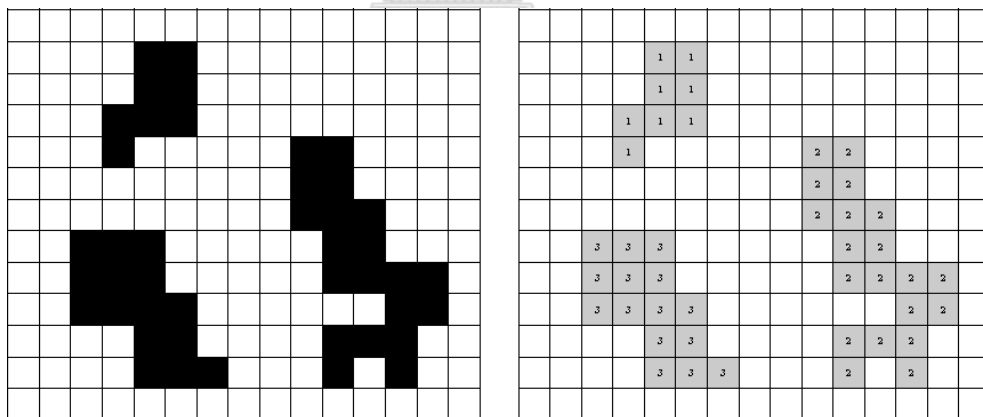
สี่เหลี่ยมในการปิดล้อมวัตถุคือกล่องสี่เหลี่ยมซึ่งสร้างขึ้นสำหรับล้อมรอบบริเวณวัตถุหรือตำแหน่งที่ต้องการในภาพ (Region of Interest : ROI) ยกตัวอย่างการรู้จัก (Feature recognition) หรือตรวจจับลักษณะเฉพาะ (Feature detection) ในภาพจะได้ผลลัพธ์โดยมีสี่เหลี่ยมในการปิดล้อมวัตถุที่ต้องการล้อมรอบ ซึ่งประกอบไปด้วยพิกัดของจุดภาพ รวมถึงความกว้าง ความยาว พื้นที่ และจุดศูนย์กลางของวัตถุนั้น ๆ ดังรูปที่ 2-7 เป็นการใช้สี่เหลี่ยมในการปิดล้อมรูปทรงเรขาคณิตที่ต้องการ



รูปที่ 2-7 การตีกรอบสี่เหลี่ยมปิดล้อมรูปทรงเรขาคณิตในภาพ (ก) ภาพต้นฉบับ (ข) ภาพผลลัพธ์ที่ตีกรอบปิดล้อมวัตถุ

2.6 การติดป้ายส่วนประกอบที่เชื่อมติดกัน (Connected component labeling)

การติดป้ายส่วนประกอบที่เชื่อมติดกันเป็นการพิจารณาจุดภาพทั้งหมดบนภาพถ่ายว่า จุดภาพใดบ้างที่มีการเชื่อมติดกันและดำเนินการติดป้ายให้กลุ่มของจุดภาพนั้น ๆ นอกจากนี้ยังทราบขนาด จำนวน และตำแหน่งวัตถุในภาพที่พิจารณาจากจุดภาพที่เชื่อมติดกัน แสดงดังรูปที่ 2-8 ซึ่งภาพที่ใช้ในการติดป้ายส่วนประกอบที่เชื่อมติดกัน เป็นภาพขาวดำที่มีค่า 0 และ 1 (Binary image) โดย จะทำการติดป้ายบริเวณจุดภาพส่วนที่เป็นสีขาว

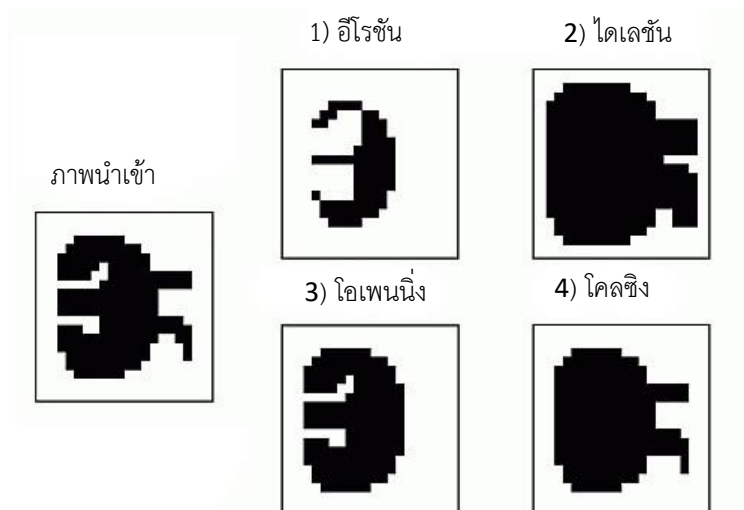


รูปที่ 2-8 การติดป้ายส่วนประกอบที่เชื่อมติดกัน [10]

2.7 การเปลี่ยนแปลงลักษณะโครงสร้างของภาพ (Morphological processing)

การเปลี่ยนแปลงโครงสร้างของภาพ คือ เซตของการประมวลผลทางภาพทางด้านโครงสร้าง โดยนำตัวประกอบโครงสร้าง (Structure element) ซึ่งทำหน้าที่เป็นหน้ากาก (Mask) ที่มีค่า 0 และ 1 และมีหลายลักษณะ เช่น วงกลม สี่เหลี่ยม เส้นตรง จาน หรือสี่เหลี่ยมขนมเปียกปูน เป็นต้น กระทำ

กับภาพนำเข้าที่เป็นภาพไบนารีจะได้ภาพผลลัพธ์ที่มีขนาดเท่ากับภาพนำเข้า โดยกระบวนการเปลี่ยนแปลงลักษณะโครงร่างของภาพมีอยู่ 4 ชนิด ดังรูปที่ 2-9



รูปที่ 2-9 การเปลี่ยนแปลงลักษณะโครงร่างของภาพทั้ง 4 ชนิด [11]

1) อีโรชัน (Erosion) คือ การกร่อนวัตถุในภาพให้มีขนาดเล็กลง ด้วยการกำหนดตัวประกอบโครงสร้าง และนำไปกราดบนข้อมูลภาพทั้งหมด ซึ่งจะเลื่อนไปทุกตำแหน่งเพื่อเปรียบเทียบกับข้อมูลภาพ ถ้ามีค่าตรงกันจะกำหนดค่าข้อมูลภาพ โดยใช้ในการลดขนาดวัตถุให้มีขนาดเล็กลง และใช้ในการกำจัดสัญญาณรบกวนที่อยู่ในภาพ

2) ไดเลชัน (Dilation) คือ การขยายวัตถุในภาพให้มีขนาดใหญ่ขึ้น มีวิธีคล้ายกับการกร่อน ด้วยการกำหนดตัวประกอบโครงสร้าง และนำไปกราดบนข้อมูลภาพทั้งหมด ซึ่งจะเลื่อนไปทุกตำแหน่งเพื่อเปรียบเทียบกับข้อมูลภาพ ถ้ามีค่าตรงกันจะกำหนดค่าข้อมูลภาพ โดยใช้ในการเชื่อมระหว่างวัตถุ 2 วัตถุที่อยู่ใกล้กันหรือเชื่อมวัตถุที่ขาดออกจากกัน รวมถึงใช้ในการปิดรู และใช้ในการกำจัดสัญญาณรบกวนที่อยู่ในภาพ

3) โอเพนนิ่ง (Opening) คือ การนำกระบวนการกร่อนมาดำเนินการกับภาพ และตามด้วยกระบวนการ dilate โดยใช้ในการลบสัญญาณรบกวน และขอบวัตถุที่ไม่ต้องการ

4) โคลซิง (Closing) คือ การนำกระบวนการขยายมาดำเนินการกับภาพ และตามด้วยกระบวนการ erode โดยใช้ในการเชื่อมวัตถุที่แยกจากกัน

2.8 การเรียนรู้ของเครื่อง (Machine learning)

การเรียนรู้ของเครื่องเป็นการที่ให้คอมพิวเตอร์เรียนรู้ด้วยตัวเองเหมือนพฤติกรรมการเรียนรู้ของมนุษย์โดยเรียนรู้จากข้อมูลจำนวนมาก ซึ่งคอมพิวเตอร์สังเกตรูปแบบข้อมูล (Patterns) ก่อนหน้าและตัดสินใจ คาดการณ์ผลลัพธ์หรือตอบคำถามที่ตั้งไว้ได้ โดยที่ผู้วิเคราะห์ข้อมูลไม่ต้องเขียนโปรแกรมเพื่อแก้ปัญหาทางนั้น ๆ แบบเฉพาะเจาะจง การเรียนรู้ของเครื่องเป็นวิธีที่แตกต่างไปจากปัญญาประดิษฐ์ (Artificial Intelligence : AI) ที่นักพัฒนาสร้างกฎแบบทีละขั้นตอนเพื่อให้คอมพิวเตอร์ทำงานตามชุดคำสั่งที่สร้างไว้ ในทางกลับกันการเรียนรู้ของเครื่องจะให้คอมพิวเตอร์เรียนรู้จากข้อมูลด้วยตัวเองและสามารถพัฒนาประสิทธิภาพในการแก้ไขปัญหาที่ตั้งไว้ได้อย่างฉลาด ตัวอย่างเช่น ระบบเทคโนโลยียานยนต์อัจฉริยะ ซึ่งสามารถขับเคลื่อนเองได้โดยไม่ต้องมีคนขับ หรือเทคโนโลยีการจดจำเสียงสิริ (Siri voice recognition) ที่เปรียบเสมือนผู้ช่วย ซึ่งผู้ใช้งานสามารถใช้คำสั่งเสียงสั่งการ หรือขอความช่วยเหลือได้ทันทีแบบทันกาล เป็นต้น การเรียนรู้ของเครื่องแบ่งออกเป็น 5 ประเภทหลัก ดังนี้

2.8.1 การเรียนรู้แบบมีผู้สอน (Supervised learning)

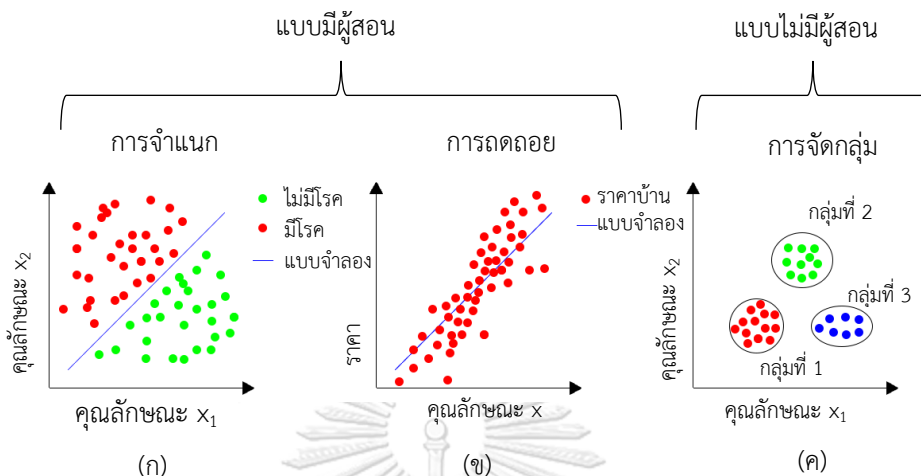
เป็นการเรียนรู้ของเครื่องจากข้อมูลตัวอย่างการสอนที่ติดป้ายกำกับล่วงหน้า (Pre-labeled training) เพื่อนำไปสร้างอัลกอริทึมหรือแบบจำลองในการทำนายผลลัพธ์ รวมถึงเพื่อเรียนรู้ความสัมพันธ์ระหว่างข้อมูลนำเข้าและผลลัพธ์ ซึ่งแบบจำลองจากการเรียนรู้ประเภทนี้จะสามารถนำไปใช้กับข้อมูลนำเข้าชนิดใหม่ได้และการเรียนรู้แบบมีผู้สอนจะสามารถประเมินผลลัพธ์ได้ว่ามีความถูกต้องหรือผิดพลาด

2.8.2 การเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning)

เป็นการเรียนรู้ของเครื่องจากข้อมูลตัวอย่างที่ไม่ได้มีการติดป้ายกำกับ (Unlabeled) เพื่อนำไปค้นหาและจำลองรูปแบบหรือโครงสร้างพื้นฐานที่ซ่อนอยู่ในข้อมูลนำเข้า ซึ่งตรงกันข้ามกับการเรียนรู้แบบมีผู้สอนที่คอมพิวเตอร์ไม่ทราบค่าประเภทของตัวอย่างในข้อมูลสอนล่วงหน้าและพยายามเรียนรู้เพื่อระบุประเภทหรือกลุ่มที่อยู่ในข้อมูลตัวอย่าง

ในกรณีของการเรียนรู้แบบมีผู้สอนจะรวบรวมในรูปแบบป้ายชื่อกำกับ (Label) ส่วนกรณีของการเรียนรู้แบบไม่มีผู้สอนสามารถเรียกได้ว่าเป็นการจัดกลุ่ม (Clustering) ซึ่งจะจัดกลุ่มข้อมูลที่ไม่ได้ถูกป้ายชื่อให้มีความหมายตามกลุ่มของข้อมูลโดยไม่ได้อาศัยความรู้ก่อนหน้า (Priori knowledge) สำหรับผลลัพธ์ของการเรียนรู้แบบมีผู้สอนอยู่ในรูปแบบการจำแนก (Classification) และการถดถอย

(Regression) ส่วนผลลัพธ์ของการเรียนรู้แบบไม่มีผู้สอนอยู่ในรูปแบบการจัดกลุ่มและความสัมพันธ์ ดังรูปที่ 2-10



รูปที่ 2-10 ตัวอย่างของปัญหาในชีวิตจริงของผลลัพธ์จากการเรียนรู้แบบมีผู้สอนและไม่มีผู้สอน (ก) การจำแนกประเภทผู้ป่วยว่ามีโรคหรือไม่ (ข) การทำนายราคาบ้าน (ค) การจัดกลุ่มพฤติกรรมกรรมการซื้อของลูกค้า

2.8.3 การเรียนรู้แบบกึ่งมีผู้สอน (Semi-supervised learning)

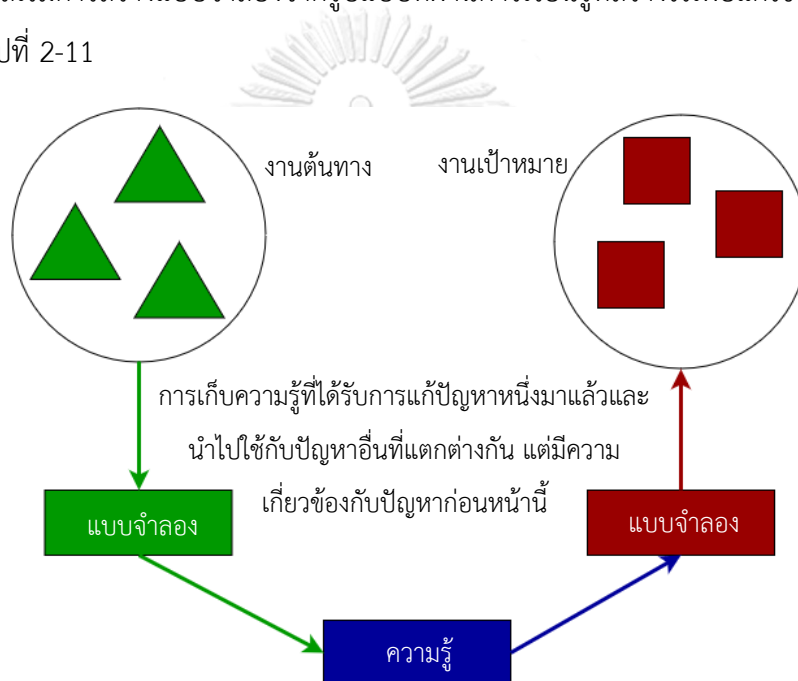
เป็นการเรียนรู้ข้อมูลแบบผสมผสานระหว่างการเรียนรู้แบบมีผู้สอนและไม่มีผู้สอน โดยข้อมูลมีทั้งข้อมูลที่ผู้วิเคราะห์กำหนดลักษณะของประเภทข้อมูลเอง (Label data) จำนวนหนึ่งและข้อมูลที่ไม่ได้มีการกำหนดลักษณะหรือติดป้ายของภาพไว้ (Unlabeled data) จำนวนมาก จากนั้นข้อมูลเหล่านี้นำไปเข้าขั้นตอนวิธีเพื่อจัดกลุ่มของข้อมูลต่อไป

2.8.4 การเรียนรู้แบบเสริมกำลัง (Reinforcement learning)

เป็นการเรียนรู้แบบตั้งเป้าหมายให้เครื่อง ซึ่งเรียนรู้จากสภาพแวดล้อมรอบข้างในรูปแบบต่าง ๆ เช่น ข้อมูลนำเข้า หรือตัวรับรู้ (Sensor) เป็นต้น โดยให้รางวัลหรือการลงโทษเป็นผลตอบแทน และการเรียนรู้จะไม่มีกรอบว่าการกระทำไหนถูกหรือผิด ยกตัวอย่างเช่น ในการเล่นเกมส์ผู้เล่นทราบว่าเมื่อไหร่จะชนะหรือแพ้ แต่กลับกันผู้เล่นจะไม่ทราบวิธีการเล่นในแต่ละขั้นตอน สำหรับการเรียนรู้แบบเสริมกำลังนี้จะเรียนรู้จากประสบการณ์และสภาพแวดล้อมรอบข้าง ทำให้ทราบถึงขั้นตอนในการเล่นเกมส์ล่วงหน้าได้

2.8.5 การถ่ายโอนการเรียนรู้ (Transfer learning)

เป็นวิธีในการเรียนรู้เครื่องที่ใช้ความรู้จากการเรียนรู้ที่ได้มีการแก้ปัญหาได้แล้วจากสภาพแวดล้อมหนึ่งมาใช้ในสภาพแวดล้อมใหม่ที่มีปัญหาใกล้เคียงกัน สำหรับการถ่ายโอนการเรียนรู้นิยมใช้ในด้านการเรียนรู้แบบเชิงลึก (Deep learning) เนื่องจากแบบจำลองที่ได้รับการสอนล่วงหน้า (Pre-trained model) ถูกนำไปใช้เป็นจุดเริ่มต้นของงานด้านคอมพิวเตอร์วิทัศน์และการประมวลผลภาษาธรรมชาติ (Natural language processing : NLP) ซึ่งการใช้ Pre-trained model สามารถช่วยลดการคำนวณและประหยัดเวลาในการพัฒนาโครงข่ายนิวรอลเน็ตเวิร์ก (Neural network) เพราะสามารถเริ่มการสร้างแบบจำลองจากรูปแบบที่ผ่านการเรียนรู้ที่สร้างไว้เพื่อแก้ไขปัญหาทางก่อนหน้านี้ได้ดังรูปที่ 2-11

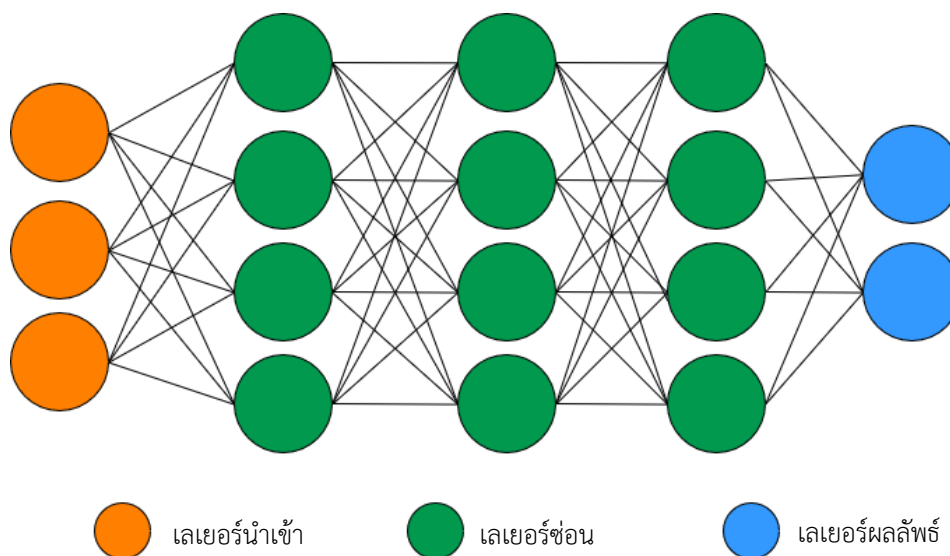


รูปที่ 2-11 การประยุกต์ความรู้ที่เรียนรู้ก่อนหน้าไปยังแบบจำลองใหม่

2.9 การเรียนรู้เชิงลึก (Deep learning)

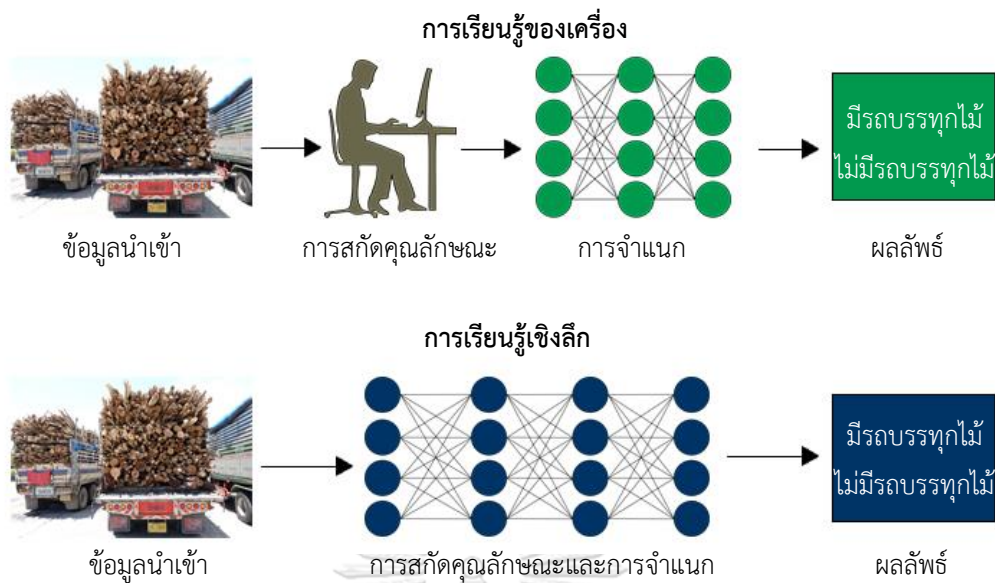
การเรียนรู้เชิงลึกเป็นแขนงย่อยของการเรียนรู้เครื่อง ซึ่งแต่ละงานแบ่งย่อยและกระจายไปยังแต่ละอัลกอริทึมของการเรียนรู้เครื่องที่ถูกจัดให้อยู่ในรูปแบบของลำดับชั้นที่ต่อเนื่องกัน โดยในแต่ละเลเยอร์ถูกสร้างขึ้นจากผลลัพธ์จากเลเยอร์ก่อนหน้า จากนั้นแต่ละเลเยอร์ประกอบกันเป็นโครงข่ายประสาทเทียม ซึ่งเหมือนกับวิธีการกระจายการแก้ปัญหาของเซลล์ประสาทในสมองมนุษย์ หรือเรียกได้ว่าเป็นการสะท้อนการทำงานของสมองคน อัลกอริทึมของการเรียนรู้เชิงลึกนั้นคล้ายกับ

การทำงานของโครงสร้างระบบประสาท โดยที่แต่ละเซลล์ประสาทเชื่อมต่อซึ่งกันและกัน รวมถึงมีการส่งข้อมูลระหว่างเซลล์ประสาท



รูปที่ 2-12 ตัวอย่างโครงสร้างเลเยอร์ของแบบจำลองการเรียนรู้เชิงลึก

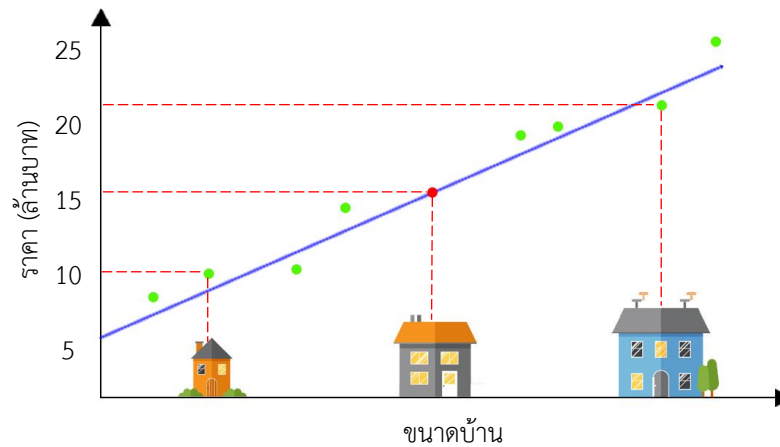
แบบจำลองของการเรียนรู้เชิงลึกทำงานเป็นลำดับขั้นดังรูปที่ 2-12 โดยที่แบบจำลองทั่วไปจะมีเลเยอร์ 3 เลเยอร์เป็นอย่างน้อย ซึ่งแต่ละเลเยอร์รับข้อมูลมาจากเลเยอร์ก่อนหน้าและส่งต่อไปยังเลเยอร์ถัดไป นอกจากนี้แบบจำลองเชิงลึกมีแนวโน้มที่จะทำงานได้ดีกับข้อมูลจำนวนมาก ในขณะที่การเรียนรู้ของเครื่องจะหยุดพัฒนาประสิทธิภาพหลังจากถึงจุดอิ่มตัว ความแตกต่างระหว่างการเรียนรู้ของเครื่องและการเรียนรู้เชิงลึกอยู่ที่การสกัดคุณลักษณะของข้อมูลดังรูปที่ 2-13 โดยการเรียนรู้ของเครื่องจะถูกสกัดคุณลักษณะโดยมนุษย์ ส่วนการเรียนรู้เชิงลึกจะสกัดคุณลักษณะออกมาด้วยตัวเอง และในการเรียนรู้เชิงลึกมีทฤษฎีที่เกี่ยวข้อง ซึ่งเป็นพื้นฐานของการเรียนรู้เชิงลึก ดังนี้



รูปที่ 2-13 เปรียบเทียบความแตกต่างในการสกัดคุณลักษณะของข้อมูล ระหว่างการเรียนรู้ของเครื่อง และการเรียนรู้เชิงลึก

2.9.1 การถดถอยเชิงเส้น (Linear regression)

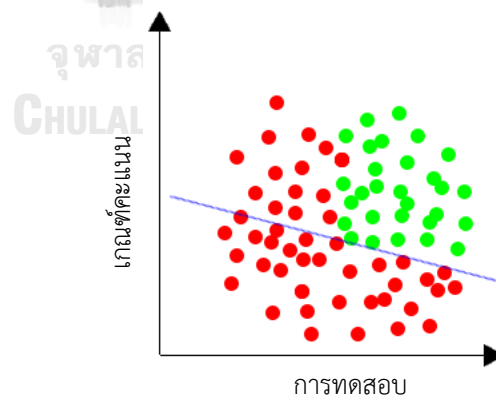
เป็นวิธีการทางสถิติที่ช่วยสรุปและศึกษาความสัมพันธ์ระหว่างตัวแปรต่อเนื่องจำนวน 2 ตัวแปร ยกตัวอย่างในกรณีที่มีข้อมูลขนาดของบ้านย้อนหลัง ซึ่งวาดและกำหนดจุดลงในกราฟดังรูปที่ 2-14 โดยการถดถอยเชิงเส้นเป็นเทคนิคการหาเส้นตรงระหว่างจุดที่วาดบนกราฟให้มีข้อผิดพลาดน้อยลงและเมื่อเส้นตรงมีข้อผิดพลาดน้อยลง ทำให้สามารถคาดการณ์ราคาบ้านโดยอ้างอิงตามขนาดของบ้านได้



รูปที่ 2-14 ตัวอย่างแสดงการทำงานของ การถดถอยเชิงเส้น ซึ่งประมาณการราคาบ้านที่ดีที่สุด

2.9.2 การถดถอยโลจิสติก (Logistic regression)

เป็นวิธีการทางสถิติสำหรับวิเคราะห์ชุดข้อมูลที่มีตัวแปรอิสระอย่างน้อย 1 ตัวเป็นตัวกำหนดผลลัพธ์ โดยผลลัพธ์ที่ได้มีเพียง 2 ผลลัพธ์ที่เป็นไปได้ ได้แก่ จริงหรือเท็จ ยกตัวอย่างในกรณีที่มีข้อมูลนักเรียนที่ผ่านและไม่ผ่านเกณฑ์คะแนนและผลการทดสอบที่ตั้งไว้ ถ้าต้องการทราบว่าจะนักเรียนจะผ่านหรือไม่จะขึ้นอยู่กับเกณฑ์คะแนนและผลการทดสอบ ซึ่งในกรณีนี้สามารถนำการถดถอยโลจิสติกมาประยุกต์โดยหาเส้นตรงที่จำแนกข้อมูลทั้ง 2 ประเภทนี้ (ผ่านและไม่ผ่าน) ออกจากกันได้ดีที่สุดดังรูปที่ 2-15 ซึ่งมีลักษณะคล้ายกับการถดถอยเชิงเส้น



รูปที่ 2-15 ตัวอย่างแสดงการทำงานของ การถดถอยโลจิสติก ซึ่งคำนวณหา นักเรียนที่ผ่านเกณฑ์คะแนนและผลการทดสอบ

2.9.3 ฟังก์ชันการกระตุ้น (Activation function)

เป็นฟังก์ชันที่ทำหน้าที่ตัดสินใจและกำหนดขอบเขตสำหรับตัวแปรขาออก (Output) ที่ออกมาจากในแต่ละโหนด (Node) ของเซลล์ประสาทให้เป็นมาตรฐานเดียวกันและช่วยทำให้ตัวแปรขาออกมีประสิทธิภาพมากขึ้น โดยฟังก์ชันการกระตุ้นสามารถกำหนดขอบเขตค่าที่เป็นไปได้ของตัวแปรขาออกให้น้อยลงที่สุดจนถึงขอบของกลุ่มข้อมูลนั้น ยกตัวอย่างฟังก์ชันเฮฟวิไซด์ (Heaviside step function) เป็นตัวอย่างของฟังก์ชันกระตุ้นที่ง่ายที่สุด กล่าวคือฟังก์ชันจะคืนค่าเท่ากับ 0 เมื่อผลรวมเชิงเส้นมีค่าน้อยกว่า 0 และจะคืนค่าเท่ากับ 1 เมื่อผลรวมเชิงเส้นเป็นจำนวนเต็มบวกหรือเท่ากับศูนย์ ซึ่งคำนวณได้จากสมการที่ 2

$$f(h) = \begin{cases} 0 & \text{if } h < 0 \\ 1 & \text{if } h \geq 0 \end{cases} \quad (2)$$

โดยที่

$f(h)$ คือ ผลลัพธ์จากตัวแปรขาออก

h คือ ค่าจากตัวแปรขาเข้าไปยังตัวแปรขาออก

2.9.4 ค่าน้ำหนัก (Weight)

เป็นค่าตัวแปรที่เกี่ยวข้องกับการเชื่อมต่อระหว่างเซลล์ประสาทในแต่ละครั้ง ซึ่งบ่งบอกถึงความสำคัญของค่าตัวแปรขาเข้า เมื่อมีตัวแปรขาเข้าเข้ามาในเซลล์ประสาทจะคูณด้วยค่าน้ำหนักที่กำหนดให้กับข้อมูลขาเข้านี้ โดยปกติค่าน้ำหนักเริ่มต้นจะถูกตั้งค่าแบบสุ่มและเมื่อเซลล์ประสาทเริ่มเรียนรู้ จะมีการปรับค่าน้ำหนักเพื่อให้ได้ผลลัพธ์ที่ใกล้เคียงกับคำตอบที่ต้องการมากที่สุด ยกตัวอย่างเช่น มหาวิทยาลัยมีเกณฑ์การทดสอบนักศึกษาโดยใช้คะแนนสอบ (Test scores) และระดับคะแนน (Grades) ดังนั้นเซลล์ประสาทนี้จึงมีค่าน้ำหนักที่เกี่ยวข้องจำนวน 2 ค่าและสามารถปรับเปลี่ยนค่าน้ำหนักในแต่ละค่าได้

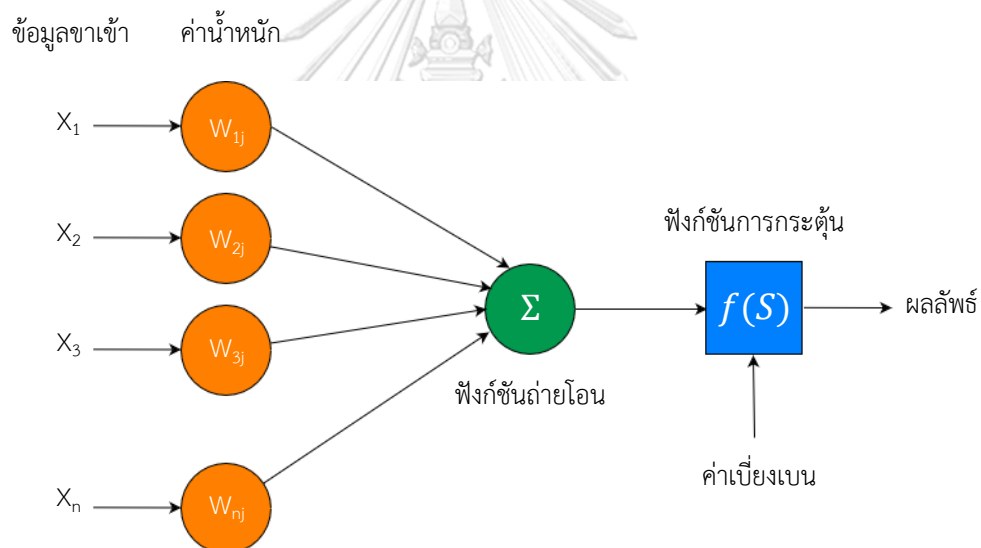
ถ้าความถูกต้องหรือประสิทธิภาพมีความแม่นยำสูงหรือใกล้เคียงกับผลลัพธ์ที่ต้องการ จะไม่มีการปรับค่าน้ำหนัก แต่ถ้าประสิทธิภาพต่ำ ค่าน้ำหนักจะถูกปรับผ่านการคำนวณเฉพาะ

2.9.5 ค่าเบี่ยงเบน (Bias)

เป็นค่าตัวแปรที่ช่วยปรับค่าตัวแปรขาเข้าให้กราฟขยับไปทางซ้ายหรือทางขวา (0-1) โดยค่าเบี่ยงเบนจะมีการปรับค่าในขณะที่มีการเรียนรู้ทุกครั้งเช่นเดียวกับค่าน้ำหนักและทำหน้าที่หลีกเลี่ยงไม่ให้เซลล์ประสาทส่งผลลัพธ์ที่ไม่มีค่า (เท่ากับ 0) ไปยังเลเยอร์ผลลัพธ์

2.9.6 โครงข่ายประสาทเทียม (Artificial Neural Network : ANN)

เป็นโครงข่ายของเซลล์ประสาทเทียมหรือโหนดที่เชื่อมโยงถึงกันโดยที่แต่ละเซลล์ประสาทแสดงถึงหน่วยประมวลผลข้อมูล ซึ่งโหนดที่เชื่อมโยงกันเหล่านี้จะมีปฏิสัมพันธ์ ส่งข้อมูลให้แกกัน และดำเนินการบางอย่างก่อนที่จะส่งข้อมูลต่อไปเหมือนกับการทำงานของสมองมนุษย์ สำหรับการดำเนินการที่กล่าวไว้ก่อนหน้านี้นี้จะกระทำโดยฟังก์ชันกระตุ้นในข้อที่ 2.8.3 จากนั้นจะแปลงข้อมูลขาเข้าเป็นผลลัพธ์ ซึ่งสามารถใช้ผลลัพธ์นี้เป็นข้อมูลขาเข้าของโหนดอื่น ๆ ได้



รูปที่ 2-16 ภาพตัวอย่างการทำงานของโครงข่ายประสาทเทียม

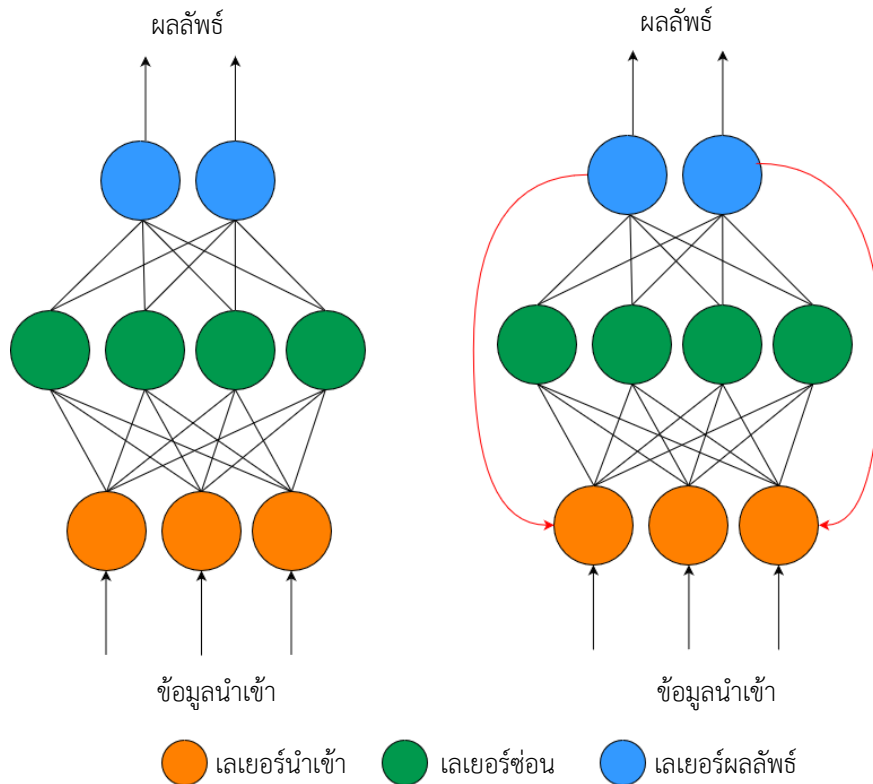
จากรูปที่ 2-16 เป็นการทำงานของโครงข่ายประสาทเทียม โดยเซลล์ประสาทเทียมจะรับข้อมูลขาเข้าจากโหนดและใช้ฟังก์ชันการกระตุ้นเพื่อรวมค่าน้ำหนักของข้อมูลขาเข้าโดยใช้ฟังก์ชันถ่ายโอน (Transfer function) และส่งข้อมูลไปเป็นผลลัพธ์ที่ต้องการ นอกจากนี้ค่าขีดแบ่งหรือค่าเบี่ยงเบนถูกนำมาใช้รวมกับผลรวมของค่าน้ำหนักเพื่อหลีกเลี่ยงไม่ให้ค่าผลลัพธ์มีค่าเท่ากับ 0

เทคนิคโครงข่ายประสาทเทียมถูกใช้ในการประมวลผลภาพและการค้นคืนภาพ ซึ่งเป็นระบบที่อาศัยการคำนวณโดยใช้เซลล์ประสาทที่แบ่งเป็นแต่ละเลเยอร์ หน้าที่หลักของเซลล์ประสาทเพื่อเอาไว้อรับค่าฟังก์ชันกระตุ้นจากเพื่อนบ้านใกล้เคียง นั่นคือผลลัพธ์จากเซลล์ประสาทอื่น ๆ และนำมาประมวลผลออกเป็นผลลัพธ์ โดยขึ้นอยู่กับค่าน้ำหนักของตัวแปรขาเข้าที่เข้ามา แบบจำลองโครงข่ายประสาทเทียมถูกศึกษามาเป็นระยะเวลาานาน โดยมีจุดประสงค์ให้มีการเรียนรู้เหมือนกับที่มนุษย์คิดหรือฉลาดกว่ามนุษย์ สามารถสร้างเครื่องมือในการเรียนรู้จัดจำรูปแบบ รวมถึงการสร้างความรู้ใหม่ ยกตัวอย่างในกรณีที่ต้องการค้นหารูปแบบจากคุณลักษณะในภาพ โดยสามารถใส่ภาพทดสอบเข้าไปยังโครงข่ายประสาทเทียมเพื่อให้ระบบเรียนรู้รูปแบบที่ต้องการสืบค้น และหลังจากนั้นก็ทดลองใส่รูปภาพที่ไม่ใช่ชุดทดสอบเข้าไป ถ้าตรวจสอบพบรูปแบบที่ต้องการแสดงว่าโครงข่ายประสาทเทียมสามารถเรียนรู้จากรูปภาพตัวอย่างได้ดี

สำหรับการเรียนรู้แบบเชิงลึกจะมีชั้นของเซลล์ประสาทหลายชั้น (มากกว่า 100 ชั้น) เชื่อมต่อและส่งข้อมูลให้กันและกันในรูปแบบป้อนไปข้างหน้า (Feedforward) หรือแบบย้อนกลับ (Feedback) ดังรูปที่ 2-17

1) แบบป้อนไปข้างหน้า (Feedforward) เป็นประเภทของโครงข่ายประสาทเทียมที่ง่ายที่สุด การเชื่อมต่อระหว่างชั้นไม่สามารถวนกลับได้ ข้อมูลขาเข้าจะถูกส่งตรงไปยังผลลัพธ์ในทิศทางเดียว โดยผ่านค่าน้ำหนักของแต่ละการเชื่อมต่อ มักนิยมใช้ในงานด้านการรู้จำรูปแบบ (Pattern recognition)

2) แบบย้อนกลับ (Feedback) เป็นการเชื่อมต่อในลักษณะที่วนกลับได้ ผลลัพธ์ที่ได้รับจากโครงข่ายจะถูกส่งวนกลับเข้าโครงข่ายอีกครั้งเพื่อปรับปรุงประสิทธิภาพของผลลัพธ์ โดยโครงข่ายแบบย้อนกลับนี้อาจมีความซับซ้อนมากและมีประสิทธิภาพมากกว่าแบบป้อนไปข้างหน้า นิยมใช้ในงานหลากหลายด้าน



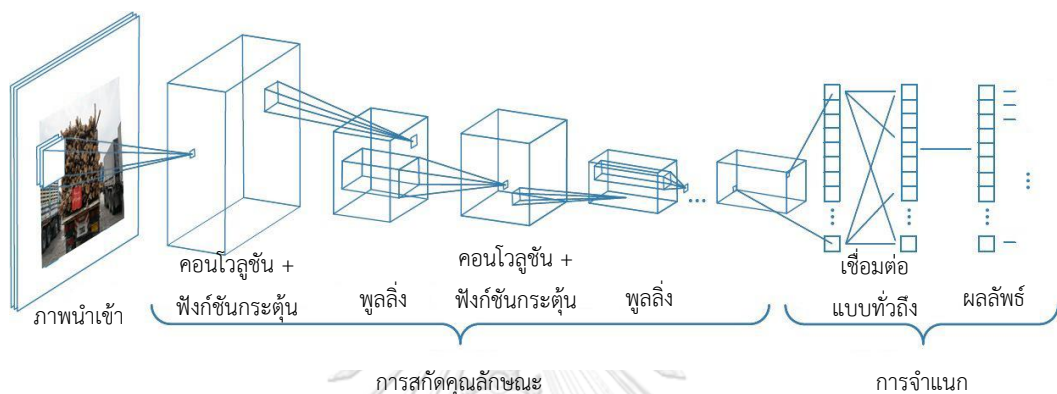
รูปที่ 2-17 ประเภทของโครงข่ายประสาทเทียม (ก) แบบป้อนไปข้างหน้า (ข) แบบย้อนกลับ

2.10 นิวรอลเน็ตเวิร์กคอนโวลูชัน (Convolutional Neural Network : CNN or ConvNet)

ในการเรียนรู้เชิงลึกนั้น มีประเภทเฉพาะของโครงข่ายประสาทเทียมที่ใช้กันอย่างกว้างขวาง นิวรอลเน็ตเวิร์กคอนโวลูชันนั้นมีความคล้ายคลึงกับโครงข่ายประสาทเทียมทั่วไป แต่โครงสร้างได้ถูกออกแบบมาโดยเฉพาะสำหรับงานที่ใช้ภาพเป็นข้อมูลขาเข้า โดยเฉพาะ ชั้นของนิวรอลเน็ตเวิร์กคอนโวลูชันแตกต่างจากโครงข่ายประสาทเทียมทั่วไป คือ ชั้นมีเซลล์ประสาทที่จัดอยู่ในรูปแบบ 3 มิติ ได้แก่ ความกว้าง ความสูง และความลึก ซึ่งมีความเหมาะสมสำหรับข้อมูลเชิงพื้นที่ การรู้จำวัตถุ และการวิเคราะห์ภาพโดยใช้โครงสร้างเซลล์ประสาทแบบหลายมิติ (Multidimensional neurons structures) ตัวอย่างการนำนิวรอลเน็ตเวิร์กคอนโวลูชันไปใช้ เช่น ระบบขับเคลื่อนรถยนต์ด้วยตัวเอง งานด้านคอมพิวเตอร์วิทัศน์ หรือการวิเคราะห์ข้อมูล เป็นต้น

นิวรอลเน็ตเวิร์กคอนโวลูชันถูกใช้ในการประเมินข้อมูลขาเข้าที่ผ่านชั้นคอนโวลูชัน ในกรณีของภาพจากแบบจำลองอาร์จีบีสามารถแยกออกได้เป็น 3 สีพื้นฐาน คือ สีแดง สีเขียว และสีน้ำเงิน

ซึ่งในการแยกของแบบจำลองสี่ประเภทนี้จะมีความลึก (Depth) เพิ่มเข้ามาในข้อมูลภาพทำให้ข้อมูลนำเข้าของภาพจากแบบจำลองอาร์จีบีเป็นข้อมูลภาพแบบ 3 มิติ จากนั้นข้อมูลขาเข้าถูกแปลงด้วยตัวกรอง (Filter) เพื่อสกัดคุณลักษณะของชั้นคอนโวลูชันจากข้อมูลขาเข้าออกมา ซึ่งขั้นตอนนี้ทำให้โครงข่ายสามารถตรวจหาขอบและคุณลักษณะระดับล่าง (Low-level features) ในขั้นก่อนหน้านี้ รวมถึงตรวจจับคุณลักษณะที่ซับซ้อนของชั้นที่ลึกขึ้นในโครงข่าย



รูปที่ 2-18 ตัวอย่างโครงสร้างของนิรอลเน็ตเวิร์กคอนโวลูชัน โดยตัวกรองถูกนำไปใช้ในการสอนที่ความละเอียดต่างกันและผลลัพธ์ของแต่ละภาพจากชั้นคอนโวลูชันจะถูกนำไปใช้เป็นข้อมูลนำเข้าของชั้นถัดไป [12]

โครงสร้างของนิรอลเน็ตเวิร์กคอนโวลูชันประกอบด้วยชั้นที่ดำเนินการเปลี่ยนแปลงข้อมูล โดยเรียนรู้จากคุณลักษณะของข้อมูลขาเข้า ชั้นของนิรอลเน็ตเวิร์กคอนโวลูชันที่พบมากที่สุดประกอบด้วยชั้นคอนโวลูชันตามด้วยชั้นพูลลิงและชั้นสุดท้ายเป็นชั้นเชื่อมต่อแบบทั่วถึงดังรูปที่ 2-18 ซึ่งเป็นชั้นที่แปลงผลลัพธ์จากการทำคอนโวลูชันและทำหน้าที่ในการสร้างผลลัพธ์สุดท้ายของนิรอลเน็ตเวิร์กคอนโวลูชัน โดยแต่ละชั้นมีรายละเอียด ดังนี้

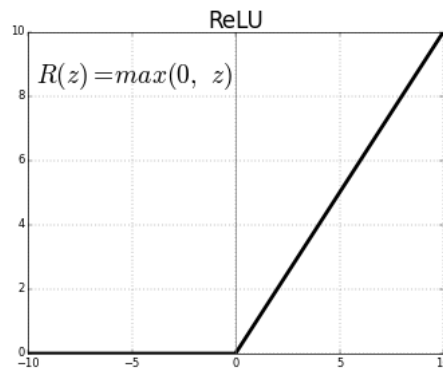
1) ชั้นข้อมูลขาเข้า (Input layer) มักใช้กับข้อมูลภาพ ซึ่งทุกภาพเป็นเมทริกซ์ของค่าจุดภาพ โดยช่วงของค่าจุดภาพที่สามารถเข้ารหัสได้ในแต่ละจุดภาพขึ้นอยู่กับขนาดบิตของภาพนั้น ๆ โดยทั่วไปภาพจะมีขนาด 8 บิตหรือ 1 ไบต์ ดังนั้นช่วงที่เป็นไปได้ของค่าจุดภาพสามารถแสดงได้ตั้งแต่ค่า 0 – 255 อย่างไรก็ตามในภาพสีโดยเฉพาะแบบจำลอง RGB จะดำเนินการแยกช่องสีออกเป็น 3 ช่อง โดยแต่ละช่องจะมีเขตข้อมูลความลึกเพิ่มเข้ามาทำให้ข้อมูลขาเข้าเป็นข้อมูล 3 มิติ ประกอบด้วย ความกว้าง x ความสูง x ความลึก ของแต่ละช่องสีในภาพ ยกตัวอย่างสำหรับภาพ RGB ขนาด 255 x

255 (ความกว้าง \times ความสูง) จุดภาพ ชั้นข้อมูลขาเข้าจะมี 3 เมทริกซ์ที่เชื่อมโยงกับภาพ แต่ละเมทริกซ์แสดงถึงแต่ละช่องสี ดังนั้นภาพ 1 ภาพจึงประกอบด้วยโครงสร้าง 3 มิติที่เรียกว่าปริมาตรข้อมูลขาเข้าซึ่งมีค่าเท่ากับ $255 \times 255 \times 3$

ตัวกรอง (Filters) หรือแกนกลาง (Kernels) แสดงด้วยเวกเตอร์ของค่าน้ำหนักที่ใช้ในการรวมกับข้อมูลขาเข้า โดยสามารถเพิ่มจำนวนตัวกรองของปริมาตรข้อมูลขาเข้าเพื่อเพิ่มจำนวนพีเจอร์แมพ (Feature map) ซึ่งแต่ละพีเจอร์แมพจะพยายามเรียนรู้มุมมองที่ต่างกันในภาพ เช่น ขอบภาพหรือจุดสี เป็นต้น ตัวกรองหลายตัวถูกใช้ในการสกัดคุณลักษณะในชั้นคอนโวลูชันของโครงข่ายประสาทเทียม และมีการก้าวข้าม (Stride) คอยควบคุมตัวกรองที่ดำเนินการรอบปริมาตรข้อมูลขาเข้า

2) ชั้นคอนโวลูชัน (Convolution layer) ประกอบด้วยตัวกรองที่สามารถเรียนรู้ได้ โดยตัวกรองมีลักษณะเป็นเมทริกซ์ประกอบกัน ได้แก่ ความกว้าง ความยาว และความลึก จากนั้นตัวกรองจะเลื่อนผ่านเมทริกซ์ของภาพไปเรื่อย ๆ จนได้ภาพคอนโวลูชันที่เป็นภาพจากตัวกรองของภาพขาเข้า ในแบบจำลองสามารถมีตัวกรองได้มากกว่า 1 ตัวกรองเพื่อรองรับคุณลักษณะของภาพขาเข้าที่แตกต่างกันโดยที่ค่าเมทริกซ์ของตัวกรองจะมีการเรียนรู้ในระหว่างขั้นตอนการสอนของแบบจำลองอีกด้วย

3) ชั้นกระตุ้น (Activation layer) เป็นฟังก์ชันที่กำหนดให้ข้อมูลขาเข้าส่งต่อไปยังโหนดและกำหนดผลลัพธ์ของโหนดนั้น ๆ เนื่องจากเป็นฟังก์ชันที่ตัดสินใจผลลัพธ์ที่แท้จริง จึงเรียกผลลัพธ์ของชั้นนี้ว่าการกระตุ้น (Activation) โดยฟังก์ชันการกระตุ้นที่เป็นที่นิยมใช้ในคอนโวลูชัน คือ ฟังก์ชันเรกติไฟด์เชิงเส้น (Rectified linear unit : ReLU) ที่ช่วยลดจำนวนตัวแปรในโครงข่ายซึ่งส่งผลทำให้โครงข่ายเรียนรู้ได้เร็วและมีประสิทธิภาพมากยิ่งขึ้นโดยการเปลี่ยนค่าลบจากผลลัพธ์ให้เป็น 0 และคงค่าที่เป็นบวกไว้ดังรูปที่ 2-19 และสมการที่ 3 และจะมีเพียงผลลัพธ์ของคุณลักษณะเท่านั้นที่ถูกส่งต่อไปยังขั้นถัดไป



รูปที่ 2-19 กราฟแสดงสมการฟังก์ชันเรกติไฟต์เชิงเส้น [35]

$$R(z) = \max(0, z) \quad (3)$$

โดยที่

$$\text{ผลลัพธ์เป็นบวกก็ต่อเมื่อ } R(1) = \max(0, 1) = 1$$

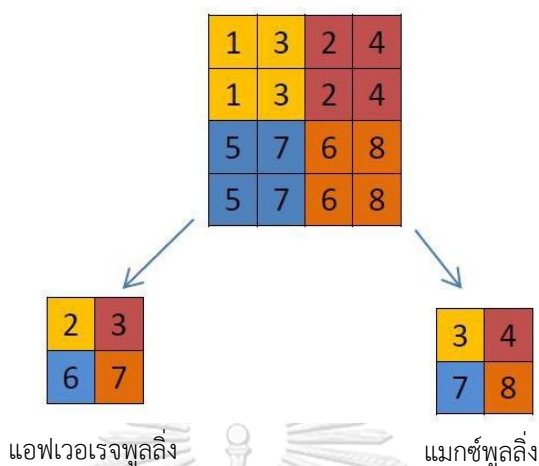
$$\text{ผลลัพธ์เป็นลบก็ต่อเมื่อ } R(-1) = \max(0, -1) = 0$$

นอกจากนี้ฟังก์ชันค่าสูงสุดอย่างอ่อน (Softmax function) เป็นฟังก์ชันการถดถอยโลจิสติกที่ใช้สำหรับการจำแนกประเภทหลายชั้นดังสมการที่ 4 โดยฟังก์ชันค่าสูงสุดอย่างอ่อนถูกใช้ในชั้นสุดท้ายของโครงข่ายเพื่อใช้คาดการณ์และใช้ความน่าจะเป็นสูงสุดจากข้อมูลประเภท

$$f(z)_j = \frac{e^{z_j}}{\sum_{i=1}^K e^{z_i}} \quad (4)$$

4) ชั้นพูลลิ่ง (Pooling layer) เป็นชั้นที่ช่วยลดขนาดเฉพาะความกว้างและความสูงของภาพ แต่ความลึกยังคงค่าไว้เหมือนเดิม ส่วนใหญ่พูลลิ่งประกอบด้วย 3 ประเภท ได้แก่ แมกซ์พูลลิ่ง (Max pooling) มินพูลลิ่ง (Min pooling) และแอฟเวอเรจพูลลิ่ง (Average pooling) โดยชั้นพูลลิ่งดำเนินการกับแต่ละชั้นของความลึกได้อย่างอิสระและปรับขนาดของภาพดังตัวอย่างรูปที่ 2-20 โดยใช้การดำเนินการทางคณิตศาสตร์ในแต่ละประเภทของพูลลิ่งดังที่ได้ระบุไว้ก่อนหน้านี้ ชั้นพูลลิ่งมี

ประโยชน์ในการลดความยุ่งยากของผลลัพธ์โดยการสุ่มตัวอย่างแบบไม่เชิงเส้นและลดจำนวนของตัวแปรที่โครงข่ายต้องการเรียนรู้



รูปที่ 2-20 ตัวอย่างการทำแอฟเวอเรจพูลลิ่งและแมกซ์พูลลิ่ง [13]

5) ชั้นเชื่อมต่อแบบทั่วถึง (Fully connected layer) เป็นชั้นที่ดำเนินการหลังจากที่คุณลักษณะในหลายชั้นของนิวรอลเน็ตเวิร์กคอนโวลูชันเปลี่ยนไปสู่การจำแนก เซลล์ประสาทในชั้นนี้มีการเชื่อมต่อการกระตุ้นทั้งหมดในชั้นก่อนหน้าและสามารถคำนวณการกระตุ้นเหล่านี้ด้วยการคูณเมทริกซ์ด้วยค่าไบแอส (Bias) ในชั้นนี้จะแสดงเวกเตอร์ผลลัพธ์ของ K มิติ โดยที่ K เป็นจำนวนประเภทที่โครงข่ายต้องการที่จะคาดการณ์ผลลัพธ์ เวกเตอร์ที่กล่าวถึงนี้ประกอบไปด้วยแต่ละประเภทของภาพที่กำลังจะถูกรวบรวมอยู่ในรูปแบบของกลุ่มประเภท

การดำเนินการของชั้นต่าง ๆ ของนิวรอลเน็ตเวิร์กคอนโวลูชันจะเกิดขึ้นแบบวนซ้ำโดยแต่ละชั้นจะเรียนรู้คุณลักษณะต่าง ๆ ที่ต้องการค้นหาจากข้อมูลนำเข้า ข้อดีของการใช้โครงข่ายประเภทนี้คือการแบ่งปันและการเชื่อมโยงตัวแปร ซึ่งการแบ่งปันตัวแปรสามารถช่วยลดจำนวนของตัวแปรค่าน้ำหนักให้เหลือเพียง 1 ชั้นโดยที่ไม่ส่งผลกระทบต่อความแม่นยำของผลลัพธ์ที่จะคาดการณ์ นอกจากนี้คอนโวลูชันยังแบ่งคุณลักษณะข้อมูลขาเข้าให้เป็นคุณลักษณะที่มีขนาดเล็กลง ทำให้แต่ละค่าผลลัพธ์ขึ้นอยู่กับจำนวนข้อมูลขาเข้าขนาดเล็ก ส่งผลให้สามารถปรับค่าได้อย่างรวดเร็ว

2.11 การรู้จำภาพ (Image Recognition)

การรู้จำภาพเป็นกระบวนการในการระบุและตรวจจับวัตถุหรือลักษณะเฉพาะที่ต้องการในภาพนั้น ๆ และเป็นเรื่องพื้นฐานสำหรับการประมวลผลภาพ โดยมีจุดประสงค์เพื่อช่วยระบุวัตถุใน

ภาพว่าเป็นภาพอะไร เช่น การรู้จำสัญลักษณ์ป้ายจราจรบนท้องถนน การรู้จำใบหน้ามนุษย์ หรือการรู้จำป้ายทะเบียนรถยนต์ เป็นต้น การรู้จำภาพถูกนำไปใช้ในการแก้ปัญหาทางด้านคอมพิวเตอร์วิทัศน์ (Computer vision) ซึ่งการรู้จำภาพสามารถแตกแขนงออกได้ 4 วิธี ดังนี้

2.11.1 การจำแนกประเภทข้อมูลภาพ และการจัดกลุ่มภาพ (Image classification and image clustering)

การจำแนกประเภทข้อมูลภาพเป็นปัญหาทั่วไปที่สามารถพบในงานด้านคอมพิวเตอร์วิทัศน์ กล่าวคือเป็นการจำแนกประเภทของภาพออกเป็นหลาย ๆ กลุ่ม ซึ่งกลุ่มที่ถูกจำแนกออกนี้เรียกว่ากลุ่มข้อมูลประเภท (Class) โดยที่แต่ละกลุ่มจะต้องมีลักษณะเฉพาะที่เหมือนกัน หรือมีคุณสมบัติที่คล้ายคลึงกัน และแยกกลุ่มข้อมูลประเภทที่มีลักษณะเฉพาะที่แตกต่างให้ออกจากกัน

2.11.2 การหาตำแหน่ง (Localization)

การหาตำแหน่งมีความหมายที่เหมือนกันกับการจำแนกประเภทข้อมูลและการจัดกลุ่มในข้อ 2.7.2 ซึ่งเป็นการหาตำแหน่งของคุณลักษณะหรือวัตถุที่อยู่ในภาพ โดยการหาตำแหน่งเป็นวิธีที่ใช้ประยุกต์เข้ากับการจำแนกประเภทข้อมูลภาพเพื่อจัดกลุ่มคุณลักษณะหรือวัตถุในภาพออกเป็นกลุ่มหลายกลุ่มตามประเภทข้อมูลภาพ ซึ่งผลลัพธ์ของการหาตำแหน่งสามารถระบุป้ายชื่อวัตถุว่าอยู่ในกลุ่มข้อมูลประเภทใด รวมถึงตีกรอบสี่เหลี่ยมในการปิดล้อมวัตถุของวัตถุในภาพอีกด้วย

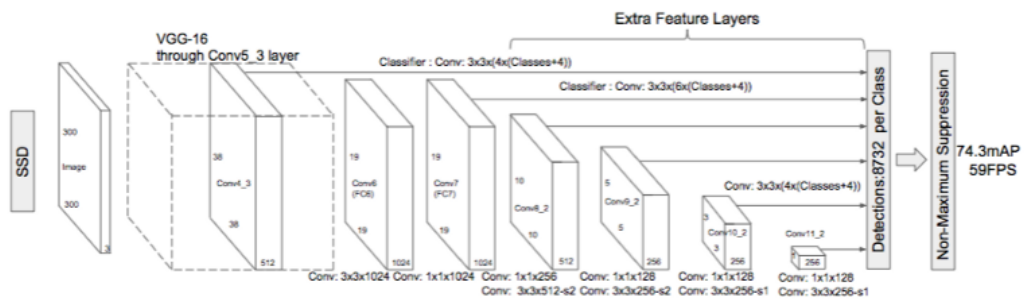
2.11.3 การตรวจจับวัตถุ (Object Detection)

การตรวจจับวัตถุเป็นการจัดการปัญหาของการหาตำแหน่งคุณลักษณะหรือวัตถุในแต่ละกลุ่มข้อมูลประเภท ซึ่งใช้หลักการค้นหาและจัดกลุ่มโดยใช้ค่าตัวแปรของวัตถุหรือคุณลักษณะในภาพ ความแตกต่างที่ชัดเจนของการตรวจจับวัตถุและวิธีทั้ง 2 วิธีที่กล่าวมาก่อนหน้านี้ คือ การตรวจจับวัตถุจะใช้ค่าตัวแปร ซึ่งเป็นค่าความแม่นยำ (Accuracy) ในการตำแหน่งและระบุวัตถุหลาย ๆ วัตถุในภาพเดียวกัน ในการตรวจจับวัตถุมีหลายโครงสร้าง (Framework) ที่ถูกนำมาใช้ สำหรับในวิทยานิพนธ์ฉบับนี้ผู้วิจัยใช้เทคนิคการเรียนรู้แบบเชิงลึกด้วยโครงสร้างซึ่งเกิดชอตทีเทคเตอร์ (Single Shot Detector : SSD) [14] และการสกัดคุณลักษณะจากแบบจำลองโมบายเน็ต (MobileNet) [15] เนื่องจากเป็นสถาปัตยกรรมที่ทำงานได้อย่างรวดเร็วและมีประสิทธิภาพโดยไม่ต้องใช้ความสามารถด้านการคำนวณของหน่วยประมวลผลในการตรวจจับวัตถุมากนัก ซึ่ง SSD จะกำหนดความน่าจะเป็นของสี่เหลี่ยมในการปิดล้อมวัตถุทั้งหมดในการประมวลผลครั้งเดียวทำให้การทำงานนั้นเร็วขึ้นกว่าโครงสร้างอื่น ๆ และแบบจำลอง MobileNet ถูกนำมาใช้เป็นตัวสกัดคุณลักษณะสำหรับ

สถาปัตยกรรมโครงข่ายนิวรอลเน็ตเวิร์ก ซึ่งถูกออกแบบมาเพื่อใช้ในแอปพลิเคชันมือถือ โดยโครงสร้าง SSD มีหลักการทำงาน ดังนี้

ปัจจุบันโครงสร้างซิงเกิลชอตดีเทคเตอร์ (Single Shot Detector : SSD) เป็นหนึ่งในโครงสร้างที่ถูกใช้ในการตรวจจับวัตถุ โดยมีค่าความแม่นยำสูง พัฒนาโดย W.Liu และคณะ [14] เพื่อทำนายทุกกรอบสี่เหลี่ยมในการปิดล้อมวัตถุ (Bounding boxes) และความน่าจะเป็นของข้อมูลประเภทด้วยสถาปัตยกรรมคอนโวลูชันแบบปลายทางถึงปลายทาง โครงสร้าง SSD ใช้ภาพนำเข้าที่ผ่านชั้นคอนโวลูชันหลายชั้นพร้อมตัวกรองขนาดต่าง ๆ ได้แก่ 10×10 , 5×5 และ 3×3 จากนั้นการกระจายคุณลักษณะจากชั้นคอนโวลูชันที่ตำแหน่งต่าง ๆ ของโครงข่ายถูกใช้เพื่อทำนายขอบเขตของกรอบสี่เหลี่ยมในการปิดล้อมวัตถุ โดยกรอบสี่เหลี่ยมเหล่านี้ถูกสร้างขึ้นโดยชั้นคุณลักษณะพิเศษ (Extra feature layers) ซึ่งประมวลผลโดยชั้นคอนโวลูชันกับตัวกรองแบบ 3×3 ซึ่งแต่ละกรอบสี่เหลี่ยมประกอบด้วย 4 ตัวแปร ได้แก่ พิกัดของจุดศูนย์กลาง ความกว้าง ความสูง และเวกเตอร์ที่สอดคล้องกับความเชื่อมั่นในแต่ละข้อมูลประเภทของวัตถุ

นอกจากนี้วิธีการขจัดค่าที่ไม่มากที่สุด (Non-maxima suppression) ถูกนำมาใช้ที่ส่วนท้ายของโครงสร้าง SSD เพื่อให้มีสี่เหลี่ยมในการปิดล้อมวัตถุที่เกี่ยวข้องมากที่สุดและจัดกลุ่มสี่เหลี่ยมในการปิดล้อมวัตถุที่ซ้อนทับกันให้เหลือเพียงสี่เหลี่ยมชิ้นเดียว วิธีการขจัดค่าที่ไม่มากที่สุดจะเก็บกรอบสี่เหลี่ยมในการปิดล้อมวัตถุที่มีค่าความเชื่อมั่นสูงสุดและละทิ้งส่วนที่เหลือ จากนั้นโครงสร้าง SSD ใช้เทคนิคการทำเหมืองข้อมูลเชิงลบอย่างหนัก (Hard Negative Mining : HNM) เพื่อสร้างสมดุลของข้อมูลประเภทระหว่างการสอน โดยชุดย่อยของตัวอย่างเชิงลบที่มีค่าสูญเสียการสอน (Training loss) สูงสุดจะถูกใช้ในแต่ละรอบของการสอน และโครงสร้าง SSD จะควบคุมอัตราส่วนของค่าเชิงลบต่อค่าเชิงบวกมีค่าเท่ากับ 3 ต่อ 1

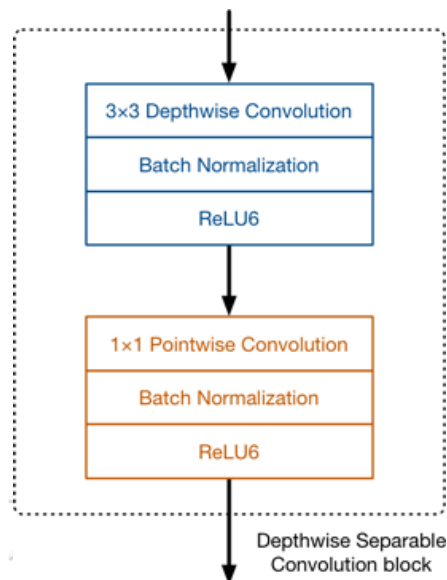


รูปที่ 2-21 สถาปัตยกรรมของโครงสร้าง SSD โดยมีภาพนำเข้าขนาด 300 x 300 จุดภาพ [36]

จากรูปที่ 2-21 แสดงสถาปัตยกรรมของโครงสร้าง SSD ถูกสร้างขึ้นจากสถาปัตยกรรม VGG16 [16] แต่ละทิ้งชั้นเชื่อมต่อแบบทั่วถึง โดยเหตุผลที่นำสถาปัตยกรรม VGG16 มาใช้เป็นโครงข่ายพื้นฐานนั้น เนื่องจากมีประสิทธิภาพที่ดีในงานจำแนกภาพที่มีคุณภาพสูงและเป็นโครงข่ายที่ได้รับความนิยมสำหรับปัญหาที่ใช้เทคนิคการถ่ายโยงการเรียนรู้ช่วยในการปรับปรุงผลลัพธ์ จากเดิมปกติเป็นชั้นเชื่อมต่อแบบทั่วถึงในโครงข่าย VGG ได้ถูกแทนที่ด้วยชั้นคอนโวลูชันเสริมจากชั้น conv6 เป็นต้นไป ด้วยเหตุนี้จึงช่วยให้สามารถสกัดคุณลักษณะที่มีขนาดหลายขนาดได้และยังค่อย ๆ ลดขนาดของข้อมูลนำเข้าในแต่ละชั้นตามลำดับได้ การทำงานของโครงสร้าง SSD เริ่มจากใช้โครงข่ายขอบเขตที่ต้องการเพื่อสร้างพื้นที่ที่สนใจ จากนั้นชั้นคอนโวลูชันที่ไวต่อตำแหน่ง (Position-sensitive convolution layers) ถูกนำมาเพื่อจำแนกพื้นที่เหล่านั้น โดยโครงสร้าง SSD ดำเนินการทั้ง 2 ชั้นตอนนี้พร้อมกันในครั้งเดียวซึ่งให้ผลลัพธ์สุดท้ายอยู่ในรูปแบบของการทำนายกรอบสี่เหลี่ยมในการปิดล้อมวัตถุและประเภทข้อมูลทั้งหมดที่ปรากฏอยู่ในภาพนำเข้า

โมบายเน็ต (MobileNet) [15] เป็นตัวสกัดคุณลักษณะที่ออกแบบมาเพื่อประมวลผลบนอุปกรณ์เคลื่อนที่ซึ่งใช้เพื่อรองรับกับคุณลักษณะระดับสูงของภาพ โดย MobileNet พยายามที่จะลดความต้องการในการคำนวณของอัลกอริทึมการเรียนรู้เชิงลึกเพื่อให้รองรับกับการ์ดแสดงผลภาพของอุปกรณ์เคลื่อนที่ซึ่งมีอยู่ในสมาร์ทโฟนจำนวนมาก รวมถึงอุปกรณ์อิเล็กทรอนิกส์พกพาแบบอื่น ๆ เช่น ทีวีโอเกมซึ่งมีขนาดหน่วยความจำของการ์ดแสดงผลเพียงจำนวนเล็กน้อยเมื่อเทียบกับหน่วยความจำบนเครื่องคอมพิวเตอร์ ความแตกต่างระหว่างสถาปัตยกรรมของ MobileNet และนิวโรลเน็ตเวิร์กคอนโวลูชัน คือ แทนที่จะเป็นชั้นคอนโวลูชันขนาด 3 x 3 เดียวตามด้วยกลุ่มนอร์มัลไลเซชัน (Batch normalization) และฟังก์ชัน ReLU แต่ในทางกลับกัน MobileNet แบ่งคอนโวลูชันออกเป็นเดฟไวส์คอนโวลูชัน (Depthwise) ขนาด 3 x 3 และพอยน์ทไวส์คอนโวลูชัน (Pointwise) ขนาด 1 x 1 ที่ทำหน้าที่ในการรวมค่าตัวกรองเหล่านี้เพื่อสร้างคุณลักษณะใหม่ โดยทั้งคอนโวลูชัน depthwise และ

pointwise สามารถเรียกได้ว่าคอนโวลูชันแบบเดฟไวส์เซพเรเบิล (Depthwise separable) ซึ่งคล้ายคลึงกับนิวรอลเน็ตเวิร์กคอนโวลูชันแบบดั้งเดิมแต่ด้วยโครงสร้างของ MobileNet ส่งผลให้มีความเร็วกว่าเดิมมาก



รูปที่ 2-22 ชั้นของคอนโวลูชันแบบ Depthwise separable [17]

สถาปัตยกรรมของ MobileNet ดังรูปที่ 2-22 ประกอบด้วยคอนโวลูชันขนาด 3×3 อยู่ที่ชั้นแรกตามด้วยกลุ่มของคอนโวลูชันแบบ Depthwise separable จำนวน 13 กลุ่มดังรูปที่ 2-23 ซึ่งในระหว่างกลุ่มชั้นของ Depthwise separable จะไม่มีชั้นพูลลิ่งอยู่ แต่ในบางชั้น Depthwise มีการก้าวข้ามที่ 2 เพื่อลดมิติเชิงพื้นที่ของข้อมูล และเมื่อเกิดเหตุการณ์นี้ขึ้น ชั้น Pointwise ที่สอดคล้องกันจะเพิ่มจำนวนช่องของผลลัพธ์ออกเป็น 2 เท่า โดยปกตินิวรอลเน็ตเวิร์กคอนโวลูชันใช้ฟังก์ชันกระตุ้น ReLU แต่สำหรับ MobileNet จะใช้ฟังก์ชัน ReLU6 แทนที่ ซึ่งมีลักษณะเหมือนกับ ReLU แต่จะช่วยป้องกันไม่ให้การกระตุ้นนั้นมีขนาดใหญ่จนเกินไป สำหรับตัวจำแนกประเภทบน MobileNet ประกอบด้วยชั้นแอพเวอเรจพูลลิ่งแบบครอบคลุมที่ส่วนท้ายตามด้วยชั้นจำแนกประเภทเชื่อมต่อแบบทั่วถึงหรือเทียบเท่ากับคอนโวลูชันขนาด 1×1 และฟังก์ชันค่าสูงสุดอย่างอ่อน

2.11.4 การจำแนกทางความหมาย (Semantic Segmentation)

การจำแนกทางความหมาย คือ การกำหนดจุดภาพแต่ละจุดภาพในภาพนำเข้าให้เป็นข้อมูลประเภทที่มีความหมายเพื่อนำไปจำแนกกลุ่มความหนาแน่นของจุดภาพในแต่ละกลุ่มดังรูปที่ 2-23 โดยการจำแนกทางความหมายเป็นส่วนหนึ่งของงานด้านคอมพิวเตอร์วิทัศน์ตั้งแต่ปีพ.ศ. 2550 และมี

การเปลี่ยนแปลงและพัฒนาที่สำคัญเกิดขึ้นเมื่อมีการนำโครงข่ายคอนโวลูชันแบบทั่วถึง (Fully Convolutional Networks : FCNs) มาใช้ครั้งแรกจากงานวิจัยของ Long และคณะ [18] โดยมีจุดประสงค์เพื่อทำการแบ่งส่วนแบบปลายทางถึงปลายทาง (End-to-end segmentation) ซึ่งการจำแนกทางความหมายเป็นขั้นตอนตามธรรมชาติในการพัฒนาจากการอนุมานที่หายวับไปจนถึงการอนุมานที่ดี โดยเริ่มจากการจำแนกประเภทซึ่งประกอบด้วยการทำงานจากข้อมูลขาเข้าทั้งหมด ขั้นตอนถัดมา คือ การหาตำแหน่งหรือการตรวจจับซึ่งไม่เพียงจำแนกประเภทเท่านั้น แต่ยังเพิ่มข้อมูลตำแหน่งของแต่ละวัตถุที่อยู่ในข้อมูลประเภทเหล่านั้นด้วย ขั้นตอนสุดท้ายเป็นการจำแนกทางความหมายที่จะรับการอนุมานที่ดีโดยใช้การคาดการณ์แบบหนาแน่นเพื่อติดป้ายให้กับทุก ๆ จุดภาพ เพื่อให้แต่ละจุดภาพมีการกำกับป้ายชื่อของแต่ละประเภทล้อมรอบวัตถุนั้น ๆ



(ก)

(ข)

รูปที่ 2-23 ตัวอย่างการจำแนกทางความหมายของภาพ (ก) ภาพนำเข้า (ข) ภาพผลลัพธ์ของการจำแนกทางความหมาย [19]

โครงสร้างของการจำแนกทางความหมายทั่วไปสามารถใช้โครงข่ายเข้ารหัส (Encoder) ตามด้วยโครงข่ายถอดรหัส (Decoder) สำหรับตัวเข้ารหัสมักจะเป็นโครงข่ายการจำแนกที่ถูกฝึกไว้แล้ว เช่น โครงข่ายวีจีจี (VGG) [16] หรือเรสเน็ต (ResNet) [20] เป็นต้น ตามด้วยโครงข่ายถอดรหัส ส่วนทำงานของตัวถอดรหัสเป็นการสร้างความแตกต่างของคุณลักษณะ (ความละเอียดต่ำ) จากตัวเข้ารหัสบนพื้นที่ของจุดภาพ (ความละเอียดสูง) เพื่อจำแนกประเภทของจุดภาพที่หนาแน่นออกมา แตกต่างจากการจำแนกประเภทที่ผลลัพธ์สุดท้ายของโครงข่ายเชิงลึกเป็นสิ่งสำคัญเท่านั้น ส่วนการจำแนกทางความหมายไม่เพียงแค่แบ่งข้อมูลภาพที่ระดับจุดภาพเท่านั้น แต่ยังมีกระบวนการในการคัดเลือกคุณลักษณะในขั้นตอนต่าง ๆ ของตัวเข้ารหัสบนพื้นที่ของจุดภาพอีกด้วย สำหรับวิธีในการถอดรหัสจะ

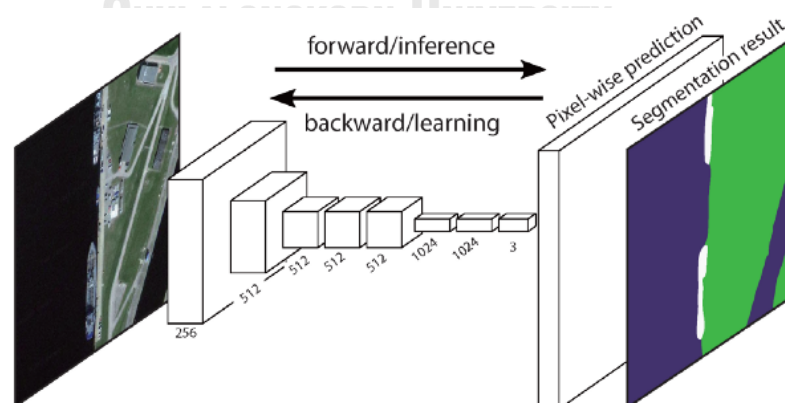
ใช้เทคนิคที่แตกต่างกัน ในงานวิจัยนี้ผู้วิจัยใช้วิธีการจำแนกทางความหมายตามโครงข่ายคอนโวลูชันแบบทั่วถึง (Fully convolutional network-based semantic segmentation) ซึ่งมีคุณลักษณะสำคัญของโครงข่าย ดังนี้

1) การสุ่มตัวอย่างขึ้น (Upsampling) ของคุณลักษณะจากชั้นตอนต่าง ๆ ในตัวแปลงรหัสถูกรวมเข้าด้วยกันซึ่งจะมีความแตกต่างกันไปตามความหมายของข้อมูลความหมาย

2) การเรียนรู้คุณลักษณะที่ความละเอียดต่ำเพื่อหาประเภทของทุก ๆ จุดภาพในภาพทำได้โดยใช้ดีคอนโวลูชัน (Deconvolution) ที่เริ่มต้นตัวกรองแบบการสุ่มระดับสี่ (Bilinear interpolation)

3) ตัวอย่างของการถ่ายโยงการเรียนรู้จากโครงข่ายจำแนกประเภทที่ทันสมัย ได้แก่ โครงข่ายวีจีจี 16 (VGG16) [16] หรือโครงข่ายอเล็กซ์เน็ต (Alexnet) [21] เป็นต้น

โครงข่ายคอนโวลูชันแบบทั่วถึงจะเรียนรู้ด้วยวิธีการค้นหาจุดภาพแต่ละจุดภาพโดยไม่ต้องใช้กระบวนการสกัดบริเวณที่ต้องการ สำหรับการดำเนินงานแบบสายท่อของโครงข่ายคอนโวลูชันแบบทั่วถึงเป็นส่วนขยายของคอนโวลูชันแบบเดิม ซึ่งมีแนวคิดหลักเพื่อปรับคอนโวลูชันแบบเดิมให้รองรับกับภาพนำเข้าที่มีขนาดไม่เจาะจงดังรูปที่ 2-24 ข้อจำกัดของคอนโวลูชันแบบเดิมนั้นยอมรับและตัดป้ายเฉพาะภาพนำเข้าที่มีขนาดเฉพาะเจาะจงจากชั้นเชื่อมต่อแบบทั่วถึงเท่านั้น ตรงกันข้ามกับโครงข่ายคอนโวลูชันแบบทั่วถึงที่มีชั้นคอนโวลูชันและชั้นพูลลิ่งซึ่งทำให้มีความสามารถในการคาดการณ์ภาพนำเข้าที่มีขนาดไม่เจาะจงได้



รูปที่ 2-24 โครงสร้างทำนายความหนาแน่นของโครงข่ายคอนโวลูชันแบบทั่วถึงแบบปลายทางถึงปลายทาง [18]

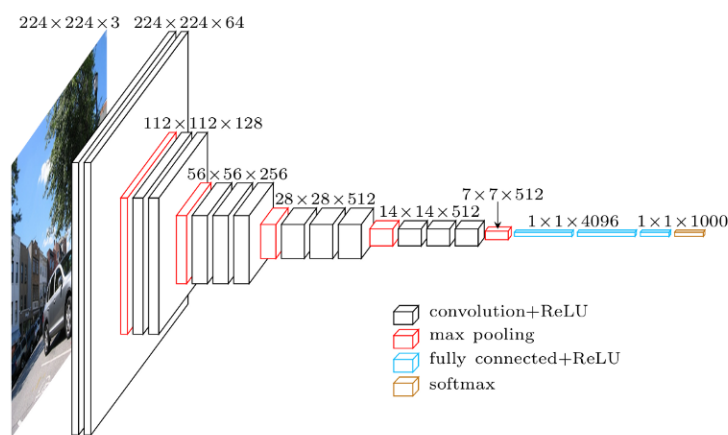
จุดมุ่งหมายสุดท้ายของโครงข่ายคอนโวลูชันแบบทั่วถึง คือ การจำแนกทางความหมายซึ่งภาพผลลัพธ์จะมีขนาดเดียวกับภาพนำเข้าและมีคุณสมบัติคล้ายกับภาพนำเข้าต้นฉบับ แตกต่างกันในแต่ละจุดภาพในภาพประกอบด้วยสี่ C ซึ่งคือจำนวนประเภทข้อมูลของภาพที่ต้องการจำแนก ตัวอย่างดังรูปที่ 2-23 มีค่า $C = 3$ ประกอบด้วยม้า คน และพื้นหลัง สำหรับแนวคิดพื้นฐานของโครงข่ายคอนโวลูชันแบบทั่วถึงเป็นการคอนโวลูชันแบบทั่วถึงซึ่งในทุกชั้นเป็นชั้นแบบคอนโวลูชันภายในโครงสร้างของโครงข่ายคอนโวลูชันแบบทั่วถึงไม่มีส่วนใดที่เชื่อมต่อกับชั้นเชื่อมต่อแบบทั่วถึงในส่วนท้าย ซึ่งมักใช้กับการจำแนกประเภท ในทางกลับกันโครงข่ายคอนโวลูชันแบบทั่วถึงใช้ชั้นคอนโวลูชันเพื่อจัดกลุ่มแต่ละจุดภาพในภาพ ดังนั้น ชั้นผลลัพธ์สุดท้ายจะมีความกว้างและความสูงเท่ากับภาพนำเข้า แต่จำนวนช่องจะเท่ากับจำนวนประเภทที่จำแนกได้ ยกตัวอย่างในกรณีที่ต้องการจำแนกแต่ละจุดภาพเป็น 1 ใน 15 กลุ่มประเภทที่แตกต่างกัน ดังนั้นชั้นผลลัพธ์สุดท้ายเท่ากับความกว้าง \times ความสูง $\times 15$ ประเภท โดยในชั้นสุดท้ายมีฟังก์ชันความน่าจะเป็นค่าสูงสุดอย่างอ่อนที่สามารถค้นหาข้อมูลประเภทที่น่าสนใจที่สุดสำหรับแต่ละจุดภาพ

การเรียนรู้ได้เองของการสุ่มตัวอย่างขึ้นในโครงข่ายคอนโวลูชันแบบทั่วถึงจะทำให้ชั้นกลางของโครงสร้างมีขนาดเล็กกลงโดยใช้พูลลิงในการลดความกว้างและความสูงของเทนเซอร์ (Tensors) และโครงข่ายนี้ใช้คอนโวลูชันหรือคอนโวลูชันแบบย้อนกลับ (Backwards convolution) เพื่อขยายเทนเซอร์ที่ชั้นกลางเพื่อให้ขนาดภาพทั้งความกว้างและความสูงตรงกับภาพต้นฉบับ เนื่องจากชั้นคอนโวลูชันนั้นเป็นการสลับคอนโวลูชันไปอีกด้านหนึ่งเท่านั้น เพราะฉะนั้นค่าน้ำหนักจึงสามารถเรียนรู้ได้เหมือนชั้นคอนโวลูชันปกติ

Long และคณะ [18] ประสบความสำเร็จในการแปลงโครงข่ายที่ได้รับการยอมรับ เช่น VGG [16] หรือ AlexNet [21] เป็นต้น ลงในโครงข่ายคอนโวลูชันแบบทั่วถึงโดยการแทนที่ชั้นสุดท้ายด้วยโครงข่ายดังกล่าว แต่เกิดปัญหา คือ การสุ่มตัวอย่างขึ้นจากเทนเซอร์คอนโวลูชันสุดท้ายที่ดูเหมือนจะไม่มีคามแม่นยำ เนื่องจากมีข้อมูลเชิงพื้นที่สูญหายไปจากการสุ่มตัวอย่างลงทั้งหมดในโครงข่าย ด้วยเหตุนี้ Long และคณะ [18] จึงรวมการสุ่มตัวอย่างขึ้นจากตัวกลางของเทนเซอร์สุดท้ายเข้ากับการสุ่มตัวอย่างขึ้นจากเทนเซอร์ก่อนหน้า เพื่อให้ได้ข้อมูลเชิงพื้นที่ที่แม่นยำมากขึ้น

ในงานวิจัยฉบับนี้ผู้วิจัยได้ใช้โครงข่าย VGG16 เป็นแบบจำลองของโครงข่ายคอนโวลูชันสำหรับการรู้จำภาพ ซึ่งได้รับการพัฒนาโดย Simonyan และ Zisserman [16] ในกลุ่ม Visual geometry จากมหาวิทยาลัยออกซ์ฟอร์ด โครงข่าย VGG นี้มีความแม่นยำในการทดสอบสูงติด 5

อันดับแรกด้วยความแม่นยำเท่ากับ 92.7% ในคลังข้อมูลภาพ ImageNet [22] ซึ่งเป็นชุดข้อมูลที่มีภาพมากกว่า 14 ล้านภาพใน 1,000 ชุดข้อมูลประเภท สถาปัตยกรรมของ VGG16 ประกอบด้วยชั้นคอนโวลูชันจำนวน 16 ชั้น มีคอนโวลูชันตัวกรองแบบ 3×3 กับการก้าวข้าม (Stride) ที่ 1 ในชั้นคอนโวลูชัน และใช้การเสริมเติม (Padding) เดียวกันกับชั้นพูลลิ่ง 2×2 กับการก้าวข้ามที่ 2 รวมถึงมีตัวกรองจำนวนมาก ดังรูปที่ 2-25 แสดงโครงสร้างของ VGG16 ซึ่งยกตัวอย่างชั้นภาพนำเข้าที่มีขนาด $224 \times 224 \times 3$ และใช้ชั้นคอนโวลูชันขนาด 3×3 ซ้อนกันบริเวณด้านบนบนของแต่ละชั้นในความลึกที่เพิ่มขึ้น ซึ่งมีชั้นผลลัพธ์เป็นการคาดการณ์ค่าสูงสุดอย่างอ่อนใน 1,000 ชุดข้อมูลประเภทและมีชั้นเชื่อมต่อแบบทั่วถึงจำนวน 2 ชั้น แต่ละชั้นมีจำนวน 4,096 โหนด ตามด้วยชั้นแมคพูลลิ่ง (Max pooling) ซึ่งมีหน้าที่ในการลดขนาดและถูกติดป้ายกำกับโดยมีขนาด $7 \times 7 \times 512$ ถือว่าเป็นส่วนของการสกัดคุณลักษณะของแบบจำลอง ในขณะที่ส่วนที่เหลือของโครงข่ายถือเป็นส่วนของการจำแนกประเภทข้อมูลของแบบจำลอง โครงข่าย VGG16 เป็นทางเลือกที่ดีที่สุดสำหรับงานในด้านการสกัดคุณลักษณะจากภาพ การกำหนดค่าน้ำหนักของโครงข่าย VGG นั้นสามารถเข้าถึงได้อย่างสาธารณะและมีการใช้งานอย่างแพร่หลายในหลายงานประยุกต์ รวมถึงเป็นพื้นฐานของตัวสกัดคุณลักษณะ

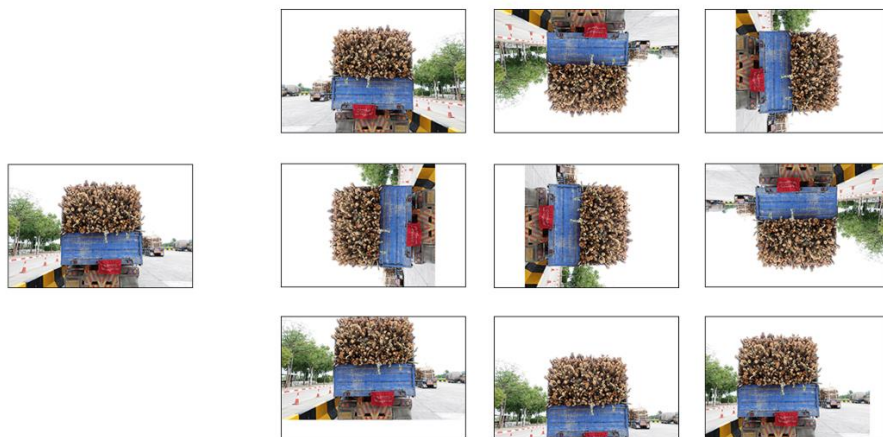


รูปที่ 2-25 สถาปัตยกรรมของโครงข่ายแบบ VGG [23]

2.12 การเพิ่มคุณค่าของข้อมูล (Data augmentation)

การเพิ่มคุณค่าของข้อมูลเป็นเทคนิคในการสร้างตัวอย่างสอนใหม่โดยใช้วิธีการสุ่มตัวอย่างจากภาพต้นฉบับที่ได้รับการแก้ไข เพื่อเพิ่มความแม่นยำในงานจำแนกประเภท [24] รวมถึงเป็นการเพิ่มชุดข้อมูลเพื่อเป็นข้อมูลนำเข้าสำหรับการเรียนรู้ของเครื่องโดยเฉพาะโครงข่ายนิรวลเน็ตเวิร์กและการเรียนรู้เชิงลึก เนื่องจากข้อมูลนำเข้าเป็นเหมือนกุญแจสำคัญในแบบจำลองของการเรียนรู้ด้วย

เครื่อง การเพิ่มคุณค่าของข้อมูลจึงถูกนำมาประยุกต์ใช้เพื่อลดการขาดหายและความหลากหลายของข้อมูล ซึ่งสามารถเพิ่มชุดข้อมูลสอนได้ 10 เท่าหรือมากกว่าดังรูปที่ 2-26 ยังมีข้อมูลสอนในแบบจำลองมากก็จะส่งผลให้แบบจำลองนั้นมีประสิทธิภาพมากขึ้นและป้องกันการเกิดโอเวอร์ฟิต (Overfitting) การที่ข้อมูลต้องมีความหลากหลายนั้นเป็นเรื่องที่ดีเนื่องจากวัตถุที่สนใจนั้นจะมีขนาดสภาพแสง และท่าทางที่แตกต่างกัน การเพิ่มคุณค่าของข้อมูลช่วยทำให้แบบจำลองมีคุณสมบัติทั่วไป (Generalizes) ที่สามารถนำไปประยุกต์ใช้ได้ดีในขั้นตอนการทดสอบ (Testing) หรือขั้นตอนการใช้งาน (Deployment)



รูปที่ 2-26 ตัวอย่างการเพิ่มคุณค่าของข้อมูลด้วยวิธีการแปลงข้อมูลภาพ

2.13 การวัดประสิทธิภาพของการจำแนกประเภท (Performance Evaluation of Classification)

การคำนวณหาค่า Precision Recall และ F-Measure คำนวณได้จาก

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$F - Measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (7)$$

โดยที่

TP (True Positive) คือ จำนวนหน้าตัดของปลายท่อนที่ทำนายว่ามีหน้าตัดของปลายท่อนและผลลัพธ์คือมีหน้าตัดของปลายท่อน

FP (False Positive) คือ จำนวนหน้าตัดของปลายท่อนที่ทำนายว่ามีหน้าตัดของปลายท่อนและผลลัพธ์คือไม่มีหน้าตัดของปลายท่อน

TN (True Negative) คือ จำนวนหน้าตัดของปลายท่อนที่ทำนายว่าไม่มีหน้าตัดของปลายท่อนและผลลัพธ์คือมีหน้าตัดของปลายท่อน

FN (False Negative) คือ จำนวนหน้าตัดของปลายท่อนที่ทำนายว่าไม่มีหน้าตัดของปลายท่อนและผลลัพธ์คือไม่มีหน้าตัดของปลายท่อนจริง ๆ

2.14 งานวิจัยที่เกี่ยวข้อง

งานวิจัยนี้มีจุดมุ่งหมายในการตรวจจับหน้าตัดของปลายท่อนซุง โดยใช้เทคนิคการประมวลผลจากภาพถ่ายด้านท้ายรถบรรทุกไม้อูคาลิปตัส ผู้วิจัยได้ทำการสำรวจงานวิจัยที่เกี่ยวข้องพบว่าม้งานวิจัยที่เกี่ยวกับการตรวจจับหน้าตัดของปลายท่อนซุงโดยใช้อัลกอริทึมบนภาพถ่ายค่อนข้างน้อยโดยเฉพาะอย่างยิ่งจากภาพถ่ายด้านท้ายรถบรรทุก อย่างไรก็ตามพบว่ามีเทคนิคการประมวลผลจากภาพถ่ายหลายเทคนิคถูกนำมาประยุกต์ใช้ในการตรวจจับหน้าตัดของปลายท่อนซุง โดยสามารถแบ่งออกได้เป็น 2 ประเภท คือ ประเภทที่ใช้ภาพถ่ายและประเภทที่ใช้แสงเลเซอร์

2.14.1 ประเภทที่ใช้การประมวลผลจากแสงเลเซอร์

ปี 2009 Nylinder และคณะ (Nylinder, Kubenka, and Hultnas, 2009) [25] ได้สำรวจวิธีการวัดท่อนซุงด้วยแสงเลเซอร์โดยใช้ระบบ 3D laser scanner ที่มีชื่อว่า Logmeter 4000 จากบริษัท Woodtech ในประเทศชิลี ได้มีการถูกนำไปใช้ในอุตสาหกรรมไม้หลายประเทศ ขั้นตอนเริ่มจากรถบรรทุกจะขับไปยังบริเวณสถานที่วัดด้วยแสงเลเซอร์ โดยอุปกรณ์แสงเลเซอร์จะถูกติดตั้งอยู่กับแท่นสถานี มีอยู่ด้วยกัน 3 ตำแหน่ง คือ บริเวณซ้าย ขวา และบนแท่นเหนือรถบรรทุก ซึ่งจำเป็นต้องขับผ่านอย่างช้า ๆ จะช่วยให้ประมวลผลได้แม่นยำมากขึ้น ระบบ Logmeter สามารถให้ข้อมูลไม้ที่สำคัญได้ เช่น ความยาว เส้นผ่านศูนย์กลาง และปริมาตรของท่อนซุง การตรวจจับและวัดขนาดไม้ด้วยแสงเลเซอร์เป็นเรื่องที่ไม่ลำบากนัก และสามารถดำเนินการได้อย่างอัตโนมัติ โดยมนุษย์จะมีส่วน

ร่วมกับกระบวนการนี้ไม่ถึง 5% และสามารถตรวจได้อย่างรวดเร็วภายใน 1 นาที ซึ่งรองรับรถบรรทุกได้สูงสุดมากถึง 600 คันต่อวัน เมื่อรถบรรทุกเคลื่อนไปยังบริเวณที่ติดตั้งเลเซอร์ไว้ เลเซอร์จะทำการกวาดไปทั่วทั้งคันเพื่อสร้างภาพตัดขวางและแบบจำลองมุมมอง 3 มิติ จากนั้นจะใช้เทคนิคการประมวลผลภาพถ่ายเพื่อกำจัดสิ่งที่ไม่จำเป็นออกไปจากภาพ เช่น ส่วนประกอบของรถบรรทุก หลังจากที่ได้ภาพตัดขวางแล้วมีการนำภาพไปวิเคราะห์ค้นหาเส้นผ่านศูนย์กลางของไม้ทั้งหมดเพื่อนำไปสร้างแบบจำลอง โดยอ้างอิงจากเส้นเส้นผ่านศูนย์กลางของแต่ละเส้นรอบวง ขั้นตอนสุดท้ายระบบสร้างผลลัพธ์โดยคำนวณตัวแปรจากผลลัพธ์ก่อนหน้านี้นี้ โดยผลลัพธ์ที่ได้ประกอบด้วยข้อมูลของไม้ ได้แก่ แบบจำลอง 3 มิติ ความยาว เส้นผ่านศูนย์กลาง และปริมาตรของไม้ตามลำดับ

ปี 2011 Molder และ Martens (Molder and Martens, 2011) [26] ได้เสนอว่าการประมวลผลภาพถ่ายมีประโยชน์ในด้านอุตสาหกรรมไม้ ปกติในอุตสาหกรรมจะมีงานที่ใช้ระบบเดิมที่ใช้มนุษย์วัดขนาดไม้อยู่เป็นจำนวนมาก ดังนั้นเราสามารถนำเทคนิคการประมวลผลภาพถ่ายมาประยุกต์ใช้ได้ และได้สำรวจวิธี 3D laser scanner เป็นระบบที่ใช้กล้องดิจิทัลวิเคราะห์สิ่งของและสร้างเป็นแบบจำลอง 3 มิติขึ้นมา มีทั้งหมด 2 วิธีหลัก ได้แก่

1. Laser signal delay : เป็นการส่งสัญญาณกระตุ้นออกจากเลเซอร์ แล้วสัญญาณก็จะสะท้อนกลับจากวัตถุไปยังเซนเซอร์ภาพ โดยมีการจับเวลาและวัดระยะห่างของจุดที่วัด มุมและตำแหน่งของเลเซอร์จะเปลี่ยนไปเรื่อย ๆ เพื่อวัดจุดต่อไป จุดเด่น คือ ความแม่นยำในการตรวจจับ จุดด้อย คือ การเคลื่อนที่ของวัตถุ เพราะตามโรงงานอาจมีฝุ่น และตกลงมาทำความเสียหายให้กับอุปกรณ์ ระหว่างการวัดได้

2. Laser projection : ใช้เส้นเลเซอร์และกล้องดิจิทัลถูกตั้งไว้ที่จุด ๆ หนึ่ง เส้นเลเซอร์จะถูกจับด้วยกล้องดิจิทัล หลังจากทีเส้นเลเซอร์ถูกจับภาพก็จะใช้เทคนิคการประมวลผลภาพถ่ายเพื่อตรวจจับเส้นจากรูปภาพ ถ้าทราบมุมระหว่างเลเซอร์และกล้องดิจิทัล และระยะห่างจากวัตถุก็จะทำให้ทราบความกว้างและความสูงของวัตถุนั้น ๆ ได้ จุดเด่น คือ มีการเคลื่อนที่ของวัตถุน้อย จุดด้อย คือ มีปัจจัยภายนอกเข้ามามีผลกระทบ เช่น สภาพแสง เงา เป็นต้น ทำให้ความแม่นยำลดลง และด้วยจุดเด่นของวิธีนี้จึงเป็นที่นิยมใช้มากที่สุดในอุตสาหกรรม

2.14.2 ประเภทที่ใช้ภาพถ่าย

ปี 2011 Rahman และคณะ (Rahman, Yella, and Dougherty, 2011) [2] ได้เสนอวิธีในการนับจำนวนท่อนซุงบนรถบรรทุกและรวบรวมข้อมูลการขนส่งไม้ก่อนเข้าไปยังโรงเลื่อยไม้ โดยใช้ภาพถ่ายจากท้ายรถบรรทุกไม้เพื่อดำเนินการด้วย 3 กระบวนการย่อย เริ่มจากกระบวนการแรกเป็นการประมวลผลภาพถ่ายเบื้องต้นด้วยการแปลงแบบจำลองสีจาก RGB เป็นแบบจำลอง HSV โดยนำ

เฉพาะค่าความสว่าง (Value) ออกมาจากแบบจำลองสี HSV เพื่อปรับความสว่างและความเปรียบต่าง ซึ่งผู้ใช้เป็นคนกำหนดค่าเอง จากนั้นนำค่าความสว่างไปรวมกับค่าสีและความอิ่มตัวสีและแปลงกลับเป็นแบบจำลองสี RGB ต่อมาเป็นการดำเนินการกับค่าจุดภาพด้วยวิธีนอร์มัลไลซ์เพื่อปรับความเปลี่ยนแปลงของแสง สำหรับกระบวนการที่สองเป็นการหาขอบภาพด้วยวิธี Canny edge และนำไปเข้ากระบวนการขยายภาพ 1 รอบ เพื่อให้ขอบภาพมีเส้นที่หนาขึ้น กระบวนการสุดท้ายเป็นการแบ่งส่วนภาพ (Image Segmentation) เพื่อหาบริเวณปลายท่อนซุง ซึ่งเป็นระบบย่อยหลักของงานวิจัยนี้ โดยนำภาพในกระบวนการแรกมาแปลงกลับเป็นแบบจำลองสี HSV อีกครั้งและกำหนดค่าขีดแบ่งช่วงค่าสีเพื่อเก็บจุดภาพที่มีช่วงค่าสีอยู่ในค่าขีดแบ่งที่กำหนดไว้ และนำขอบภาพที่ได้จากกระบวนการที่สองมาซ้อนทับกับภาพในกระบวนการสุดท้ายเพื่อสกัดส่วนที่ไม่ใช่บริเวณปลายท่อนซุงออกจากภาพ หลังจากนั้นนำไปเข้ากระบวนการคอนทัวร์เพื่อหาพื้นที่จุดเชื่อมกันทั้งหมดบนเส้น และคำนวณหาพื้นที่ในแต่ละพื้นที่ ซึ่งจะตัดพื้นที่ที่มีขนาดน้อยกว่า 300 จุดภาพออกไป หลังจากนั้นมีการติกรอบสี่เหลี่ยมบริเวณคอนทัวร์ที่หาได้เพื่อนำไปหาจุดศูนย์กลางวงกลม มีการวาดวงกลมใหม่โดยอ้างอิงจากจุดศูนย์กลางของคอนทัวร์ และกระบวนการสุดท้ายเป็นการป้ายชื่อจุดภาพที่อยู่ติดกัน เพื่อนับจำนวนปลายท่อนซุงทั้งหมดในภาพ สำหรับงานวิจัยของ Rahman และคณะสามารถทราบจำนวนท่อนซุงที่บรรทุกมา โดยผลการประเมินความถูกต้องของการนับไม้เฉลี่ยคิดเป็น 97% ของไม้ทั้งหมด

ปี 2014 Shvarts และ Tamre (Shvarts and Tamre, 2014) [3] ได้เสนอระบบพหุภาพสำหรับใช้ประเมินปริมาตรของวัตถุนขนาดใหญ่ ซึ่งใช้เทคโนโลยีแมชชีนวิชัน (Machine vision) ประยุกต์เข้ากับขั้นตอนวิธีและอุปกรณ์เคลื่อนที่อย่างสมาร์ตโฟน ซึ่งอัลกอริทึมใหม่นี้ถูกทดลองนำไปใช้ในการประเมินปริมาตรของท่อนซุงที่กองไว้บนพื้น กระบวนการของระบบนี้ เริ่มจากผู้ใช้เป็นคนถ่ายภาพท่อนซุงและต้องควบคุมคุณภาพการถ่ายภาพด้วยตัวเอง เช่น การโฟกัส ความไวชัตเตอร์ หรือขนาดรูรับแสง เป็นต้น รวมไปถึงการกำหนดขอบเขตที่สนใจบนภาพ (Region of Interest : ROI) จากนั้นจะมีการนอร์มัลไลซ์ภาพ ซึ่งภาพผลลัพธ์ที่ได้จะมีการปรับปรุงภาพให้มีความสว่างที่มากขึ้นกว่าเดิม เพื่อส่งต่อไปยังขั้นตอนต่อไป คือ การแบ่งส่วนภาพ ที่จำเป็นต้องใช้กระบวนการปรับความสว่างเพื่อช่วยในการตรวจจับวัตถุที่ต้องการได้ดีขึ้น ในการแบ่งส่วนภาพของงานวิจัยนี้ใช้แบบจำลองสี HSV เช่นเดียวกับงานวิจัยของ (Rahman, Yella, & Dougherty, 2011) ที่ได้กล่าวไว้ข้างต้น โดยจะวาดกราฟเพื่อดูการกระจายตัวของค่าสีเพื่อทำการสกัดบริเวณที่คาดการณ์ว่าเป็นท่อนซุง หลังจากนั้นใช้เทคนิคการเปลี่ยนแปลงลักษณะโครงสร้างภาพโดยกระบวนการโคลชิงและโอเพนนิ่ง ถูกนำมาใช้เพื่อปรับปรุงภาพผลลัพธ์ให้ดีขึ้น จากนั้นจะคำนวณหาเส้นผ่านศูนย์กลางของปลายท่อนซุงหนึ่งท่อน แล้วนำไปคำนวณหาปริมาตรทั้งหมดของกองไม้ ภาพผลลัพธ์จากการทดลองพบว่าความแม่นยำขึ้นอยู่กับ

ขนาดของวัตถุในภาพ งานวิจัยของ (Shvarts & Tamre, 2014) สามารถทำงานได้ดีสำหรับท่อนซุงที่มีขนาดเส้นผ่านศูนย์กลางมากกว่า 10 จุดภาพถ้าต่ำกว่านี้จะสามารถคำนวณเป็นพื้นที่รวมของกองวัตถุเท่านั้น

ปี 2004 Knyaz และคณะ (Knyaz, Visilter, and Zheltov, 2004) [27] ได้เสนอวิธีการวัดด้วยเทคนิคการประมวลผลภาพ 3 มิติเพื่อนำไปใช้ในอุตสาหกรรมงานไม้ การพัฒนาแบ่งออกเป็น 2 ส่วนหลัก คือ การประเมินปริมาตรของไม้ และคำนวณปริมาตรไม้บนสายพานลำเลียง ในส่วนแรกกระบวนการเริ่มจากการปรับเทียบมาตรฐาน (Calibration) โดยมีกล้องดิจิทัลจำนวน 2 กล้อง เพื่อทำการกำหนดจุดภาพอ้างอิงในภาพกับขนาดของวัตถุจริง ช่วยให้ทราบได้ว่าขนาดของวัตถุที่ต้องการวัดมีขนาดเท่ากับกี่จุดภาพ จากนั้นนำภาพที่ได้ไปเข้าอัลกอริทึมเพื่อตรวจจับหาวงกลมในภาพด้วยเทคนิคเวกเตอร์เกรเดียนต์ โดยสกัดจุดของเวกเตอร์เกรเดียนต์ที่มีค่าไม่ถึงค่าสูงสุดท้องถิ่น (Local maxima) ในทิศทางนั้นๆ วิธีการนี้จะทำให้คอนทัวร์ของวัตถุมีความคมชัดขึ้น มีการทดสอบหาจุดศูนย์กลางของวงกลมโดยพิจารณาจากระดับของเวกเตอร์จากจุดหนึ่งไปยังจุดศูนย์กลางเพื่อหาจำนวนจุดศูนย์กลางที่เป็นไปได้ในคอนทัวร์ จากค่าสูงสุดท้องถิ่นจะมีค่าสูงสุด คือ ค่าสูงสุดสากล (Global maxima) ซึ่งจะใส่จุดศูนย์กลางวงกลมและรัศมีไปในวงกลมที่มีจุดศูนย์กลางทั้งหมด โดยจะเริ่มจากวงกลมขนาดใหญ่ไปจนถึงวงกลมขนาดเล็ก และสร้างคอนเวกซ์ (Convex) ในการกำหนดขอบเขตของปลายท่อนซุงทั้งหมดเพื่อนำไปคำนวณหาพื้นที่ทั้งหมด สำหรับกระบวนการถัดมาเป็นการวัดปริมาตรของท่อนซุงที่ลำเลียงอยู่บนสายพาน โดยจะยิงด้วยแสงเส้นตรงไปยังท่อนซุงเพื่อสร้างแบบจำลอง 3 มิติขึ้นมาใหม่ ซึ่งสร้างได้ประมาณ 30-40% ของพื้นผิวท่อนซุง

ปี 2014 Knyaz และ Maksimov (Knyaz and Maksimov, 2014) [4] ได้ปรับปรุงวิธีของ (Knyaz, Visilter, and Zheltov, 2004) [27] โดยเสนอวิธีประมาณปริมาตรของกองท่อนไม้ โดยใช้อุปกรณ์ปรับเทียบมาตรฐานประกอบด้วยกล้อง DSLR จำนวน 2 กล้อง วางห่างจากกองไม้เป็นระยะ 2-20 เมตร และถ่ายภาพเพื่อนำไปประมวลผลจุดที่ตรงกันของทั้ง 2 กล้องนี้จากจุดอ้างอิงของภาพผ่านระบบ Photogrammetric ซึ่งจะช่วยในการวัดขนาดปลายท่อนซุงได้ หลังจากนั้นก็นำภาพไปหาขอบภาพด้วยเวกเตอร์เกรเดียนต์และสกัดจุดที่มีเกรเดียนต์ต่ำ หมายความว่า ณ ส่วนประกอบใด ๆ ในสนามเกรเดียนต์ที่มีค่าขีดแบ่งต่ำกว่าให้ส่วนประกอบบริเวณนั้นมีค่าเป็น 0 ขั้นตอนถัดมาเป็นการค้นหาวงกลมจากจุดเส้นรอบวงโดยใช้วิธีการค้นหาแบบวน (Iterative search) ของจุดสูงสุดสัมพัทธ์ (Local maxima) วงกลมเหล่านี้ถูกนำมาใช้เพื่อหาค่าปริมาตรของกองท่อนไม้

ปี 2014 Herbon และคณะ (Herbon, Tonnie, and Stock, 2014) [28] ได้นำเสนอวิธีในการตรวจจับวัตถุที่คล้ายกัน (Clustered objects detection) โดยที่ไม่ใช่วิธีสถิติในการสร้างแบบจำลอง แต่ใช้กระบวนการ iterative detect and merge ซึ่งเป็นการหาแบบวิธีทำซ้ำเฉพาะวัตถุ

ที่สนใจใหม่ในภาพ และคงเหลือไว้ซึ่งแบบจำลองผสมแบบเกาส์เซียนของการกระจายตัวที่ต้องการหา (Gaussian mixture model) เพื่อนำไปตรวจหาวัตถุใหม่ ๆ ในรอบถัดไป สำหรับวิธีการตรวจจับวัตถุขึ้นอยู่กับลักษณะเฉพาะที่ (Local features) ณ บริเวณนั้น ๆ โดยมีตัวสกัดคุณลักษณะที่นำมาใช้คือ LBP และ HOG

ปี 2014 Zhao และ Ren (Zhao and Ren, 2014) [29] ได้นำเสนอวิธีตรวจจับหน้าตัดของปลายท่อนซุงโดยใช้ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm : GA) ในการปรับปรุงภาพ และการกรองแบบโฮโมมอร์ฟิก (Homomorphic filter) ในการแบ่งส่วนภาพ กระบวนการเริ่มต้นจากปรับปรุงภาพด้วยเทคนิค GA โดยใช้แบบจำลองเชิงเส้นแบบช่วง (Piecewise linear transformation) มาช่วยในการหาปลายท่อนซุงที่ต้องการค้นหา จากนั้นคำนวณหาเส้นทางที่ดีที่สุดของ GA เพื่อกำหนดจุดเปลี่ยนทั้ง 3 ส่วนของภาพระดับเทา ได้แก่ ต่ำ ปานกลาง และสูง หลังจากนั้นจะสร้างฟังก์ชันความเหมาะสม (Fitness function) ให้ครอบคลุมกับภาพตัวอย่างโดยพิจารณาจากค่า mean square error เอนโทรปี (Entropy) และค่าอัตราส่วนของสัญญาณระหว่างสัญญาณที่ต้องการกับสัญญาณรบกวน (Signal to Noise Ratio : SNR) ขั้นตอนถัดมาเป็นกระบวนการกำจัดแสงเงาด้วยเทคนิคการกรองแบบโฮโมมอร์ฟิก ที่จะช่วยกำจัดแสงเงาในขอบเขตความถี่เดียวกันได้ ซึ่งภาพผลลัพธ์ของการแบ่งส่วนภาพจากวิธี GA และการกรองแบบโฮโมมอร์ฟิก มีภาพผลลัพธ์ที่ดีกว่าเมื่อเทียบกับวิธีการแบ่งส่วนภาพแบบการจัดกลุ่ม (K-means clustering : KMA) ซึ่งวิธีการแบ่งส่วนภาพแบบ KMA จำเป็นต้องกำหนดค่าเส้นผ่านศูนย์กลางและจำนวนท่อนซุงทั้งหมดที่ต้องการค้นหาในภาพก่อน หลังจากนั้นจะสุ่มเลือกข้อมูลเริ่มต้นจำนวน k ชุด และชุดข้อมูลที่เลือกมานั้นจะเป็นจุดศูนย์กลางเริ่มต้นของแต่ละกลุ่มข้อมูล หลังจากนั้นเป็นการจัดกลุ่มข้อมูลที่เหลือ โดยข้อมูลจะถูกจัดอยู่ในกลุ่มเดียวกันก็ต่อเมื่อข้อมูลนั้นมีความคล้ายกับตัวแทนของกลุ่มนั้นมากที่สุด และหาค่าเฉลี่ยของตัวแทนกลุ่มใหม่ จากนั้นดำเนินการตามขั้นตอนที่กล่าวมานั้นจนกระทั่งข้อมูลทั้งหมดถูกจัดกลุ่มอย่างสมบูรณ์ วิธี GA และการกรองแบบโฮโมมอร์ฟิกช่วยให้การแบ่งภาพปลายท่อนซุงบริเวณที่มีความสว่างน้อยได้ดีกว่าวิธีการแบ่งส่วนภาพโดยอาศัย KMA

ปี 2017 Mehrentsev และ Kruglov (Mehrentsev and Kruglov, 2017) [30] ได้เสนอวิธีการวัดท่อนซุงซึ่งระบบประกอบด้วย 3 ระบบย่อย ได้แก่ การตรวจจับวัตถุ (Object detection) การจัดกลุ่มท่อนซุง (Log clustering) และการแยกส่วนท่อนซุง (Log segmentation) เริ่มจากในการตรวจจับวัตถุ กระบวนการหาขอบภาพด้วยวิธีโซเบล (Sobel operation) ถูกใช้ในการสร้างเวกเตอร์เกรเดียนท์ (Gradient vector) โดยใช้ข้อมูลขนาดและทิศทางที่คำนวณจากจุดศูนย์กลางของวงกลม

จากนั้นจึงทำการค้นหาท่อนซุงในภาพด้วยวิธีจุดสูงสุดสัมพัทธ์ (Local maximum) สำหรับวิธีนี้สามารถแก้ไขปัญหาการซ้อนทับกันของท่อนซุงและสามารถตรวจจับท่อนซุงที่อยู่ติดกันได้โดยใช้วิธีการจัดกลุ่มแบบมีนชิฟท์ (Mean shift) ในกรณีทีกลุ่มของท่อนซุงที่ถูกแยกออกจากกัน วิธีการจัดกลุ่มแบบกราฟ (Graph clustering) ถูกนำมาใช้กับท่อนซุงที่ได้รับการตรวจจับแล้วเพื่อค้นหากลุ่มที่เชื่อมติดกันของท่อนซุง ในการแยกส่วนท่อนซุง อัลกอริทึมแบบ Watershed และ minimum s-t graph ถูกนำมาใช้เพื่อแบ่งส่วนบริเวณของท่อนซุงในแต่ละท่อน ซึ่งอาจมีท่อนซุงบางท่อนที่ระบบไม่สามารถตรวจจับได้หรือตรวจจับได้ไม่ถูกต้อง ระบบนี้อ่อนุญาตให้มีการแก้ไขด้วยตนเองเพื่อเพิ่มหรือลบท่อนซุงที่ต้องการได้

งานวิจัยที่เกี่ยวข้องกับการนำแสงเลเซอร์โดยใช้วิธี 3D laser scanner เหมาะกับท่อนซุงขนาดใหญ่ และต้องดำเนินการในสถานที่ที่ติดตั้งอุปกรณ์ไว้ รวมถึงมีค่าใช้จ่ายในการติดตั้งที่สูง วิธีนี้จึงนิยมใช้กันในแถบประเทศที่มีชื่อเสียงเรื่องอุตสาหกรรมไม้อย่างประเทศในแถบสแกนดิเนเวีย เช่น ประเทศนอร์เวย์ ประเทศสวีเดน และประเทศฟินแลนด์ เป็นต้น

งานวิจัยการตรวจจับและวัดขนาดปลายท่อนซุงในภาพถ่ายตามที่สำรวจและสรุปข้างต้นนั้นพบว่า เป็นวิธีที่ตรวจจับและวัดขนาดปลายท่อนซุงของไม้สนเป็นหลัก วิธีการส่วนใหญ่ที่กล่าวมาทำงานได้ดีกับภาพพื้นผิวหน้าตัดของปลายท่อนซุงที่สะอาดซึ่งมีสีสม่ำเสมอและปลายท่อนซุงที่มีรูปร่างคล้ายวงกลม ในทางกลับกันภาพจากรถบรรทุกไม้ยูคาลิปตัสอาจทำให้เกิดปัญหากับวิธีข้างต้นได้ ปัญหาเหล่านี้ประกอบด้วยการจัดวางท่อนซุงบนรถบรรทุก สีหน้าตัดของปลายท่อนซุงไม่สม่ำเสมอ หน้าตัดของปลายท่อนซุงซ้อนกัน แสงเงาบบังหน้าตัดของปลายท่อนซุง หรือมีวัตถุที่ไม่ต้องการปกคลุมปลายท่อนซุงในบางส่วน โดยปลายท่อนซุงไม้ยูคาลิปตัสมีวงปีค่อนข้างชัด มีขนาดเล็กกว่าไม้สน ปลายไม้ตัดไม่ได้ฉาก บางครั้งมีการแตกที่หน้าตัดไม้ และปลายของไม้ยื่นมาไม่สม่ำเสมอ นอกจากนี้ท่อนซุงมีขนาดและรูปทรงที่แตกต่างกันออกไปและอาจไม่เป็นวงกลม จากปัญหาดังกล่าวพบว่าการตรวจจับของรถบรรทุกไม้ในประเทศไทยมีความซับซ้อนโดยเฉพาะอย่างยิ่งในไม้ยูคาลิปตัส ดังรูปที่ 2-27 ซึ่งงานวิจัยที่เสนอนี้จะเป็นการตรวจจับหน้าตัดของปลายท่อนซุงจากภาพถ่ายด้านท้ายรถบรรทุกไม้ยูคาลิปตัส



รูปที่ 2-27 ตัวอย่างภาพถ่ายของรูปแบบต่าง ๆ ของรถบรรทุกไม้ยูคาลิปตัสในประเทศไทย

จากงานวิจัยที่เกี่ยวข้องพบว่า กระบวนการประมวลผลข้อมูลจากภาพถ่าย เป็นวิธีที่มีค่าใช้จ่ายในการพัฒนาระบบที่ต่ำกว่าการประมวลผลข้อมูลจากแสงเลเซอร์ และสามารถดำเนินการได้ทุกสถานที่ที่มีอุปกรณ์การถ่ายภาพ การประมวลผลข้อมูลจากภาพถ่ายจึงเป็นเทคนิคที่เหมาะสมในการตรวจจับหน้าตัดของปลายท่อนซุงจากภาพถ่ายท้ายรถบรรทุกไม้ยูคาลิปตัส

บทที่ 3 แนวคิดและวิธีดำเนินงาน

ในบทนี้ผู้วิจัยได้นำเสนอวิธีใหม่ในการตรวจจับหน้าตัดของปลายท่อนซุงไม้ยูคาลิปตัสจากภาพถ่ายหลังท้ายรถบรรทุกในประเทศไทย โดยใช้ภาพถ่ายจากโรงงานผลิตกระดาษ ณ บริษัทสยามเซลลูโลส จำกัด ซึ่งผู้วิจัยได้เข้าเยี่ยมชมดูงานในสถานที่จริงและบันทึกภาพถ่ายตัวอย่างท้ายรถบรรทุกขณะมีไม้ยูคาลิปตัสบรรทุกอยู่ ภาพที่ได้เป็นภาพสีที่มีความละเอียด 1632 x 1088 จุดภาพ ดังรูปที่ 3-1 โดยภาพถ่ายถูกบันทึกในเวลาประมาณระหว่าง 8:00 – 17:00 น. และเป็นภาพนิ่งเท่านั้น



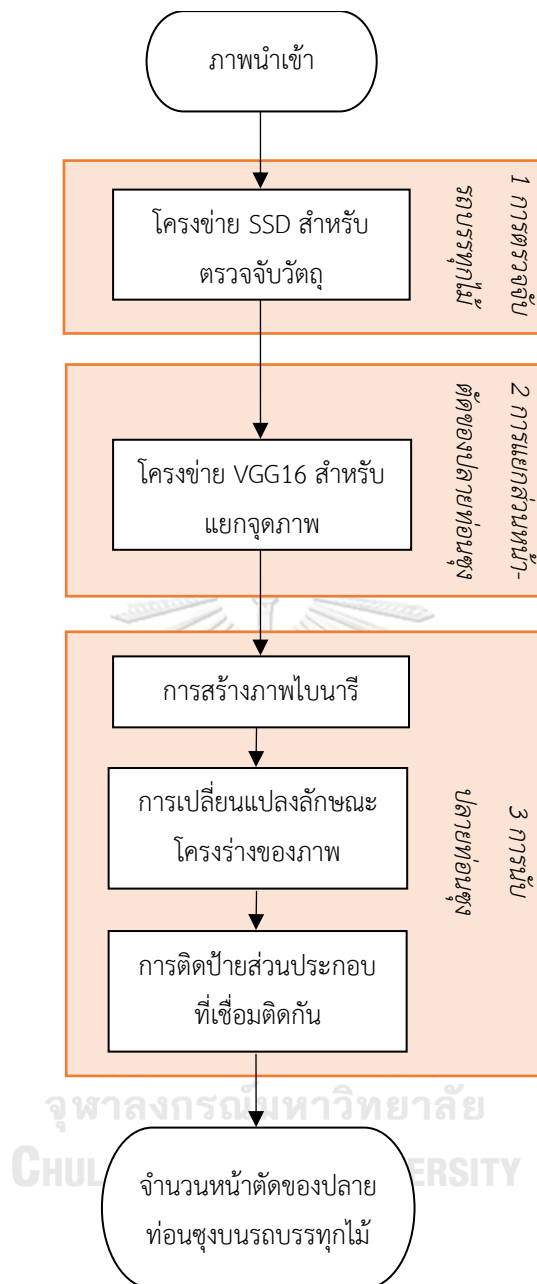
รูปที่ 3-1 ตัวอย่างภาพถ่ายด้านท้ายรถบรรทุกจากโรงงานผลิตกระดาษ ณ บริษัทสยามเซลลูโลส จำกัด

จุดประสงค์ของงานวิจัยนี้เพื่อตรวจจับว่ามีหน้าตัดของปลายท่อนซุงอยู่ในภาพหรือไม่ ซึ่งวิธีที่เสนอนี้มีขอบเขตของการวิจัย คือ ตรวจจับหน้าตัดของปลายท่อนซุงเฉพาะภาพถ่ายจากรถบรรทุกเท่านั้นและไม่สามารถทำงานร่วมกันได้กับภาพถ่ายจากรถบรรทุกแบบกึ่งพวง (Trailer truck) ดังรูปที่ 3-2 เนื่องจากวิธีการที่เสนอจะเป็นภาพถ่ายจากท้ายรถบรรทุกเท่านั้น ซึ่งภาพถ่ายจากรถกึ่งพวงถูกบันทึกที่บริเวณด้านข้างของรถกึ่งพวงและไม่รองรับกับการทำงานในสถานการณ์จริง ณ โรงงานผลิตกระดาษ เนื่องจากจำเป็นต้องใช้ระยะห่างระหว่างรถกึ่งพวงและผู้วิจัยขณะทำการบันทึกภาพมากกว่าปกติรวมถึงไม่มีพื้นที่เพียงพอต่อการบันทึกภาพ ณ บริเวณจุดตรวจรับสินค้า



รูปที่ 3-2 ตัวอย่างภาพถ่ายด้านข้างรถกึ่งพวงมลัยยูคาลิปตัส

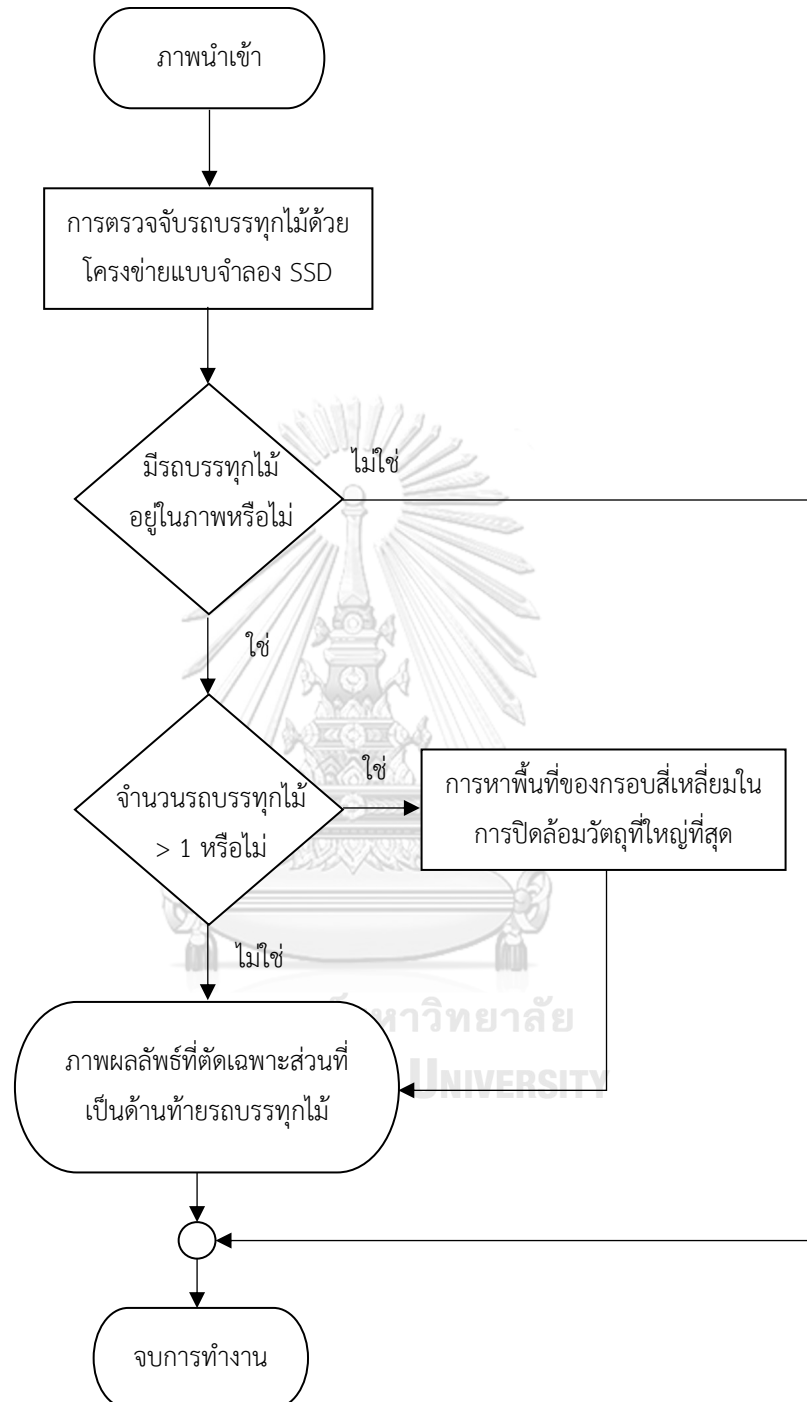
ภาพรวมของระบบนี้แสดงดังรูปที่ 3-3 เริ่มจากภาพขาเข้าถูกถ่ายโดยกล้องในช่วงเวลา กลางวัน การตรวจจับวัตถุ (Object detection) ถูกนำไปใช้ในพื้นที่ที่น่าสนใจของรถบรรทุกไม้ โดย ภาพขาเข้าถูกตัดโดยใช้พิกัดของขอบเขตที่ตรวจพบรถบรรทุกไม้ จากนั้นการแยกส่วนไม้ (Log segmentation) ถูกนำไปใช้เพื่อแบ่งส่วนปลายไม้แต่ละท่อนออกจากภาพพื้นหลังซึ่งกระบวนการนี้ ดำเนินการโดยใช้วิธีการจำแนกทางความหมายตามโครงข่ายคอนโวลูชันแบบทั่วถึง (Fully convolutional network-based semantic segmentation) นอกจากนี้ระบบยังใช้เทคนิคการ ประมวลผลภาพโดยใช้การแปลงภาพไบนารี (Binary conversion) การเปลี่ยนแปลงลักษณะโครงร่าง ของภาพ (Morphological operation) และการติดป้ายส่วนประกอบที่เชื่อมติดกัน (Connected component labeling)



รูปที่ 3-3 แผนภาพแสดงภาพรวมของระบบการตรวจจับหน้าตัดของปลายท่อนซุง

ระบบการตรวจจับหน้าตัดของปลายท่อนซุงที่เสนอสามารถอธิบายได้โดยแบ่งเป็น 3 ขั้นตอนสำคัญ ได้แก่ การตรวจจับรถบรรทุกไม้ การแยกส่วนหน้าตัดของปลายท่อนซุง และการนับหน้าตัดของปลายท่อน

3.1 การตรวจจับรถบรรทุกไม้ (Timber truck detection)



รูปที่ 3-4 แผนภาพแสดงกระบวนการตรวจจับรถบรรทุกไม้ยุคาลิปตัส

การตรวจจับรถบรรทุกไม่เป็นกระบวนการที่ตรวจจับรถบรรทุกไม่ในภาพถ่ายจากด้านท้ายรถบรรทุกแสดงดังรูปที่ 3-4 กระบวนการนี้เริ่มจากการประยุกต์ใช้โครงสร้างเชิงเกิลชอตดีเทคเตอร์ (Single Shot Detector : SSD) ที่เสนอโดย Liu และคณะ [14] กับสถาปัตยกรรมของตัวสกัดคุณลักษณะ MobileNet [15] ถูกนำมาใช้เพื่อตรวจสอบสี่เหลี่ยมในการปิดล้อมวัตถุ (Bounding boxes) ของรถบรรทุกไม่ในภาพ เริ่มจากการติดป้ายกำกับข้อมูลประเภทใหม่ในชุดข้อมูลสอนที่ประกอบด้วยภาพรถบรรทุกไม่และดำเนินการสอนชุดข้อมูลนี้บนแบบจำลองที่ได้รับการสอนล่วงหน้า SSD Mobilenet v1 COCO จากนั้นแบบจำลอง SSD Mobilenet v1 COCO ถูกใช้เพื่อหาพื้นที่ที่เป็นไปได้ของสี่เหลี่ยมในการปิดล้อมวัตถุที่มีรถบรรทุกไม่อยู่ในบริเวณนั้น เมื่อขั้นตอนการสอนเสร็จสิ้น ขั้นตอนต่อไป คือ การทดสอบผลลัพธ์ในชุดข้อมูลทดสอบและตอบผลลัพธ์ในลักษณะภาพที่มีกรอบสี่เหลี่ยมในการปิดล้อมวัตถุซ้อนทับบริเวณที่มีรถบรรทุก จากนั้นระบบจะดำเนินการตัดภาพส่วนเฉพาะด้านท้ายรถบรรทุกไม่ในภาพโดยอ้างอิงจากพิกัดของกรอบสี่เหลี่ยมในการปิดล้อมวัตถุ ดังรูปที่ 3-5 (ก) (ค) และ (จ)

อย่างไรก็ตามอาจมีภาพนำเข้าบางภาพที่มีรถบรรทุกไม่หลายคันปรากฏอยู่ในภาพ ระบบนี้จะกำจัดรถบรรทุกคันอื่นที่ไม่สนใจออกไปด้วยการคำนวณพื้นที่ของแต่ละกรอบสี่เหลี่ยมที่ปิดล้อมรถบรรทุก โดยพื้นที่ของกรอบสี่เหลี่ยมที่ปิดล้อมที่มีขนาดใหญ่ที่สุดจะได้รับเลือกเป็นพื้นที่ที่สนใจเพื่อนำไปดำเนินการต่อในขั้นตอนต่อไป ดังรูปที่ 3-5 (ข) (ง) และ (ฉ) เหตุผลที่ผู้วิจัยกำจัดรถบรรทุกไม่คันอื่นออกไป เนื่องจากในสถานการณ์จริง ณ โรงงานผลิตกระดาษจะอนุญาตให้รถบรรทุกไม่เคลื่อนรถเพื่อเข้าไปดำเนินการขนถ่ายน้ำหนักและตรวจรับสินค้าไม้ครั้งละ 1 คันเท่านั้น ดังนั้นวัตถุประสงค์ของการตรวจจับรถบรรทุกไม่ คือ การกำจัดบริเวณอื่นที่ไม่ใช่รถบรรทุกไม่รวมถึงรถบรรทุกไม่คันอื่น ๆ ซึ่งไม่ใช่คันที่ผู้ตรวจสอบกำลังดำเนินการตรวจรับสินค้าอยู่

ด้วยการกำจัดรถบรรทุกไม่จากวิธีดังกล่าว ระบบสามารถลดต้นทุนการคำนวณสำหรับการจำแนกทางความหมายและกำจัดวัตถุที่ไม่ต้องการออกไปจากภาพ ยกเว้นรถบรรทุกไม่คันที่ต้องการเท่านั้น



(ก)



(ข)



(ค)



(ง)



(จ)



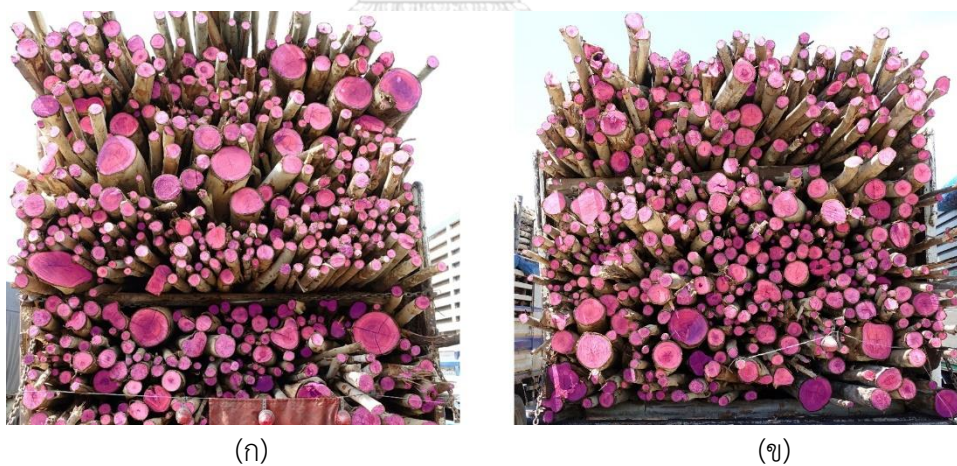
(ฉ)

รูปที่ 3-5 ภาพตัวอย่างผลลัพธ์ของขั้นตอนการตรวจจ้บรถบรรทุกไม้ (ก) ภาพนำเข้าปกติ (ข) ภาพนำเข้าที่มีรถบรรทุกไม้มากกว่า 1 คัน (ค) ผลลัพธ์ของการตรวจจ้บรถบรรทุกไม้ในภาพนำเข้าปกติ (ง) ผลลัพธ์ของการตรวจจ้บรถบรรทุกไม้ในภาพนำเข้าที่มีรถบรรทุกไม้มากกว่า 1 คัน (จ) และ (ฉ) ผลลัพธ์ภาพที่ถูกตัดภาพส่วนเฉพาะด้านท้ายรถบรรทุกไม้ยูคาลิปตัส

3.2 การแยกส่วนหน้าตัดของปลายท่อนซุง (Log segmentation)

ในขั้นตอนนี้ระบบดำเนินการด้วยวิธีการจำแนกทางความหมายตามโครงข่ายคอนโวลูชันแบบทั่วถึงซึ่งใช้กันอย่างกว้างในการพัฒนาโครงข่ายคอนโวลูชันแบบปลายทางถึงปลายทาง [18] จากขั้นตอนการตรวจจับรถบรรทุกไม้ก่อนหน้านี้ทำให้ตรวจพบพื้นที่ของรถบรรทุกไม้ วิธีการแบ่งส่วนทางความหมายจึงถูกนำมาใช้ในการสกัดจุดภาพเฉพาะที่เป็นพื้นที่ของหน้าตัดของปลายท่อนซุงเท่านั้น โดยจุดภาพที่ถูกติดป้ายกำกับในภาพของชุดข้อมูลสอนด้วยประเภทของหน้าตัดของปลายท่อนซุงและประเภทของที่ไม่ใช่หน้าตัดของปลายท่อนซุง และเพื่อปรับปรุงอัตราความแม่นยำในการเรียนรู้เทคนิคการเพิ่มคุณค่าของข้อมูลด้วยไลบรารี OpenCV ได้ถูกนำมาใช้เพื่อเพิ่มปริมาณและความหลากหลายของชุดข้อมูลสอน

หลังจากที่ขั้นตอนการเตรียมชุดข้อมูลสอน วิธีการถ่ายโยงการเรียนรู้ถูกใช้ในการปรับตัวแปรของโครงข่ายผู้วิจัยใช้แบบจำลองที่ได้รับการสอนล่วงหน้าของโครงข่าย VGG16 ซึ่งเสนอโดย Simonyan และ Zisserman [16] จากนั้นโครงข่ายคอนโวลูชันแบบทั่วถึงถูกสอนบนแบบจำลอง VGG16 ด้วยชุดข้อมูลเพื่อแยกส่วนพื้นที่จุดภาพของหน้าตัดของปลายท่อนซุงในภาพ ผลลัพธ์ของชุดข้อมูลทดสอบที่สกัดจุดภาพของหน้าตัดของปลายท่อนซุงแสดงดังรูปที่ 3-6

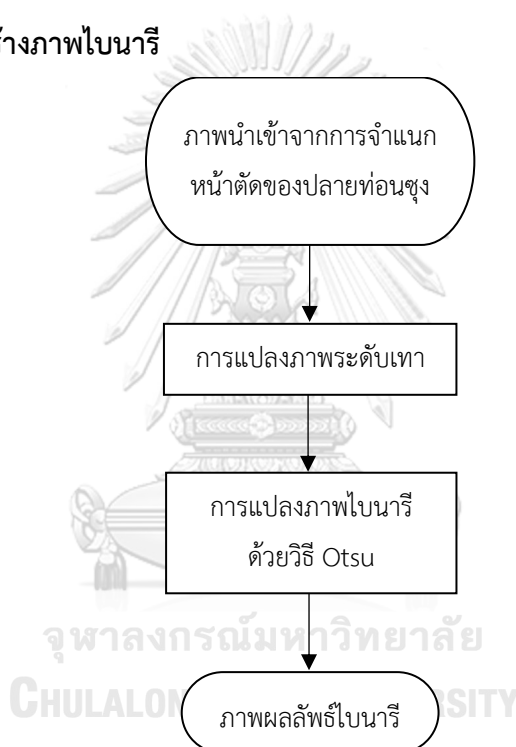


รูปที่ 3-6 ภาพผลลัพธ์ของการแยกส่วนหน้าตัดของปลายท่อนซุงด้วยโครงข่าย VGG16 (ก) ภาพนำเข้าปกติ และ (ข) ภาพนำเข้าที่มีรถบรรทุกไม้มากกว่า 1 คัน

3.3 การนับปลายท่อนซุง (Log counting)

ในขั้นตอนนี้วิธีการนับปลายท่อนซุงถูกนำเสนอโดยใช้เทคนิคการประมวลผลภาพซึ่งส่วนใหญ่ใช้ไลบรารีจาก OpenCV ในการดำเนินการ โดยภาพที่ถูกแยกส่วนเฉพาะหน้าตัดของปลายท่อนซุงจากขั้นตอนที่ 3.2 อาจมีหน้าตัดของปลายท่อนซุงบางส่วนที่ติดกันกับหน้าตัดของปลายท่อนซุงท่อนอื่น ๆ รวมถึงอาจมีส่วนที่ไม่ใช่หน้าตัดของปลายท่อนซุงปรากฏขึ้นในผลลัพธ์ ดังนั้นเพื่อแก้ปัญหาที่กล่าวมานั้นจึงได้มีการใช้วิธีการนับปลายท่อนซุงมาใช้ในการแยกปลายท่อนซุงที่อยู่ติดกันและกำจัดสิ่งรบกวนในภาพที่ไม่ใช่ส่วนของปลายท่อนซุงออกไป

3.1.1 การสร้างภาพไบนารี



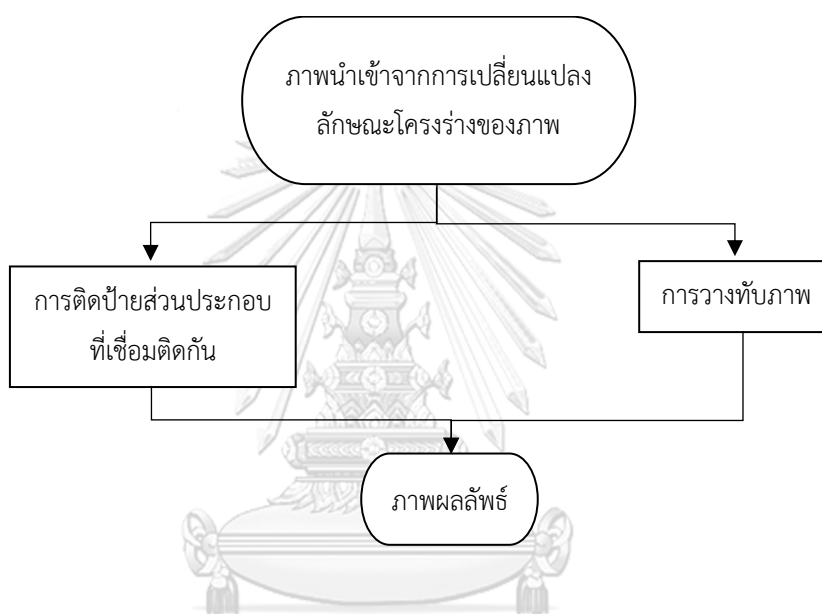
รูปที่ 3-7 แผนภาพแสดงกระบวนการสร้างภาพไบนารี

ภาพไบนารีถูกสร้างขึ้นจากภาพผลลัพธ์ของการแยกส่วนหน้าตัดของปลายท่อนซุงแสดงดังรูปที่ 3-7 โดยเริ่มจากใช้วิธีการแปลงเป็นภาพเป็นภาพระดับเทาบนภาพที่ได้รับการแยกส่วน จากนั้นใช้การกำหนดค่าขีดแบ่งครอบคลุม (Global thresholding) ของ Otsu [7] บนภาพระดับสีเทาเพื่อแปลงภาพที่ได้รับการแยกส่วนระดับเทาเป็นภาพไบนารี

3.1.2 การเปลี่ยนแปลงลักษณะโครงร่างของภาพ

ภาพผลลัพธ์ของการแยกส่วนหน้าตัดของปลายท่อนซุงอาจมีพื้นที่ของหน้าตัดของปลายท่อนซุงที่ติดกันกับปลายท่อนซุงท่อนอื่น ๆ ซึ่งบริเวณรอยต่อที่ติดกันหรือเชื่อมต่อกันเหล่านี้จะถูกแยกออกจากกันโดยใช้กระบวนการกร่อนมาดำเนินการกับภาพ และตามด้วยกระบวนการโตเลชันซึ่งดำเนินการกับภาพไบนารี ในขั้นตอนนี้สามารถสร้างช่องว่างระหว่างปลายท่อนซุงและลดส่วนที่ไม่ต้องการในภาพออกไปดังรูปที่ 3-9 (ก) และ (ข)

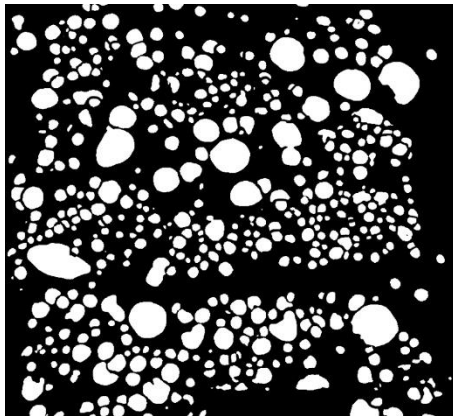
3.1.3 การติดป้ายส่วนประกอบที่เชื่อมติดกัน



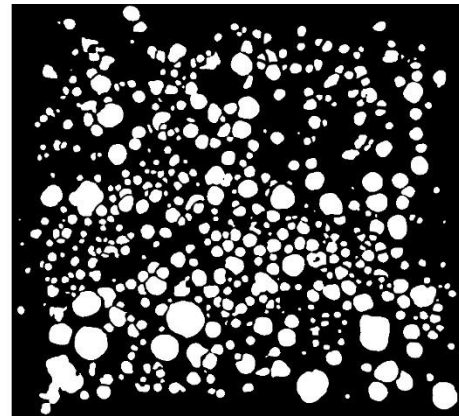
รูปที่ 3-8 แผนภาพแสดงกระบวนการติดป้ายส่วนประกอบที่เชื่อมติดกัน

CHULALONGKORN UNIVERSITY

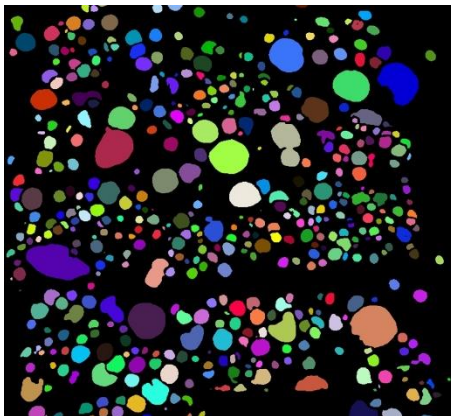
ขั้นตอนสุดท้ายเป็นการติดป้ายส่วนประกอบที่เชื่อมติดกัน ดังรูปที่ 3-8 หน้าตัดของปลายท่อนซุงถูกนับโดยใช้วิธีการติดป้ายส่วนประกอบที่เชื่อมติดกันของ OpenCV ในขั้นตอนนี้ทำให้พบหน้าตัดของปลายท่อนซุงที่ไม่ได้อยู่ติดกันและดำเนินการติดป้ายชื่อจุดภาพที่อยู่ในบริเวณเดียวกันของแต่ละหน้าตัดของปลายท่อนซุงดังรูปที่ 3-9 (ค) และ (ง) จากนั้นนำภาพผลลัพธ์จากขั้นตอนการเปลี่ยนแปลงลักษณะโครงร่างภาพนำไปวางทับ (Overlay) ด้วยภาพนำเข้าทำให้ได้ผลลัพธ์ของการติดป้ายหรือการนับแสดงดังรูปที่ 3-9 (จ) และ (ฉ) โดยผลลัพธ์สุดท้ายของการตรวจจับหน้าตัดของปลายท่อนซุงจากภาพถ่ายด้านท้ายรถบรรทุกประกอบด้วยจำนวนปลายท่อนซุงทั้งหมดที่ตรวจจับได้และภาพผลลัพธ์ของปลายท่อนซุงที่ระบบสามารถตรวจจับได้



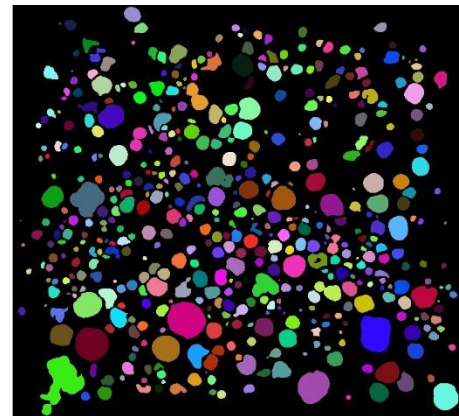
(ก)



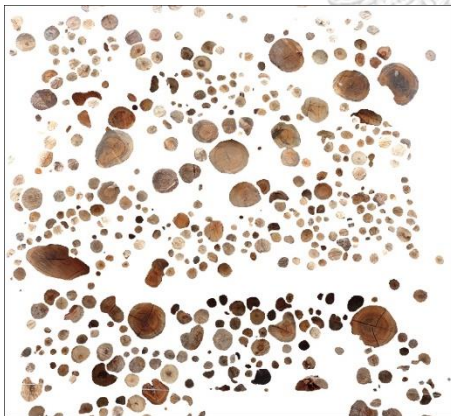
(ข)



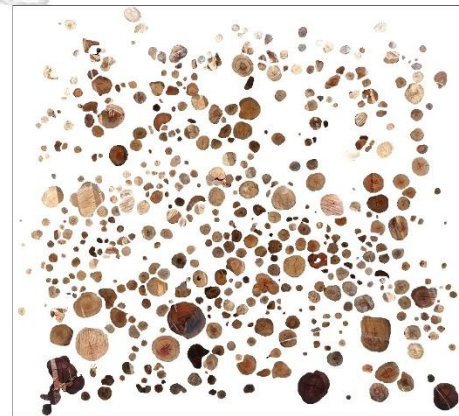
(ค)



(ง)



(จ)



(ฉ)

รูปที่ 3-9 ภาพผลลัพธ์ของระบบการตรวจจับหน้าตัดของขั้นตอนการนับปลายท่อนซุง โดยที่ (ก) และ (ข) เป็นผลลัพธ์ของขั้นตอนการเปลี่ยนแปลงลักษณะโครงสร้างของภาพที่ประยุกต์บนภาพไบนารี (ค) และ (ง) เป็นผลลัพธ์ของการนับปลายท่อนซุง (จ) และ (ฉ) เป็นการซ้อนทับระหว่างภาพนำเข้าและภาพผลลัพธ์

บทที่ 4

การทดลองและผลของการทดลอง

4.1 ภาพที่ใช้ในการทดลอง

ภาพที่ใช้ในการทดลองเป็นภาพที่ได้จากการบันทึกระหว่างเข้าเยี่ยมชมดูงาน ณ บริษัทสยามเซลลูโลส จำกัด โดยใช้กล้องดิจิทัล Sony A6000 ที่มีเลนส์ระยะ 18-105 มิลลิเมตรถูกนำมาใช้เพื่อถ่ายภาพจากด้านท้ายรถบรรทุกไม้ยูคาลิปตัสที่มีความหลากหลายของมุมที่แตกต่างกันออกไปในช่วงระยะเวลา 08:00 – 17:00 น. ซึ่งมีความละเอียดของภาพ 1632 x 1088 จุดภาพ และไฟล์ภาพมีรูปแบบเป็นภาพสี JPEG ที่มีความลึก 24 บิตในแบบจำลองสี RGB ที่มีมุมของกล้องในการบันทึกภาพเป็นมุมประมาณแนวระดับกับกลางกลุ่มท่อนซุง

ชุดข้อมูลภาพถูกแบ่งออกเป็นชุดข้อมูลสอน (Training set) สำหรับเป็นตัวอย่างข้อมูลที่ใช้เพื่อสอนแบบจำลอง ชุดข้อมูลตรวจสอบ (Validation set) สำหรับเป็นตัวอย่างข้อมูลที่ใช้เพื่อประเมินแบบจำลอง และชุดข้อมูลทดสอบ (Testing set) สำหรับเป็นตัวอย่างข้อมูลที่ใช้เพื่อทดสอบ โดยในการตรวจจับรถบรรทุกไม้ยูคาลิปตัสและการแยกส่วนหน้าตัดของปลายท่อนซุง ชุดข้อมูลประกอบด้วยกรณีต่าง ๆ ของหน้าตัดของปลายท่อนซุงตามที่ระบุไว้ในส่วนงานวิจัยที่เกี่ยวข้อง

4.2 สภาพแวดล้อมด้านซอฟต์แวร์และฮาร์ดแวร์

วิธีการที่เสนอถูกพัฒนาโดยใช้ทรัพยากรคอมพิวเตอร์เพื่อดำเนินการกับชุดข้อมูล ดังนี้

4.2.1 ฮาร์ดแวร์

หน่วยประมวลผล Intel ® Core™ i7-6700 CPU @ 3.40 GHz

การ์ดแสดงผลภาพ NVIDIA GeForce GTX 1080Ti

หน่วยความจำ 32.0 GB

4.2.2 ซอฟต์แวร์

ระบบปฏิบัติการ Ubuntu 14.04 LTS

เครื่องมือในการเรียนรู้เชิงลึก Tensorflow

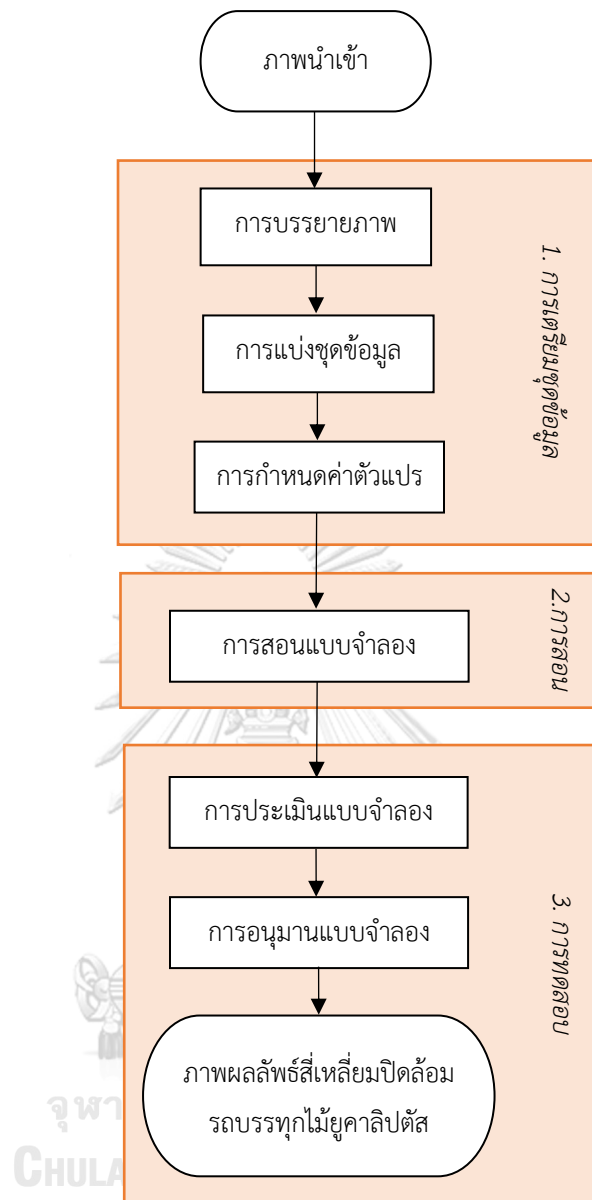
ไลบรารีทางด้านภาพ OpenCV 3

ภาษา Python 3

เครื่องมือในการช่วยพัฒนาโปรแกรม PyCharm Community Edition

4.3 การตรวจจําบรทุกไม้ยุคาลิปตัสด้วยโครงสร้าง SSD

ในการทดลองการตรวจจําบรทุกไม้ยุคาลิปตัสด้วยโครงข่ายแบบจําลอง SSD นั้น ขั้นตอนการเตรียมข้อมูลและการสอนมีความสำคัญในกระบวนการตรวจจําบรทุกไม้ยุคาลิปตัสในภาพว่ามีบรทุกไม้ยุคาลิปตัสหรือไม่ การทดลองดำเนินการโดยใช้การถ่ายโยงการเรียนรู้ (Transfer learning) แบบจําลองที่ได้รับการสอนล่วงหน้า (Pre-trained model) ของไมโครซอฟท์คอมมอนออฟเจคอินคอนเทคส์ (Common Objects in Context : COCO) [31] ที่เรียกว่าโครงสร้าง SSD และตัวสกัดคุณลักษณะ MobileNet เหตุผลที่ผู้วิจัยนำแบบจําลองที่ได้รับการสอนล่วงหน้ามาใช้เนื่องจากแทนที่จะสร้างแบบจําลองใหม่ขึ้นมาตั้งแต่เริ่มต้น การใช้แบบจําลองที่ได้รับการสอนมาแล้วสำหรับการแก้ปัญหาที่คล้ายกันนั้นสามารถใช้จุดเริ่มต้นสำหรับการสอนในโครงข่ายที่ต้องการได้ โดยในการทดลองนี้ใช้จุดตรวจของแบบจําลอง SSD MobileNet v1 COCO ที่ได้รับการสอนล่วงหน้าจาก Tensorflow zoo [32] ซึ่งแบบจําลองนี้ถูกใช้เป็นจุดเริ่มต้นสำหรับการสอนและได้รับการสอนเพิ่มเติมด้วยภาพถ่ายด้านท้ายบรทุกไม้ยุคาลิปตัส จากนั้นแบบจําลองที่ปรับนี้จะถูกใช้สำหรับการอนุมาน (Inference) ในขั้นตอนการทดสอบ โดยขั้นตอนการทดลองในการตรวจจําบรทุกไม้ยุคาลิปตัสแสดงดังรูปที่ 4-1



รูปที่ 4-1 แผนภาพแสดงการทดลองการตรวจจับไม้ยูคาลิปตัสด้วยแบบจำลอง SSD MobileNet v1
COCO

4.3.1 การเตรียมข้อมูล

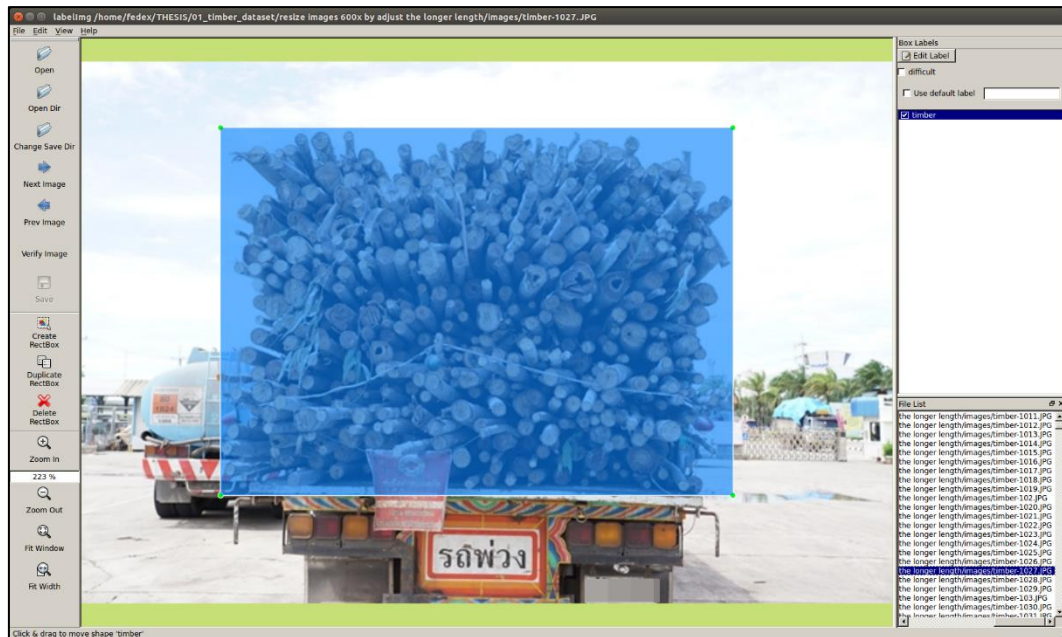
ขั้นตอนแรกในการทดลอง คือ การรวบรวมภาพนำเข้าซึ่งผู้วิจัยได้เยี่ยมชมดูงาน ณ โรงงานผลิตกระดาษเพื่อบันทึกภาพถ่ายด้านท้ายรถบรรทุกไม้ยูคาลิปตัสโดยภาพถ่ายมีความหลากหลายของมุมที่แตกต่างกันออกไปในช่วงเวลากลางวันดังที่กล่าวไปในข้อ 4.1 นอกจากนี้ผู้วิจัยยังนำภาพนำเข้าที่บันทึกด้วยสมาร์ทโฟนที่มีความละเอียดของภาพ 3264 x 2448 จุดภาพซึ่งมีระยะหรือส่วนของรถบรรทุกไม้ยูคาลิปตัสที่สามารถมองเห็น (Depth of Field : DOF) มีความคมชัดน้อยกว่า

ภาพที่ถูกบันทึกด้วยกล้องดิจิทัลเพื่อให้ระบบตรวจจับนี้สามารถทำงานได้ภายใต้สภาวะที่ต่างกัน ของแสง มุม และความคมชัดของภาพ เนื่องจากภาพนำเข้าต้นฉบับมีขนาดใหญ่ ซึ่งจำเป็นต้องใช้ หน่วยความจำจำนวนมาก ดังนั้นภาพนำเข้าทั้งหมดจึงถูกย่อขนาดด้วยฟังก์ชัน `resize` ในไลบรารี `OpenCV` ให้เหลือขนาด `600 x 400` จุดภาพเพื่อรักษาอัตราส่วนของภาพ ตัวอย่างชุดข้อมูลแสดงดัง รูปที่ 4-2



รูปที่ 4-2 ชุดข้อมูลภาพนำเข้ารถบรรทุกไม้ยูคาลิปตัส

การสอนแบบจำลองตรวจจับวัตถุ ในแต่ละภาพจำเป็นต้องทราบความกว้าง ความสูง และ ข้อมูลประเภทแต่ละประเภทโดยใช้สี่เหลี่ยมในการปิดล้อมวัตถุ ซึ่งประกอบด้วยค่า `xmin`, `xmax`, `ymin` และ `ymax` เพราะฉะนั้นสี่เหลี่ยมในการปิดล้อมวัตถุของการทดลองนี้ คือ บริเวณกรอบ สี่เหลี่ยมที่ล้อมรอบรถบรรทุกไม้ยูคาลิปตัส ซึ่งถ้าภาพใด ๆ มีรถบรรทุกไม้ยูคาลิปตัสมากกว่า 1 คันขึ้นไป ผู้วิจัยจะดำเนินการล้อมกรอบสี่เหลี่ยมในการปิดล้อมวัตถุของรถบรรทุกไม้ยูคาลิปตัสทุกคัน โดยในการบรรยายภาพผู้วิจัยใช้เครื่องมือ `labelling [33]` ดังรูปที่ 4-3 (ก) ที่จะสร้างคำอธิบายภาพในรูปแบบ `Pascal VOC` ซึ่งถูกบันทึกอยู่ในรูปแบบไฟล์ `eXtensible Markup Language (XML)` ที่ถูกสร้างขึ้นสำหรับแต่ละภาพ ไฟล์ `XML` ประกอบด้วยพิกัดของสี่เหลี่ยมในการปิดล้อมวัตถุ ชื่อไฟล์ภาพ ข้อมูลประเภท และข้อมูลอื่น ๆ ดังรูปที่ 4-3 (ข) สำหรับวัตถุแต่ละวัตถุที่ปรากฏในภาพ คำบรรยายเหล่านี้ถูกใช้เป็นสี่เหลี่ยมพื้นที่จริงขั้นพื้นฐาน (`Ground truth boxes`) สำหรับการเปรียบเทียบในขั้นตอนการสอนและการทดสอบ



(ก)

```

<annotation verified="yes">
  <folder>images</folder>
  <filename>timber-1027.JPG</filename>
  <path>/home/fedex/THESIS/01_timber_dataset/resize images 600x by adjust the longer length/images/timber-1027.JPG</path>
  <source>
    <database>Unknown</database>
  </source>
  <size>
    <width>600</width>
    <height>401</height>
    <depth>3</depth>
  </size>
  <segmented>0</segmented>
  <object>
    <name>timber</name>
    <pose>Unspecified</pose>
    <truncated>0</truncated>
    <difficult>0</difficult>
    <bndbox>
      <xmin>103</xmin>
      <ymin>49</ymin>
      <xmax>482</xmax>
      <ymax>321</ymax>
    </bndbox>
  </object>
</annotation>

```

(ข)

รูปที่ 4-3 การเตรียมข้อมูลด้วยการบรรยายภาพ (ก) เครื่องมือที่ใช้ในการบรรยายภาพในรูปแบบกรอบสี่เหลี่ยมในการปิดล้อมวัตถุ (ข) ไฟล์โครงสร้าง XML ของภาพที่สอดคล้องกันในรูปที่ 4-3 (ก)

ในชุดข้อมูลแต่ละชุดจำเป็นต้องมีการระบุแผนที่ป้ายกำกับ (Label map) ที่เกี่ยวข้องกับชุดข้อมูลนั้น ๆ ซึ่งทำหน้าที่เป็นตัวแปลงชื่อของข้อมูลประเภท (Label name) เป็นรหัสตัวเลข

(Numeric id) ซึ่งในงานวิจัยนี้มีข้อมูลประเภทเพียงข้อมูลเดียว คือ รถบรรทุกไม้ยูคาลิปตัส ดังนั้น Label map สำหรับการทดลองนี้จะมีโครงสร้างดังรูปที่ 4-4 ซึ่งระบุสิ่งที่ต้องการตรวจจับในภาพ

```
item {
  id: 1
  name: 'timber'
}
```

รูปที่ 4-4 โครงสร้างของไฟล์ Label map สำหรับการตรวจจับรถบรรทุกไม้ยูคาลิปตัส

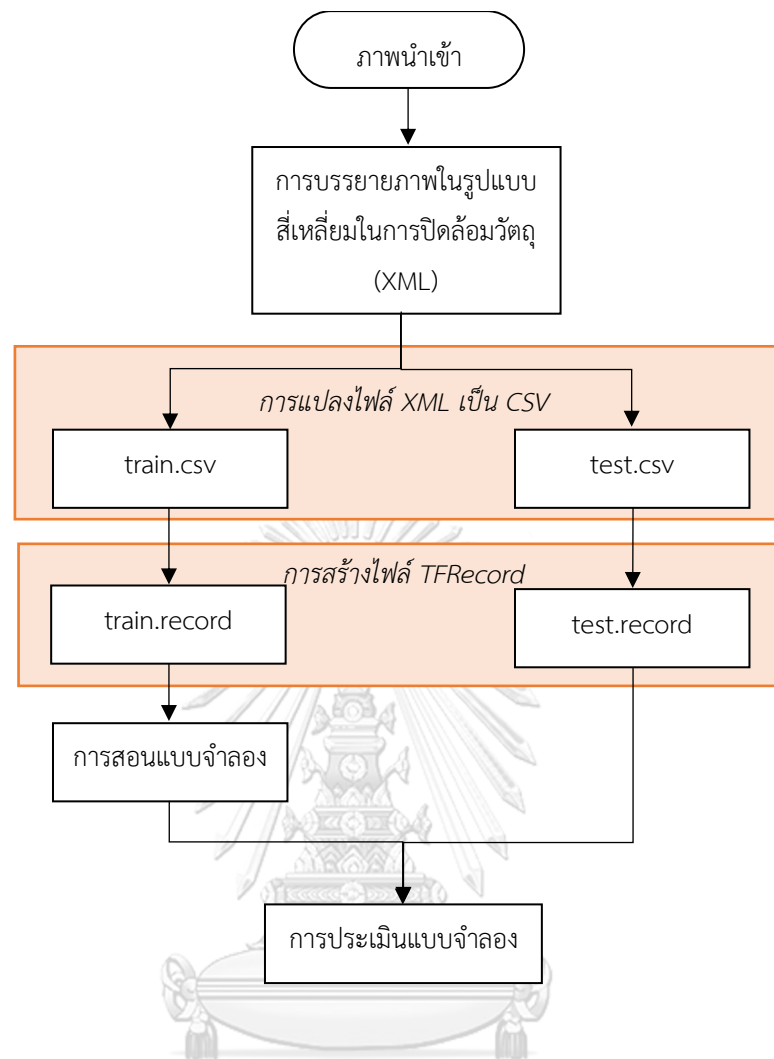
เครื่องมือการเรียนรู้เชิงลึก Tensorflow ในการตรวจจับวัตถุต้องการชุดข้อมูลที่อยู่ในรูปแบบไฟล์ TFRecord ซึ่งจะช่วยให้ประสิทธิภาพในการดึงข้อมูลภาพขาเข้าไปใช้ในการสอน ดังนั้นผู้วิจัยจำเป็นต้องแปลงชุดข้อมูลทั้งหมดเป็นรูปแบบไฟล์ TFRecord ด้วยการแปลงไฟล์ XML เป็นไฟล์ CSV ดังรูปที่ 4-5 จากนั้นจึงดำเนินการสร้างไฟล์ TFRecord จากชุดข้อมูลได้ โดยชุดข้อมูลภาพนำเข้าถูกแบ่งออกเป็นชุดข้อมูลสอนและชุดข้อมูลทดสอบ ดังรูปที่ 4-6 ซึ่งผู้วิจัยใช้ภาพจำนวน 815 ภาพสำหรับการสอน (train.records) และ 400 ภาพสำหรับการทดสอบ (test.records)



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

filename	width	height	class	xmin	ymin	xmax	ymax
timber-34.JPG	400	600	timber	1	85	450	430
timber-436.JPG	600	400	timber	110	4	505	356
timber-942.JPG	600	400	timber	78	46	525	309
timber-667.JPG	600	400	timber	113	25	487	317
timber-1071.JPG	600	400	timber	1	9	362	321
timber-465.JPG	600	400	timber	243	9	528	235
timber-1079.JPG	600	400	timber	69	15	483	351
timber-675.JPG	600	400	timber	6	15	394	349
timber-240.JPG	600	400	timber	1	43	600	324
timber-301.JPG	600	400	timber	43	29	352	317
timber-922.JPG	600	400	timber	1	17	600	365
timber-280.JPG	600	400	timber	244	55	596	338

รูปที่ 4-5 ตัวอย่างการแปลงไฟล์รูปแบบ XML เป็นไฟล์ CSV เพื่อนำไปสร้างไฟล์ TFRecord



รูปที่ 4-6 แผนภาพแสดงการบรรยายภาพและการแบ่งชุดข้อมูลสำหรับการสอน

CHULALONGKORN UNIVERSITY

เมื่อชุดข้อมูล TFRecord ถูกสร้างขึ้นแล้ว ในส่วนนี้จะกล่าวถึงการกำหนดค่าตัวแปร (Parameter configuration) เส้นทางไปยังจุดตรวจสอบ (Model checkpoint) แผนที่ย้ายกำกับ และไฟล์ TFRecord ของแบบจำลองที่ผู้วิจัยใช้ในการทดลองนี้ ซึ่งการกำหนดค่าในการสอนของแบบจำลองแสดงดังตารางที่ 4-1

ตารางที่ 4-1 การตั้งค่าตัวแปรที่เปลี่ยนแปลงสำหรับการสอนแบบจำลอง

การตั้งค่า	ตัวแปร	คำอธิบาย	ค่าตัวแปร
แบบจำลอง	num_classes	จำนวนข้อมูลประเภทในชุดข้อมูล	1
การสอน	batch_size	จำนวนรวมของตัวอย่างการสอนที่มีอยู่ในชุดชุดเดียว	24
	learning_rate	ตัวแปรที่ควบคุมว่ามีการปรับน้ำหนักของโครงข่ายมากน้อยแค่ไหน โดยให้ความสำคัญกับการสูญเสียระดับของความชัน	0.0004
	fine_tune_checkpoint	เส้นทางไปยังจุดตรวจสอบที่ได้รับการสอนล่วงหน้า	/training/
	train_input_config	เส้นทางไปยังไฟล์ TFRecords และแผนที่ป้ายกำกับสำหรับการสอนและการทดสอบ	/data/
	eval_input_config		
omp_num_threads	จำนวนเธรดของหน่วยประมวลผลกลาง	8	

4.3.2 การสอน

สำหรับการทดลองในการสอนนี้ ผู้วิจัยใช้แบบจำลอง SSD MobileNet v1 ที่ได้รับการสอนบนชุดข้อมูล COCO [31] เนื่องจากโครงสร้าง SSD นั้นจะกำหนดความน่าจะเป็นทั้งหมดของสี่เหลี่ยมในการปิดล้อมวัตถุในครั้งเดียว และสำหรับสถาปัตยกรรมนิรอรอลเน็ตเวิร์กผู้วิจัยใช้แบบจำลอง MobileNet ซึ่งออกแบบมาเพื่อใช้งานบนแอปพลิเคชันสมาร์ทโฟนโดยเฉพาะ กระบวนการเริ่มจากการใช้ขั้นสุดท้ายซึ่งมีข้อมูลประเภทจากแบบจำลองที่ได้รับการสอนแล้วแทนที่ด้วยข้อมูลประเภทบรรทัดไม้ยูคาลิปตัส โดยกระบวนการนี้จะใช้การตรวจจับคุณลักษณะที่ได้รับการสอนจากแบบจำลองก่อนหน้านี้ รวมถึงใช้คุณลักษณะของงานวิจัยนี้ในการตรวจจับข้อมูลประเภทใหม่

ในการเริ่มต้นการสอนแบบจำลอง เนื่องจากผู้วิจัยมีการ์ดแสดงผลภาพที่มีหน่วยความจำเหมาะสม จึงตัดสินใจที่จะดำเนินการสอนบนเครื่องคอมพิวเตอร์ โดยเริ่มต้งานด้านการสอนและงานด้านการประเมินแบบจำลองในเวลาเดียวกัน ดังรูปที่ 4-7 รวมถึงเริ่มการทำงานของ Tensorboard เพื่อตรวจสอบประสิทธิภาพของแบบจำลองที่ได้รับการสอนในการทดลองนี้ หลังจากดำเนินการสอน

11,960 รอบโดยใช้เวลา 1 ชั่วโมง 9 นาที 8 วินาที ผู้วิจัยสังเกตการสูญเสียทั้งหมด (Total loss) ลดลงเหลือ 0.7873 และความแม่นยำ (mean Average Precision : mAP) เพิ่มสูงขึ้นถึง 1 และค่าความสูญเสียต่ำกว่า 1 อย่างต่อเนื่อง เมื่อสังเกตภาพใน Tensorboard จะเห็นว่าแบบจำลองนั้นมีความถูกต้องและแม่นยำมากดังรูปที่ 4-8

```

fedex@fedex-ubuntu: ~/THESIS/06 object detection 2018/models-master/research
INFO:tensorflow:global step 1305: loss = 1.2173 (0.293 sec/step)
INFO:tensorflow:global step 1306: loss = 1.2372 (0.291 sec/step)
INFO:tensorflow:global step 1306: loss = 1.2372 (0.291 sec/step)
INFO:tensorflow:global step 1307: loss = 1.3160 (0.325 sec/step)
INFO:tensorflow:global step 1307: loss = 1.3160 (0.325 sec/step)
INFO:tensorflow:global step 1308: loss = 1.3321 (0.304 sec/step)
INFO:tensorflow:global step 1308: loss = 1.3321 (0.304 sec/step)
INFO:tensorflow:global step 1309: loss = 1.0406 (0.290 sec/step)
INFO:tensorflow:global step 1309: loss = 1.0406 (0.290 sec/step)
INFO:tensorflow:global step 1310: loss = 1.1888 (0.290 sec/step)
INFO:tensorflow:global step 1310: loss = 1.1888 (0.290 sec/step)
INFO:tensorflow:global step 1311: loss = 1.0609 (0.287 sec/step)
INFO:tensorflow:global step 1311: loss = 1.0609 (0.287 sec/step)
INFO:tensorflow:global step 1312: loss = 0.9916 (0.296 sec/step)
INFO:tensorflow:global step 1312: loss = 0.9916 (0.296 sec/step)
INFO:tensorflow:global step 1313: loss = 1.0564 (0.284 sec/step)
INFO:tensorflow:global step 1313: loss = 1.0564 (0.284 sec/step)
INFO:tensorflow:global step 1314: loss = 1.0903 (0.281 sec/step)
INFO:tensorflow:global step 1314: loss = 1.0903 (0.281 sec/step)
INFO:tensorflow:global step 1315: loss = 1.4685 (0.299 sec/step)
INFO:tensorflow:global step 1315: loss = 1.4685 (0.299 sec/step)
INFO:tensorflow:global step 1316: loss = 1.1336 (0.313 sec/step)
INFO:tensorflow:global step 1316: loss = 1.1336 (0.313 sec/step)

```

(ก)

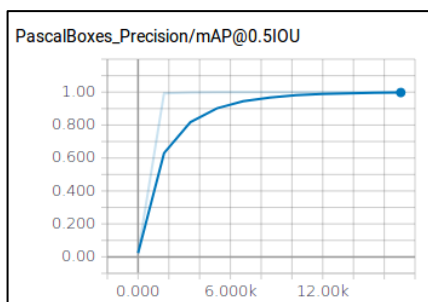
```

fedex@fedex-ubuntu: ~/THESIS/06 object detection 2018/models-master/research
6] tried to allocate 0 bytes
2018-06-12 20:48:03.653848: W tensorflow/core/common_runtime/allocator_retry.cc:
32] Request to allocate 0 bytes
2018-06-12 20:48:03.653855: E tensorflow/core/common_runtime/bfc_allocator.cc:24
6] tried to allocate 0 bytes
2018-06-12 20:48:03.653859: W tensorflow/core/common_runtime/allocator_retry.cc:
32] Request to allocate 0 bytes
2018-06-12 20:48:03.653863: E tensorflow/core/common_runtime/bfc_allocator.cc:24
6] tried to allocate 0 bytes
2018-06-12 20:48:03.653866: W tensorflow/core/common_runtime/allocator_retry.cc:
32] Request to allocate 0 bytes
2018-06-12 20:48:03.653870: E tensorflow/core/common_runtime/bfc_allocator.cc:24
6] tried to allocate 0 bytes
2018-06-12 20:48:03.653875: W tensorflow/core/common_runtime/allocator_retry.cc:
32] Request to allocate 0 bytes
2018-06-12 20:48:03.653902: E tensorflow/core/common_runtime/bfc_allocator.cc:38
1] tried to deallocate nullptr
2018-06-12 20:48:03.653909: E tensorflow/core/common_runtime/bfc_allocator.cc:38
1] tried to deallocate nullptr
2018-06-12 20:48:03.653915: E tensorflow/core/common_runtime/bfc_allocator.cc:38
1] tried to deallocate nullptr
2018-06-12 20:48:03.653920: E tensorflow/core/common_runtime/bfc_allocator.cc:38
1] tried to deallocate nullptr

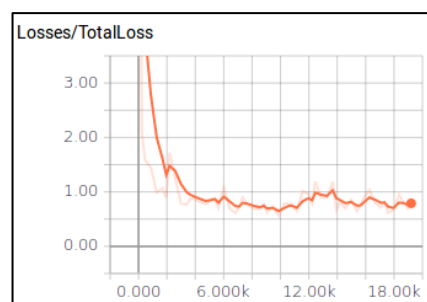
```

(ข)

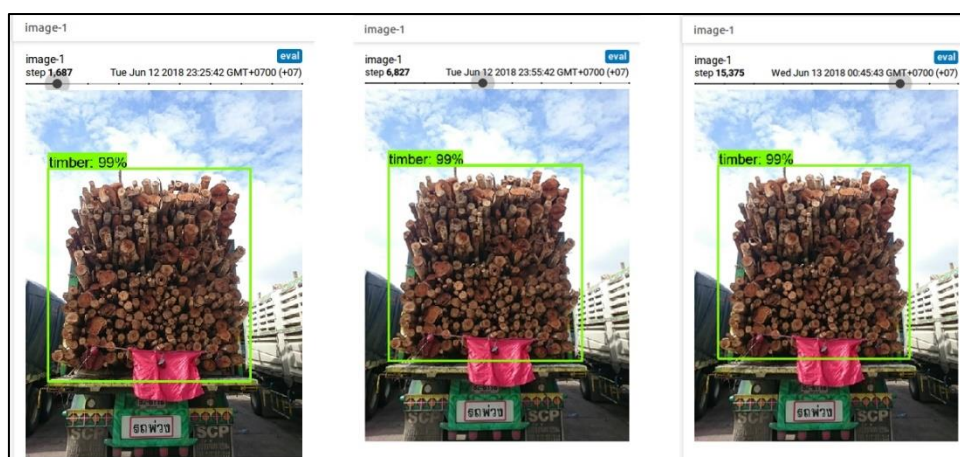
รูปที่ 4-7 การเริ่มงานด้านการสอนและงานด้านการประเมินแบบจำลองซึ่งถูกประมวลผลในเวลาเดียวกัน (ก) งานด้านการสอน (ข) งานด้านการประเมิน



(ก)



(ข)



(ค)

รูปที่ 4-8 ความแม่นยำและการสูญเสียทั้งหมดของแบบจำลอง (ก) กราฟแสดงความแม่นยำ (ข) กราฟแสดงการสูญเสียทั้งหมด (ค) การประเมินผลของภาพตัวอย่างในขณะดำเนินการสอน

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

4.3.3 การทดสอบ

หลังจากเสร็จสิ้นการสอนแล้ว ผู้วิจัยได้ส่งออกแบบจำลองการสอนในรูปแบบกราฟอนูมาน (Frozen inference graph) เพื่อใช้ในการอนุมานสำหรับกระบวนการทดสอบ โดยสำหรับ การตรวจจับรถบรรทุกไม้ยูคาลิปตัส ภาพจำนวน 815 ภาพถูกใช้ในการสอนซึ่งมีรอบสี่เหลี่ยมในการปิดล้อมวัตถุประกอบบนแบบจำลองที่ได้รับการสอนล่วงหน้า SSD Mobilenet v1 COCO และทดสอบ ด้วยชุดข้อมูลทดสอบจำนวน 300 ภาพ ดังรูปที่ 4-9 วิธีที่เสนอ คือ การตรวจจับหน้าตัดของปลาย ท่อนซุงในภาพถ่ายจากท้ายรถบรรทุกไม้ยูคาลิปตัส ในส่วนแรกกระบวนการค้นพบพื้นที่ที่ต้องการ (กรอบสี่เหลี่ยมในการปิดล้อมวัตถุ) ของไม้ยูคาลิปตัสบนรถบรรทุกเป้าหมาย ภาพจำนวน 300 ภาพ

ถูกนำไปทดสอบซึ่งใช้เวลาในการประมวลผล 13 วินาทีต่อภาพ และได้รับค่า Precision Recall และ F-Measure ดังตารางที่ 4-2

ตารางที่ 4-2 ผลลัพธ์ Precision Recall และ F-Measure จากการตรวจจับพื้นที่ที่สนใจของรถบรรทุกไม้อยู่คาลิปตัส

การตรวจจับพื้นที่ที่สนใจ (กรอบสี่เหลี่ยมในการปิดล้อมวัตถุ) ของรถบรรทุกไม้อยู่คาลิปตัส		
<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>
1	0.98	0.99



รูปที่ 4-9 ภาพผลลัพธ์ของวิธีการตรวจจับรถบรรทุกไม้ยูคาลิปด้วยด้วยโครงสร้าง SSD (ก) ภาพทดสอบ (ข) ภาพผลลัพธ์จากการทดสอบ (ค) ภาพผลลัพธ์จากการตัดภาพส่วนเฉพาะด้านท้ายรถบรรทุกไม้ในภาพโดยอ้างอิงจากพิกัดของกรอบสี่เหลี่ยมในการปิดล้อมวัตถุ

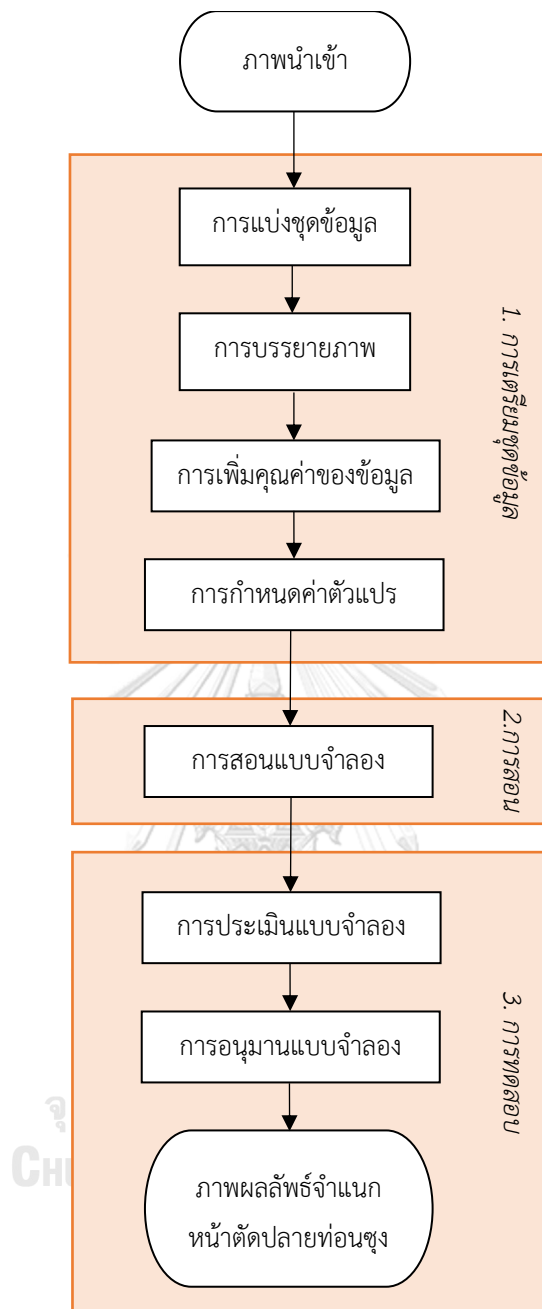
4.4 การแยกส่วนหน้าตัดของปลายท่อนซุงด้วยโครงข่าย VGG16

ในการทดลองการแยกส่วนหน้าตัดของปลายท่อนซุงนั้นมีความสำคัญต่อระบบที่เสนอ เนื่องจากในขั้นตอนนี้จะสามารถช่วยแยกส่วนจุดภาพที่ต้องการ คือ พื้นที่จุดภาพของหน้าตัดของปลายท่อนซุงออกจากพื้นหลังและส่วนที่ไม่ต้องการออกจากภาพ วิธีการที่เสนอในขั้นตอนนี้ใช้วิธีการจำแนกทางความหมายตามโครงข่ายคอนโวลูชันแบบทั่วถึง (Fully convolutional network-based semantic segmentation) ด้วยแบบจำลองของโครงข่าย VGG16 ที่ได้รับการสอนบนชุดข้อมูล Imagenet [22] ในการแยกส่วนจุดภาพหน้าตัดของปลายท่อนซุงในภาพที่ได้รับการสอนด้วยชุดข้อมูลรถบรรทุกไม้ยูคาลิปตัส โดยขั้นตอนการทดลองในการแยกส่วนหน้าตัดของปลายท่อนซุงแสดงดังรูปที่ 4-10

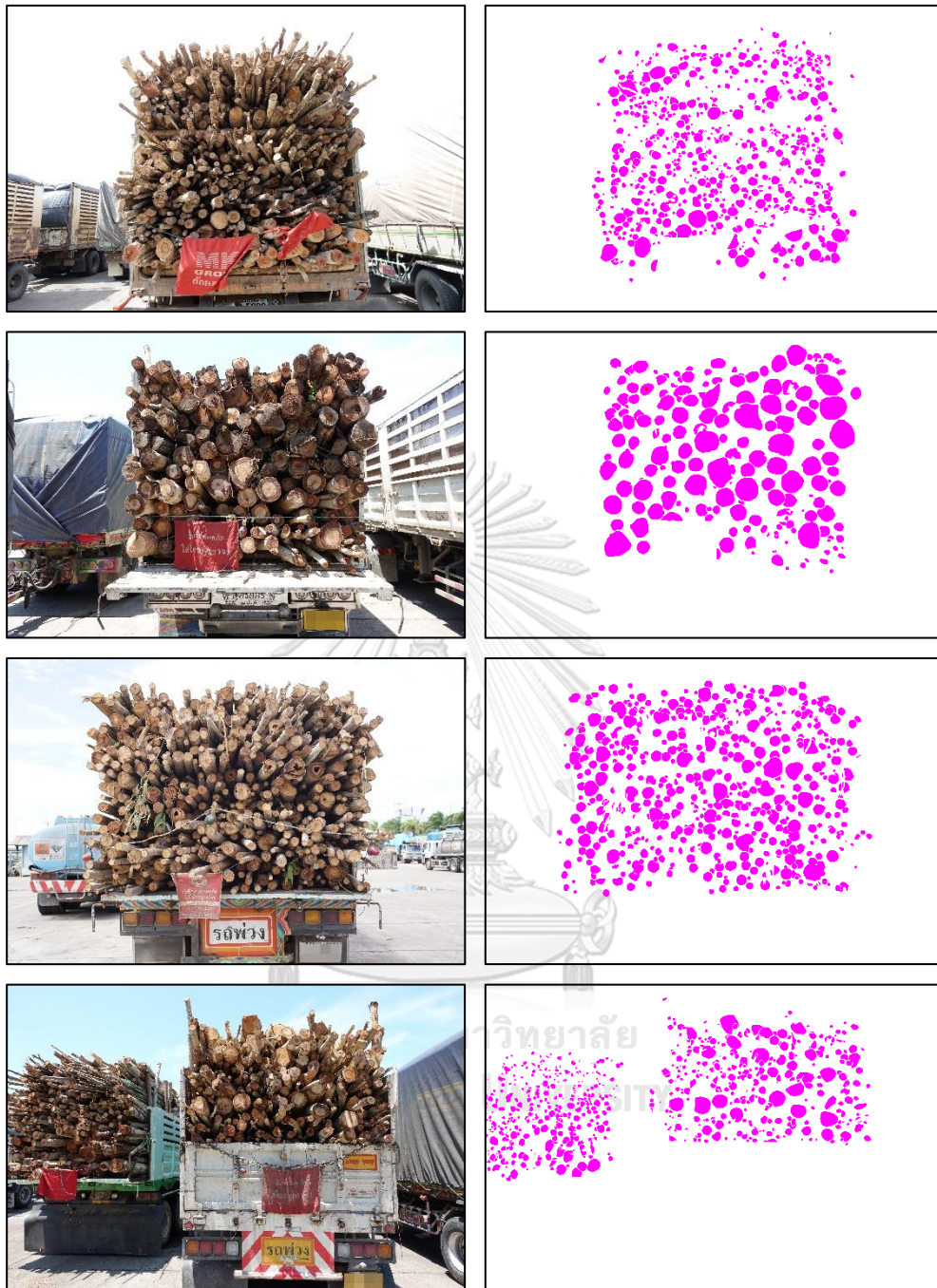
4.4.1 การเตรียมข้อมูล

ขั้นตอนแรกในการทดลองเป็นการรวมรวมภาพนำเข้าไปในลักษณะเช่นเดียวกันข้อ 4.3.1 แต่ในทางกลับกันภาพนำเข้าทั้งหมดถูกย่อขนาดด้วยฟังก์ชัน `resize` ในไลบรารี `OpenCV` ให้เหลือขนาด 1632×1088 จุดภาพเพื่อรักษาอัตราส่วนของภาพและยังคงมีรายละเอียดของภาพที่เพียงพอ โดยในการทดลองนี้ชุดข้อมูลรถบรรทุกไม้ยูคาลิปตัสถูกแบ่งออกเป็นชุดข้อมูลสอนจำนวน 50 ภาพที่ติดป้ายกำกับข้อมูลประเภทจำนวน 2 ประเภท คือ พื้นที่ที่เป็นหน้าตัดของปลายท่อนซุงและพื้นที่ที่ไม่ใช่หน้าตัดปลายท่อนซุงดังรูปที่ 4-11 และสำหรับชุดข้อมูลทดสอบถูกแบ่งออกเป็น 300 ภาพซึ่งชุดข้อมูลทั้ง 2 ชุดมีรายละเอียดหน้าตัดของปลายท่อนซุงในภาพที่แตกต่างกันไปในช่วงเวลากลางวัน

เนื่องจากชุดข้อมูลสอนที่ถูกติดป้ายกำกับข้อมูลประเภทจำเป็นต้องใช้ระยะเวลาประมาณ 12 ชั่วโมงต่อภาพในการติดป้ายกำกับในแต่ละภาพ การเพิ่มคุณค่าของข้อมูล (Data augmentation) ซึ่งเป็นแนวทางปฏิบัติมาตรฐานในด้านคอมพิวเตอร์วิทัศน์จึงถูกนำมาใช้เพื่อเพิ่มชุดข้อมูลการสอน ปรับปรุงประสิทธิภาพของโครงข่ายและเป็นการเตรียมตัวอย่างภาพให้มีความหลากหลายขึ้นจากชุดข้อมูลด้วยวิธีการแปลงข้อมูลภาพพื้นฐาน (Basic transformation) ในฟังก์ชัน `OpenCV` ดังรูปที่ 4-12 โดยเฉพาะอย่างยิ่งในกรณีของผู้วิจัยที่มีจำนวนตัวอย่างการสอนค่อนข้างน้อย สำหรับการแยกส่วนหน้าตัดของปลายท่อนซุง เทคนิคการเพิ่มคุณค่าของข้อมูลถูกนำมาใช้กับภาพในชุดข้อมูลสอนจำนวน 50 ภาพซึ่งมีการติดป้ายชื่อกำกับและได้รับผลลัพธ์จำนวน 550 ภาพสำหรับนำไปใช้ในกระบวนการสอน



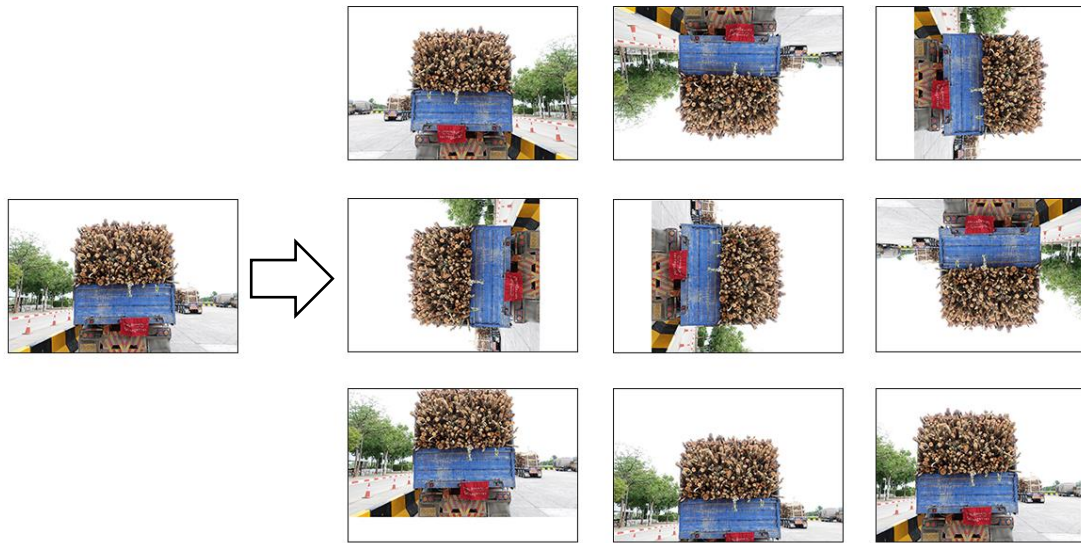
รูปที่ 4-10 แผนภาพแสดงการทดลองการแยกส่วนหน้าตัดของปลายท่อนซุงด้วยแบบจำลอง VGG16



(ก)

(ข)

รูปที่ 4-11 การบรรยายภาพด้วยการติดป้ายชื่อกำกับข้อมูลประเภท ได้แก่ หน้าตัดของปลายท่อนซุง และที่ไม่ใช่หน้าตัดของปลายท่อนซุง (ก) ภาพนำเข้าในชุดข้อมูลสอน (ข) ภาพที่ได้รับการติดป้ายกำกับข้อมูลประเภท



รูปที่ 4-12 ตัวอย่างการเพิ่มคุณค่าของข้อมูลในชุดข้อมูลรถบรรทุกไม้ยูคาลิปตัส

หลังจากการเตรียมข้อมูลภาพนำเข้าเสร็จสิ้น ในขั้นตอนสุดท้ายของการเตรียมข้อมูลก่อนเริ่มกระบวนการสอนแบบจำลองเป็นการกำหนดค่าตัวแปรที่สำคัญในการสอนของแบบจำลอง ดังตารางที่ 4-3

ตารางที่ 4-3 การตั้งค่าตัวแปรที่เปลี่ยนแปลงสำหรับการสอนแบบจำลอง

การตั้งค่า	ตัวแปร	คำอธิบาย	ค่าตัวแปร
แบบจำลอง	num_classes	จำนวนข้อมูลประเภทในชุดข้อมูล	2
การสอน	batch_size	จำนวนรวมของตัวอย่างการสอนที่มีอยู่ในชุดชุดเดียว	2
	learning_rate	ตัวแปรที่ควบคุมว่ามีการปรับน้ำหนักของโครงข่ายมากน้อยแค่ไหนโดยให้ความสำคัญกับการสูญเสียระดับของความชัน	0.001
	epochs	จำนวนรอบในกระบวนการเรียนรู้	140
	image_size	ขนาดภาพนำเข้าและภาพผลลัพธ์	1632 x 1088

ตารางที่ 4-3 การตั้งค่าตัวแปรที่เปลี่ยนแปลงสำหรับการสอนแบบจำลอง (ต่อ)

การตั้งค่า	ตัวแปร	คำอธิบาย	ค่าตัวแปร
การสอน	label_colors	พจนานุกรมแผนที่ของจำนวนข้อมูลประเภทที่แทนที่ด้วยค่าสีซึ่งใช้สำหรับการแยกส่วน	0: [0, 0, 0] 1: [255, 0, 255]
	num_training	จำนวนตัวอย่างในการสอน	550
	num_validation	จำนวนตัวอย่างในการตรวจสอบ	55

4.4.2 การสอน

การทดลองในการสอนนี้ ผู้วิจัยใช้เทคนิคการถ่ายโอนการเรียนรู้บนโครงข่ายคอนโวลูชันแบบทั่วถึงด้วยแบบจำลองของโครงข่าย VGG16 ที่ได้รับการสอนล่วงหน้าโดย Udacity [34] บนชุดข้อมูล Imagenet [22] เนื่องจากโครงข่ายคอนโวลูชันแบบทั่วถึงนั้นสามารถระบุจุดภาพพื้นที่ที่เป็นหน้าตัดของปลายท่อนซุงและพื้นที่ที่ไม่ใช่หน้าตัดของปลายท่อนซุงในภาพได้ กระบวนการเริ่มจากการนำชั้น 3 ชั้น 4 และชั้น 7 ของโครงข่าย VGG มารวมกันดังที่แสดงในบทความงานวิจัยของ Long และคณะ [18] จากนั้นใช้คอนโวลูชันขนาด 1×1 ที่บริเวณด้านบนของแต่ละชั้นสกัดเพื่อทำหน้าที่เป็นตัวแบ่งประเภทแบบท้องถิ่น เนื่องจากมีผลลัพธ์บางส่วนที่ยังคงมีขนาดเล็กกว่าภาพนำเข้า 32, 16 และ 8 เท้า ดังนั้นจึงจำเป็นต้องเพิ่มขนาดภาพผลลัพธ์ให้เท่ากับภาพนำเข้า สุดท้ายการเพิ่มค่าน้ำหนักถูกนำมาใช้เพื่อให้ได้ผลลัพธ์ ในขั้นแรกของนิรอลเน็ตเวิร์กคอนโวลูชันจะเรียนรู้คุณลักษณะทั่วไปในภาพ ส่วนขั้นที่ถูกลบไปของโครงข่ายนั้นจะเรียนรู้รูปร่างเฉพาะรวมถึงส่วนที่ได้รับการสอนนี้ คือ พื้นที่จุดภาพหน้าตัดของปลายท่อนซุงไม้ยูคาลิปตัส

ผู้วิจัยเริ่มต้นการสอนแบบจำลองบนเครื่องคอมพิวเตอร์เนื่องจากการ์ดแสดงผลภาพที่มีหน่วยความจำเหมาะสม โดยเริ่มต้นงานด้านการสอนและงานด้านการตรวจสอบแบบจำลองในเวลาเดียวกัน ดังรูปที่ 4-13 และเริ่มการทำงานของ Tensorboard เพื่อตรวจสอบประสิทธิภาพของแบบจำลองที่ได้รับการสอนในการทดลองนี้ หลังจากดำเนินการสอนบนแบบจำลองที่ได้รับการสอนล่วงหน้าของโครงข่าย VGG16 จำนวน 140 รอบโดยใช้เวลาเฉลี่ย 7.44 นาทีต่อรอบ ผู้วิจัยสังเกตการสูญเสียการสอน (Training loss) ลดลงเหลือ 0.02013 และการสูญเสียการตรวจสอบ (Validation loss) ลดลงเหลือ 0.07492 จากนั้นจึงทำการหยุดการสอนที่รอบการสอนนี้ เนื่องจากความสูญเสียในการตรวจสอบนั้นสูงขึ้นดังรูปที่ 4-14

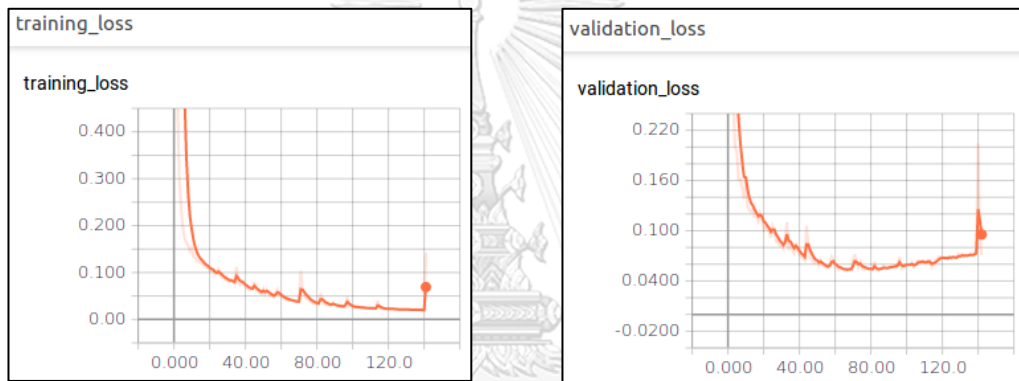
```

fedex@fedex-ubuntu: ~/THESIS/02_image-segmentation-fcn-master
[i] Epoch 92/1000: 100% | ██████████ | 312/312 [07:26<00:00, 1.43s/batches]
[i] Epoch 93/1000: 100% | ██████████ | 312/312 [07:26<00:00, 1.43s/batches]
[i] Epoch 94/1000: 100% | ██████████ | 312/312 [07:26<00:00, 1.43s/batches]
[i] Epoch 95/1000: 100% | ██████████ | 312/312 [07:26<00:00, 1.43s/batches]
Checkpoint saved: test/e95.ckpt
[i] Epoch 96/1000: 100% | ██████████ | 312/312 [07:26<00:00, 1.43s/batches]
[i] Epoch 97/1000: 100% | ██████████ | 312/312 [07:26<00:00, 1.43s/batches]
[i] Epoch 98/1000: 100% | ██████████ | 312/312 [07:26<00:00, 1.43s/batches]
[i] Epoch 99/1000: 100% | ██████████ | 312/312 [07:26<00:00, 1.43s/batches]
[i] Epoch 100/1000: 100% | ██████████ | 312/312 [07:26<00:00, 1.43s/batches]
Checkpoint saved: test/e100.ckpt
[i] Epoch 101/1000: 100% | ██████████ | 312/312 [07:26<00:00, 1.43s/batches]
[i] Epoch 102/1000: 100% | ██████████ | 312/312 [07:28<00:00, 1.44s/batches]
[i] Epoch 103/1000: 100% | ██████████ | 312/312 [07:26<00:00, 1.43s/batches]
[i] Epoch 104/1000: 100% | ██████████ | 312/312 [07:26<00:00, 1.43s/batches]
[i] Epoch 105/1000: 100% | ██████████ | 312/312 [07:26<00:00, 1.43s/batches]
Checkpoint saved: test/e105.ckpt
[i] Epoch 106/1000: 100% | ██████████ | 312/312 [07:26<00:00, 1.43s/batches]
[i] Epoch 107/1000: 100% | ██████████ | 312/312 [07:26<00:00, 1.43s/batches]
[i] Epoch 108/1000: 100% | ██████████ | 312/312 [07:28<00:00, 1.44s/batches]
[i] Epoch 109/1000: 100% | ██████████ | 312/312 [07:26<00:00, 1.43s/batches]
[i] Epoch 110/1000: 100% | ██████████ | 312/312 [07:25<00:00, 1.43s/batches]
Checkpoint saved: test/e110.ckpt
[i] Epoch 111/1000: 34% | ████████ | 106/312 [02:31<04:54, 1.43s/batches]

```

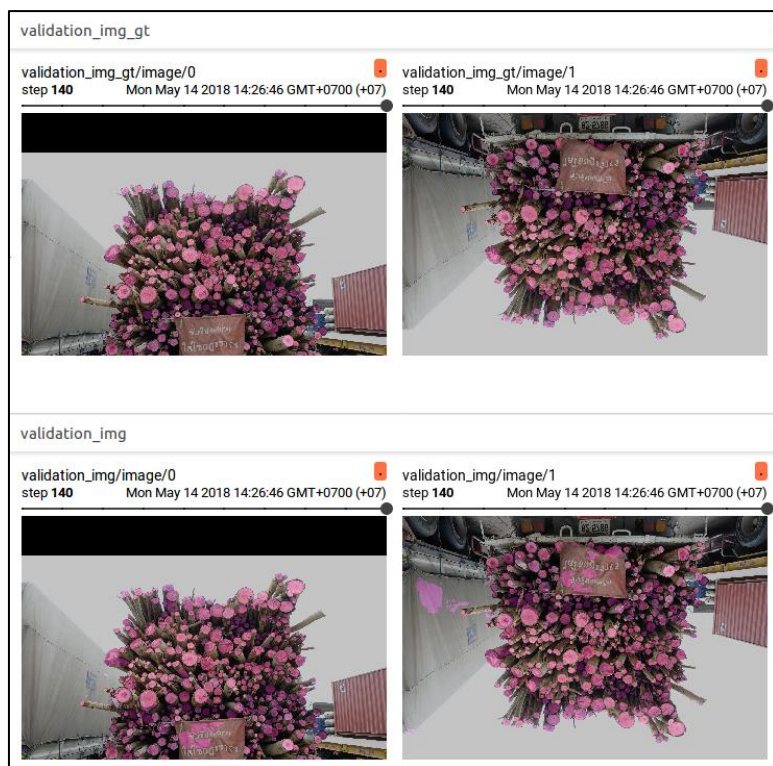
รูปที่ 4-13 การเริ่มงานด้านการสอนและงานด้านการตรวจสอบแบบจำลองซึ่งถูกประมวลผลในเวลา

เดียวกัน



(ก)

(ข)



(ค)

รูปที่ 4-14 การสูญเสียทั้งหมดของแบบจำลอง (ก) กราฟแสดงการสูญเสียของการสอน (ข) กราฟแสดงการสูญเสียของการตรวจสอบ (ค) การประเมินผลของภาพตัวอย่างของการสอนและการตรวจสอบในขณะที่ดำเนินการสอนแบบจำลอง

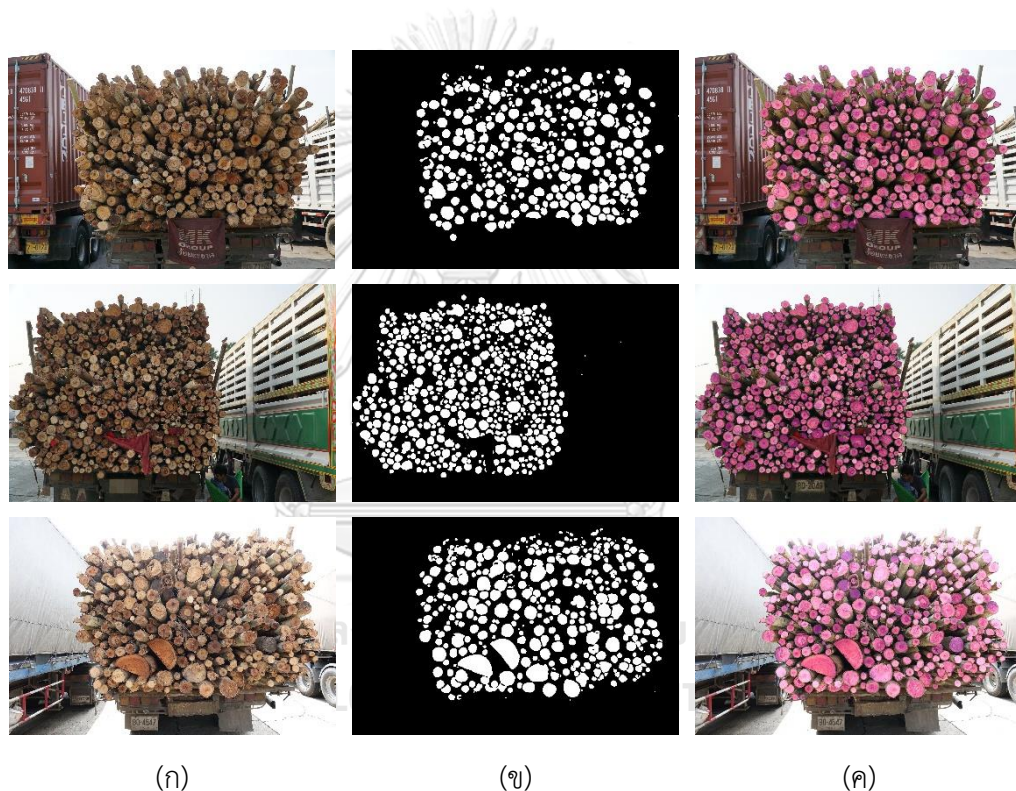
จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

4.4.3 การทดสอบ

ในส่วนที่ 2 ของระบบนี้เป็นการดำเนินการของขั้นตอนการแยกส่วนหน้าตัดของปลายท่อนซุง และการนับซุง โดยผลลัพธ์ของการนับปลายท่อนซุงถูกตรวจสอบโดยการเปรียบเทียบระหว่างวิธีการที่เสนอและการนับด้วยสายตามนุษย์ ชุดข้อมูลการทดสอบจำนวน 300 ภาพได้รับการทดสอบกับภาพรถบรรทุกไม้ยูคาลิปตัสสำหรับการแยกส่วนหน้าตัดของปลายท่อนซุงซึ่งใช้เวลาในการประมวลผล 7.4 วินาทีต่อภาพ โดยถูกนำมาตรวจสอบความแม่นยำและได้ผลลัพธ์ของการคำนวณค่า Precision Recall และ F-Measure แสดงในตารางที่ 4-4

ตารางที่ 4-4 ผลลัพธ์ Precision Recall และ F-Measure จากการแยกส่วนพื้นที่ที่สนใจหน้าตัดของ
ปลายท่อนซุง

พื้นที่ที่สนใจของการแยกส่วนหน้าตัดของปลาย ท่อนซุง (พื้นที่ของจุดภาพที่ติดป้ายชื่อ)		
<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>
0.97	0.97	0.97



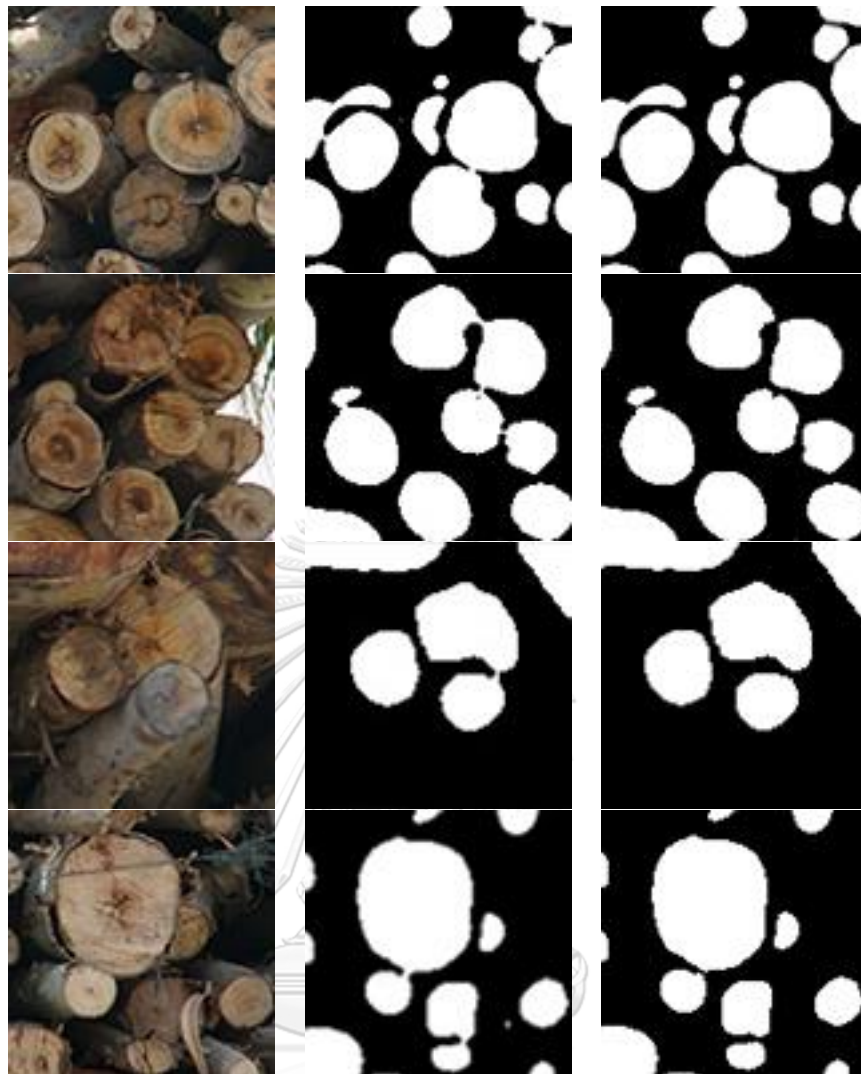
รูปที่ 4-15 ภาพผลลัพธ์ของวิธีการแยกส่วนหน้าตัดของปลายท่อนซุงด้วยแบบจำลอง VGG16 (ก) ภาพทดสอบ (ข) ภาพผลลัพธ์ของกระบวนการแยกส่วนหน้าตัดของปลายท่อนซุง (ค) ภาพผลลัพธ์จากการแยกส่วนหน้าตัดของปลายท่อนซุงวางทับบนภาพทดสอบ



รูปที่ 4-15 ภาพผลลัพธ์ของวิธีการแยกส่วนหน้าตัดของปลายท่อนซุงด้วยแบบจำลอง VGG16 (ก) ภาพทดสอบ (ข) ภาพผลลัพธ์ของกระบวนการแยกส่วนหน้าตัดของปลายท่อนซุง (ค) ภาพผลลัพธ์จากการแยกส่วนหน้าตัดของปลายท่อนซุงวางทับบนภาพทดสอบ (ต่อ)

4.5 การหาค่าแกนกลางที่เหมาะสมสำหรับการเปลี่ยนแปลงลักษณะโครงสร้างของภาพ

ในการทดลองการเปลี่ยนแปลงลักษณะโครงสร้างของภาพนั้นมีความจำเป็นต่อระบบ เนื่องจากภาพผลลัพธ์จากกระบวนการแยกส่วนหน้าตัดของปลายท่อนซุงพบว่า มีหน้าตัดของปลายท่อนซุงบางท่อนมีส่วนที่เชื่อมติดกันดังรูปที่ 4-16 (ข) ซึ่งเป็นเหตุให้ผลลัพธ์ของการนับปลายท่อนซุงคลาดเคลื่อนได้ ดังนั้นในการทดลองนี้เป็นการหาค่าแกนกลาง (Kernel) ที่เหมาะสมของกระบวนการโอเพนนิ่งในการเปลี่ยนแปลงลักษณะโครงสร้างของภาพ เพื่อเปิดช่องว่างระหว่างหน้าตัดของปลายท่อนซุงที่เชื่อมติดกัน ผู้วิจัยดำเนินการทดลองโดยใช้ฟังก์ชัน MORPH_OPEN กับองค์ประกอบโครงสร้างแบบ Elliptical-shaped ที่มีขนาดแกนกลาง (Kernel) เท่ากับ 7×7 และดำเนินการทดลองในขั้นตอนการเปลี่ยนแปลงลักษณะโครงสร้างของภาพกับภาพทดสอบทั้งหมดด้วยขนาดแกนกลางดังกล่าว และได้ผลลัพธ์ดังรูปที่ 4-16 (ค)



(ก) จุฬาลงกรณ์มหาวิทยาลัย (ข) วิทยาลัย

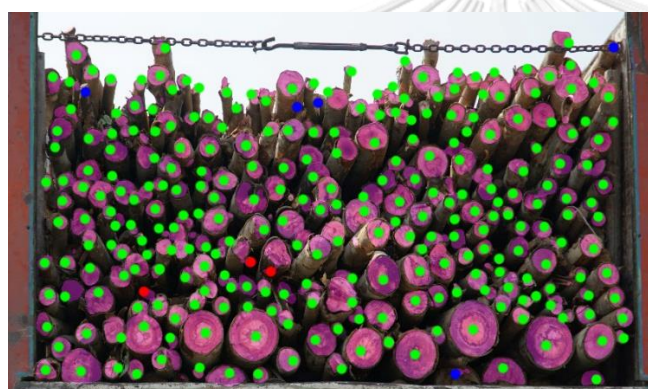
(ค)

CHULALONGKORN UNIVERSITY

รูปที่ 4-16 ตัวอย่างกรณีของหน้าตัดของปลายท่อนซุงที่เชื่อมติดกัน (ก) ภาพทดสอบ (ข) ภาพผลลัพธ์การแยกส่วนหน้าตัดของปลายท่อนซุง (ค) ภาพผลลัพธ์ของกระบวนการโอเพนนิ่งในการเปลี่ยนแปลงลักษณะโครงร่างของภาพ

4.6 ผลการทดลอง

ภาพทดสอบจำนวน 300 ภาพถูกนำมาทดสอบการนับหน้าตัดปลายท่อนซุงด้วยการเปรียบเทียบระหว่างวิธีการที่เสนอและการนับด้วยสายตามนุษย์แสดงตัวอย่างการนับดังรูปที่ 4-17 ซึ่งผู้วิจัยได้วาดวงกลมสีต่าง ๆ เพื่อตรวจสอบว่าวิธีการที่เสนอนั้นมีความถูกต้องเมื่อเปรียบเทียบกับการนับด้วยสายตามนุษย์มากน้อยแค่ไหน จากนั้นภาพทดสอบถูกนำมาตรวจสอบความแม่นยำโดยวิธีการวัดประสิทธิภาพของการจำแนกประเภท ซึ่งได้ผลลัพธ์ของจำนวนหน้าตัดที่นับได้และนับไม่ได้ทั้งหมด ดังตารางที่ 4-5 และมีความถูกต้องเฉลี่ย 94.45% ในการแยกส่วนหน้าตัดของปลายท่อนซุง ดังแสดงในตารางที่ 4-6 และมีค่าลบเท็จ (False negative) เฉลี่ย 2.71% และค่าลบจริง (False positive) เฉลี่ย 2.84%



- ตรวจสอบว่ามีไม้และมีไม้จริง ๆ
(True Positive)
- ตรวจสอบว่ามีไม้ แต่จริง ๆ ไม่มีไม้
(False Positive)
- ตรวจสอบไม่ได้ แต่จริง ๆ มีไม้
(False Negative)
- ตรวจสอบไม่ได้และไม่มีไม้จริง ๆ
(True Negative)

รูปที่ 4-17 การตรวจสอบวิธีที่เสนอเปรียบเทียบกับสายตามนุษย์

ตารางที่ 4-5 ผลลัพธ์ของจำนวนหน้าตัดที่นับได้และนับไม่ได้ทั้งหมด

	ตรวจสอบว่ามีไม้ และมีไม้จริง ๆ (True Positive)	ตรวจสอบว่ามีไม้ แต่จริง ๆ ไม่มีไม้ (False Positive)	ตรวจสอบไม่ได้ แต่จริง ๆ มีไม้ (False Negative)	ตรวจสอบไม่ได้และ ไม่มีไม้จริง ๆ (True Negative)
จำนวนหน้าตัด ทั้งหมดจากภาพ ทดสอบ 300 ภาพ	90,402	2,796	2,519	0

ตารางที่ 4-6 ผลลัพธ์ของการนับหน้าตัดของปลายท่อนซุง

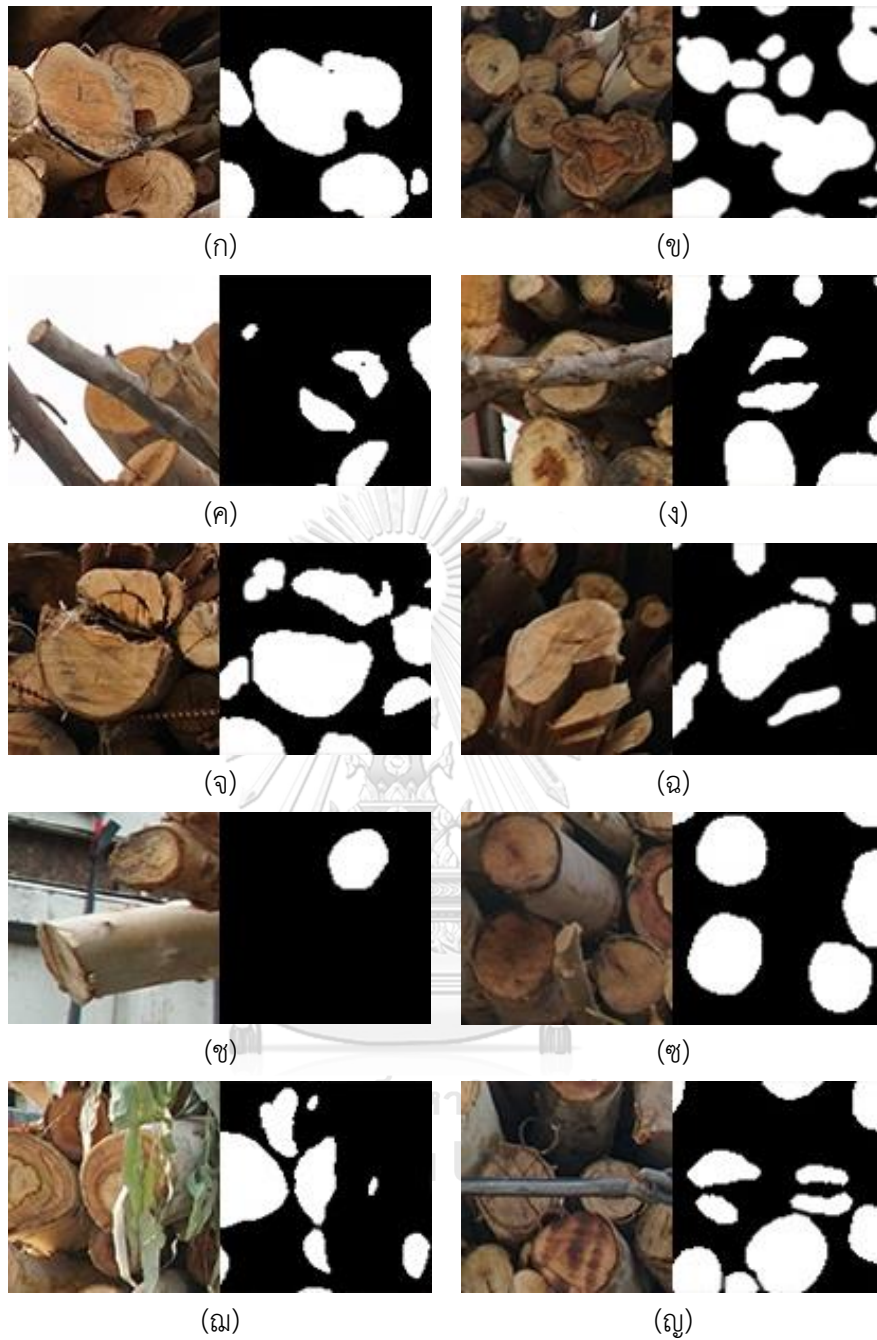
การนับปลายท่อนซุง (ขั้นตอนการแยกส่วน)	
จำนวนภาพที่ได้รับ การแยกส่วน	ความแม่นยำ
300	94.45%

4.7 การวิเคราะห์ผลการทดลอง

ผลลัพธ์ของการแยกส่วนจากภาพตัวอย่างในตารางที่ 4-5 มีความพึงพอใจมากเมื่อเปรียบเทียบกับผลการนับด้วยสายตามนุษย์ อย่างไรก็ตามมีการแยกส่วนปลายท่อนซุงบางส่วนซึ่งได้รับผลกระทบและไม่เป็นไปตามผลลัพธ์ที่ผู้วิจัยได้คาดหวังไว้ ได้แก่

1. ปลายท่อนซุงที่ติดกันบางท่อนไม่สามารถแยกออกได้ เนื่องจากสีของพื้นผิวหน้าตัดที่ต่อเนื่องกัน ทำให้นับจำนวนไม้ได้น้อยลง และน้อยกว่าความเป็นจริง
2. ปลายท่อนซุงที่มีขนาดเล็กพาดผ่านปลายท่อนซุงอื่น ๆ ทำให้นับจำนวนไม้ได้มากขึ้น และเกินความเป็นจริง
3. การแตกหักของหน้าตัดปลายไม้ทำให้นับจำนวนไม้ไม่นับได้มากขึ้นและเกินความเป็นจริง
4. หน้าตัดของปลายท่อนซุงถูกปกคลุมด้วยวัตถุบางอย่างทำให้ตรวจจับไม่เจอ ทำให้นับจำนวนไม้ไม่นับได้มากขึ้นและเกินความเป็นจริง
5. ระนาบของปลายไม้ไม่ตั้งฉากกับแนวแกนเลนส์ของกล้อง ทำให้ตรวจจับไม่เจอ และนับได้น้อยกว่าความเป็นจริง
6. ปลายท่อนซุงอยู่ลึกเข้าไปมาก ทำให้ส่วนพื้นที่หน้าตัดมืด และตรวจจับไม่เจอ

นอกจากนี้ยังมีการแยกส่วนปลายท่อนซุงบางส่วนที่ไม่ได้เป็นหน้าตัดของปลายท่อนซุงที่ปรากฏขึ้นซึ่งส่งผลให้เป็นผลบวกเท็จ (False positive) ยิ่งไปกว่านั้นในบางส่วนปรากฏผลลบลง (False negative) เกิดขึ้นเมื่อน้ำตัดของปลายท่อนซุงถูกปกคลุมด้วยวัตถุบางอย่างและหน้าตัดอยู่ห่างจากแนวตั้งฉากกับแกนของเลนส์กล้องดิจิทัล ดังรูปที่ 4-18 แสดงปัญหาที่ผู้วิจัยได้กล่าวถึงเมื่อข้างต้น



รูปที่ 4-18 ตัวอย่างของภาพนำเข้าและภาพการแยกส่วนหน้าตัดของปลายท่อนซุงที่พบปัญหา (ก) และ (ข) ปลายท่อนซุงที่เชื่อมติดกัน (ค) และ (ง) การซ้อนทับของปลายท่อนซุง (จ) และ (ฉ) การแตกของหน้าตัดของปลายท่อนซุง (ช) และ (ซ) ทิศทางของปลายท่อนซุง (ฌ) และ (ญ) วัตถุที่ไม่พึงประสงค์ปกคลุมบนหน้าตัดของปลายท่อนซุง

บทที่ 5

สรุปการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

ในงานวิจัยนี้ผู้วิจัยได้นำเสนอวิธีการตรวจจับหน้าตัดของปลายท่อนซุงไม้ยูคาลิปตัสในรถบรรทุกไม้ที่มีขนาดท่อนซุงที่แตกต่างกันรวมถึงหน้าตัดที่หลากหลายรูปแบบ โดยใช้วิธีการตรวจจับวัตถุในการตรวจจับรถบรรทุกไม้ยูคาลิปตัส การจำแนกทางความหมายเพื่อแยกหน้าตัดของปลายท่อนซุงออกจากพื้นหลังรวมถึงวัตถุที่ไม่ต้องการออกจากภาพ และเทคนิคการประมวลผลภาพดิจิทัลในการนับจำนวนปลายท่อนซุงทั้งหมด ซึ่งระบบนี้แสดงให้เห็นถึงความทนทานสำหรับใช้กับไม้ยูคาลิปตัสที่มีขนาดต่าง ๆ ซึ่งบรรทุกมาบนรถรวมถึงการวางท่อนซุงที่ไม่เป็นระเบียบบนรถบรรทุกไม้อีกด้วย

ภาพที่ใช้เป็นตัวอย่างสำหรับการสอนในโครงข่าย SSD ใช้ทั้งหมด 815 ภาพ และในโครงข่าย VGG ใช้ทั้งหมด 550 ภาพ และภาพสำหรับทดสอบทั้งหมด 300 ภาพ โดยระบบสามารถแยกส่วนหน้าตัดของปลายท่อนซุงไม้ยูคาลิปตัสได้โดยไม่มีการรบกวนและบรรลุผลลัพธ์ได้ด้วยความแม่นยำเฉลี่ยร้อยละ 94.45

ระบบที่เสนอนี้ใช้ได้ดีกับภาพที่ถูกบันทึกจากด้านท้ายของรถบรรทุกไม้ยูคาลิปตัสในช่วงกลางวันเวลาโดยประมาณ 08:00 – 17:00 น. ที่มีมุมของกล้องในการบันทึกภาพเป็นมุมประมาณแนวระดับกับกลางกลุ่มท่อนซุง การนำระบบนี้ไปใช้สามารถใช้ภาพที่มีความละเอียดสูงกว่าหรือต่ำกว่า 1632×1088 จุดภาพได้ โดยระบบจะแปลงภาพให้ทำงานกับระบบได้ที่ขนาดประมาณ 1632×1088 จุดภาพ ในกรณีที่ภาพมีความละเอียดต่ำกว่าที่ระบุไว้ ระบบจะปรับยืดภาพให้สามารถทำงานได้ แต่รายละเอียดของภาพจะน้อยลง ซึ่งอาจส่งผลต่อความแม่นยำ อย่างไรก็ตามระบบที่เสนอยังคงมีความผิดพลาดในการตรวจจับหน้าตัดของปลายท่อนซุงบางท่อนที่เชื่อมติดกัน การซ้อนทับของปลายท่อนซุงรวมถึงวัตถุที่ไม่พึงประสงค์อื่น ๆ การจัดวางท่อนซุงบนรถบรรทุกอย่างไม่เป็นระเบียบ และการแตกของหน้าตัดของปลายท่อนซุง เนื่องจากเป็นปัจจัยภายนอกที่ไม่สามารถควบคุมได้ทำให้ระบบเกิดความผิดพลาดในการตรวจจับหน้าตัดของปลายท่อนซุง

ระบบดังกล่าวคาดว่าจะช่วยลดเวลาในการทำงานและต้นทุนแรงงานในการประมาณจำนวนและราคาของท่อนซุงได้ อีกทั้งยังสามารถนำไปติดตั้งและประยุกต์ใช้กับโรงงานผลิตกระดาษได้ ในอนาคตผู้วิจัยคาดว่าจะระบบนี้จะสามารถเพิ่มความถูกต้องของการนับและช่วยเป็นข้อมูลในการวัดขนาด

ของปลายท่อนซุงและสามารถคำนวณการกระจายขนาดของหน้าตัดของปลายท่อนซุงไม้ยูคาลิปตัสได้ และนับว่าเป็นงานที่ทำหายสำหรับรูปร่างของพื้นที่ของปลายท่อนซุงที่ไม่มีลักษณะคล้ายวงกลม

5.2 ข้อเสนอแนะ

จากการทดลองวิธีการที่เสนอในงานวิจัยนี้พบว่ายังมีสิ่งที่สามารถปรับปรุงได้ เพื่อให้วิธีการที่เสนอทำงานให้มีประสิทธิภาพที่ดียิ่งขึ้นได้ ดังนี้

1. การนำวิธีที่เสนอในงานวิจัยนี้ไปประยุกต์ใช้กับภาพถ่ายหน้าตัดของปลายท่อนซุงชนิดไม้อื่น ๆ นอกเหนือจากไม้ยูคาลิปตัส สำหรับกระบวนการการสอนและการทดสอบอาจเพิ่มตัวอย่าง ภาพหน้าตัดของปลายท่อนซุงชนิดอื่น ๆ เพื่อให้ระบบมีความหลากหลายและเป็นประโยชน์ ในอุตสาหกรรมไม้มากยิ่งขึ้น
2. ในการตรวจจับหน้าตัดของปลายท่อนซุงยังเกิดความผิดพลาดในกรณีที่หน้าตัดปลายท่อนซุง ที่ติดกันบางท่อนไม่สามารถแยกออกจากกันได้ เนื่องจากสีของพื้นผิวหน้าตัดที่ต่อเนื่องกัน วิธีการเปลี่ยนแปลงลักษณะโครงร่างของภาพไม่สามารถแยกหน้าตัดของปลายท่อนซุง หากเพิ่มวิธีให้แยกหน้าตัดที่เชื่อมติดกันได้เหมาะสมมากขึ้นก็จะส่งผลให้ผลการทดลองมีความ ถูกต้องมากยิ่งขึ้น

รายการอ้างอิง

- [1] จ. วัชรินทร์รัตน์, ผลผลิตและเศรษฐกิจของไม้ยูคาลิปตัสตามลักษณะการใช้ประโยชน์. กรุงเทพฯ: มหาวิทยาลัยเกษตรศาสตร์, 1998.
- [2] A. Rahman, S. Yella, and M. Dougherty, "Image Processing Technique to Count the Number of Logs in a Timber Truck," 2012: ACTAPRESS.
- [3] M. T. Dmitry Shvarts, "BULK MATERIAL VOLUME ESTIMATION METHOD AND SYSTEM FOR LOGISTIC APPLICATIONS," presented at the 9th International DAAAM Baltic Conference, Tallinn, Estonia, 2014.
- [4] V. A. Knyaz and A. A. Maksimov, "Photogrammetric Technique for Timber Stack Volume Control," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 3, pp. 157-162, 2014/08/01/ 2014.
- [5] *TIMBER MEASUREMENT MANUAL* Republic of Ireland: the Forest Service in the Department of the Marine & Natural Resources, 1999.
- [6] R. C. Gonzales and P. Wintz, *Digital Image Processing (2Nd Ed.)*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1987.
- [7] "Color model." [Online]. Available: https://en.wikipedia.org/wiki/Color_model. [Accessed: 2018, March 05].
- [8] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62-66, 1979/01// 1979.
- [9] "Digital Watermarking." [Online]. Available: <https://www.cosy.sbg.ac.at/~pmeerw/Watermarking/>. [Accessed: 2018, March 05].
- [10] "Connected Component Labeling Algorithm for very complex and high resolution images on an FPGA platform." [Online]. Available: <https://www.semanticscholar.org/paper/Connected-Component-Labeling-Algorithm-for-very-and-Schwenk-Huber/47bca595236fee30dbf364a8d7156231b292addb>. [Accessed: 2018, March 05].

- [11] "Morphological Image Processing." [Online]. Available: <http://www.dspguide.com/ch25/4.htm>. [Accessed: 2018, March 05].
- [12] "Convolutional Neural Network." [Online]. Available: <https://www.mathworks.com/solutions/deep-learning/convolutional-neural-network.html>. [Accessed: 2018, March 15].
- [13] "Convolutional Networks for everyone." [Online]. Available: <https://medium.com/@rohantomas.me/convolutional-networks-for-everyone-1d0699de1a9d>. [Accessed: 2018, March 10].
- [14] W. Liu *et al.*, "SSD: Single Shot MultiBox Detector," *arXiv:1512.02325 [cs]*, vol. 9905, pp. 21-37, 2016 2016.
- [15] A. G. Howard *et al.*, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," *arXiv:1704.04861 [cs]*, 2017/04/16/ 2017.
- [16] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv:1409.1556 [cs]*, 2014/09/04/ 2014.
- [17] "MobileNet version 2." [Online]. Available: <http://machinethink.net/blog/mobilenet-v2/>. [Accessed: 2018, March 10].
- [18] J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," *arXiv:1411.4038 [cs]*, 2014/11/14/ 2014.
- [19] "A review of deep learning models for semantic segmentation." [Online]. Available: <https://nicolovaligi.com/deep-learning-models-semantic-segmentation.html>. [Accessed: 2018, March 15].
- [20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *arXiv:1512.03385 [cs]*, 2015/12/10/ 2015.
- [21] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," 2012, pp. 1097-1105, USA: Curran Associates Inc.
- [22] "ImageNet." [Online]. Available: <http://www.image-net.org/>. [Accessed: 2017, November 23].
- [23] "VGG in TensorFlow." [Online]. Available: <http://www.cs.toronto.edu/~frossard/post/vgg16/>. [Accessed: 2018, March 15].

- [24] L. Perez and J. Wang, "The Effectiveness of Data Augmentation in Image Classification using Deep Learning," *arXiv:1712.04621 [cs]*, 2017/12/13/ 2017.
- [25] T. K. Mats Nylinder, Mikael Hulnäs, "Roundwood measurement of truck loads by laser scanning," p. 9
- [26] A. Molder and O. Martens, "Image Processing in the Woodworking Industry: Challenges, Solutions and Platforms," *Electronics and Electrical Engineering*, 2011.
- [27] V. Knyaz, Y. Visilter, and S. Zheltov, "Photogrammetric Techniques for Measurements in Woodworking Industry," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 2004, 2004.
- [28] C. Herbon, K. Tönnies, and B. Stock, "Detection and Segmentation of Clustered Objects by Using Iterative Classification, Segmentation, and Gaussian Mixture Models and Application to Wood Log Detection," Cham, 2014, pp. 354-364: Springer International Publishing.
- [29] Y. Zhao and H. e. Ren, "Study on Genetic Algorithm and Homomorphic Filter for Logs Face Image Processing," *Journal of Information and Computational Science*, vol. 11, no. 8, pp. 2793–2800, 2014.
- [30] A. V. Mehrentsev and A. V. Kruglov, "The Algorithm and Software for Timber Batch Measurement by Using Image Analysis," *RFMI – Representation Analysis and Recognition of Shape and Motion from Image Data*, 2017.
- [31] "COCO - Common Objects in Context." [Online]. Available: <http://cocodataset.org>. [Accessed: 2017, August 24].
- [32] "Models and examples built with TensorFlow." [Online]. Available: <http://github.com/tensorflow/models>. [Accessed: 2017, August 24].
- [33] "LabelImg." [Online]. Available: <http://github.com/tzutalin/labelImg>. [Accessed: 2017, August 13].
- [34] "Contribute to CarND-Semantic-Segmentation development by creating an account on GitHub." [Online]. Available: <http://github.com/udacity/CarND-Semantic-Segmentation>. [Accessed: 2018, January 05].

- [35] "Neural networks." [Online]. Available: https://ml4a.github.io/ml4a/neural_networks/. [Accessed: 2018, March 05].
- [36] "Understanding SSD MultiBox — Real-Time Object Detection In Deep Learning." [Online]. Available: <https://towardsdatascience.com/understanding-ssd-multibox-real-time-object-detection-in-deep-learning-495ef744fab>. [Accessed: 2018, March 10].





ภาคผนวก

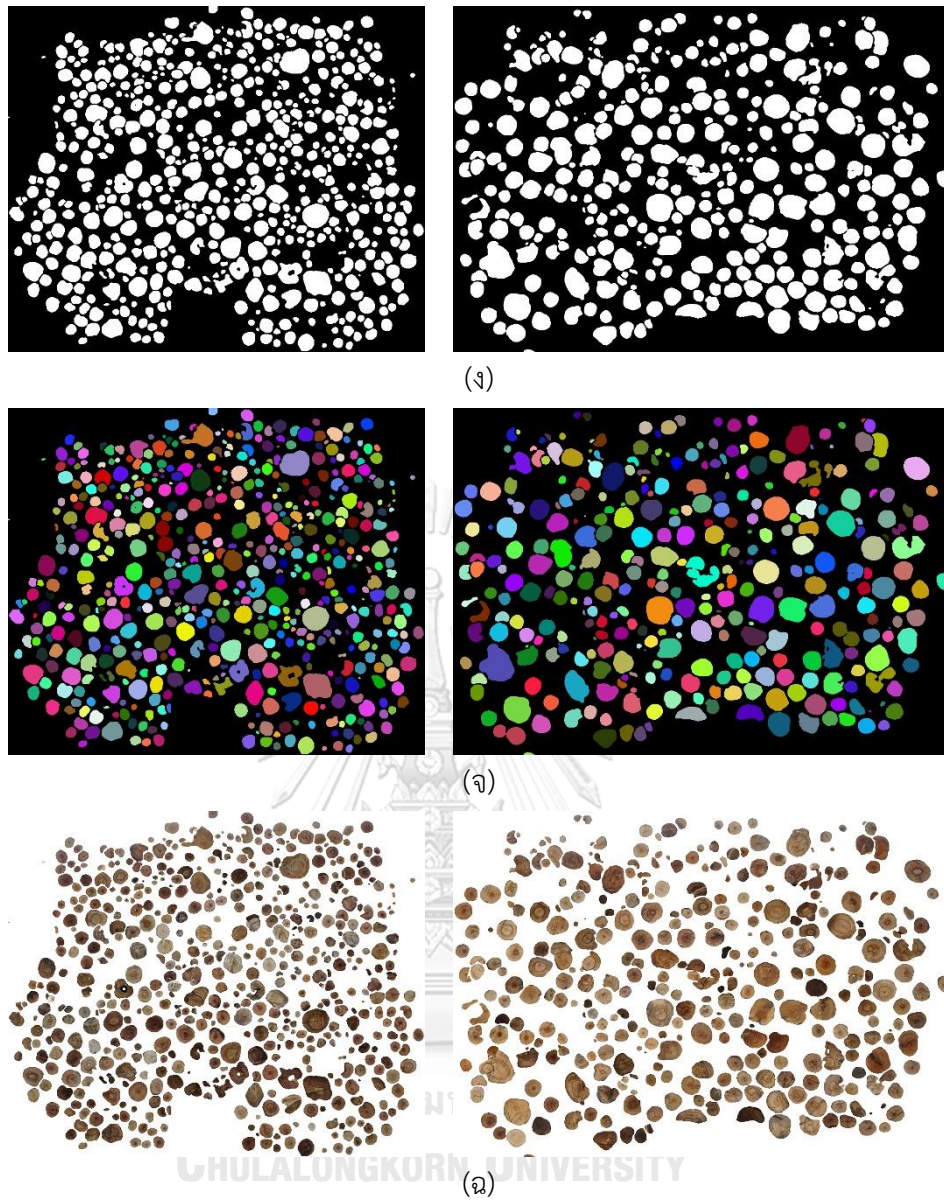
จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY



ตัวอย่างของภาพผลการทดลองเพิ่มเติมจากการทดลองในสภาวะแวดล้อมจริงในพื้นที่โรงงานผลิตกระดาษ ณ บริษัทสยามเซลลูโลส จำกัด โดยใช้ภาพจากด้านหลังรถบรรทุกไม้ยูคาลิปตัส ภาพตัวอย่างผลลัพธ์ของขั้นตอนดังที่ระบบเสนอแสดงในรูปที่ ก-1



รูปที่ ก-1 ตัวอย่างผลลัพธ์ของขั้นตอนที่ได้จากการทดลอง (ก) ภาพนำเข้า (ข) ผลลัพธ์ของขั้นตอนการตรวจจึบรถบรรทุกไม้ (ค) ผลลัพธ์ของการแยกส่วนหน้าตัดของปลายท่อนซุง



รูปที่ ก-1 ตัวอย่างผลลัพธ์ของขั้นตอนที่ได้จากการทดลอง (ต่อ)

(ง) ผลลัพธ์ของการเปลี่ยนแปลงลักษณะโครงร่างของภาพ (จ) ผลลัพธ์ของการนับปลายท่อนซุง

(ฉ) การซ้อนทับระหว่างภาพนำเข้าและภาพผลลัพธ์



(ก)

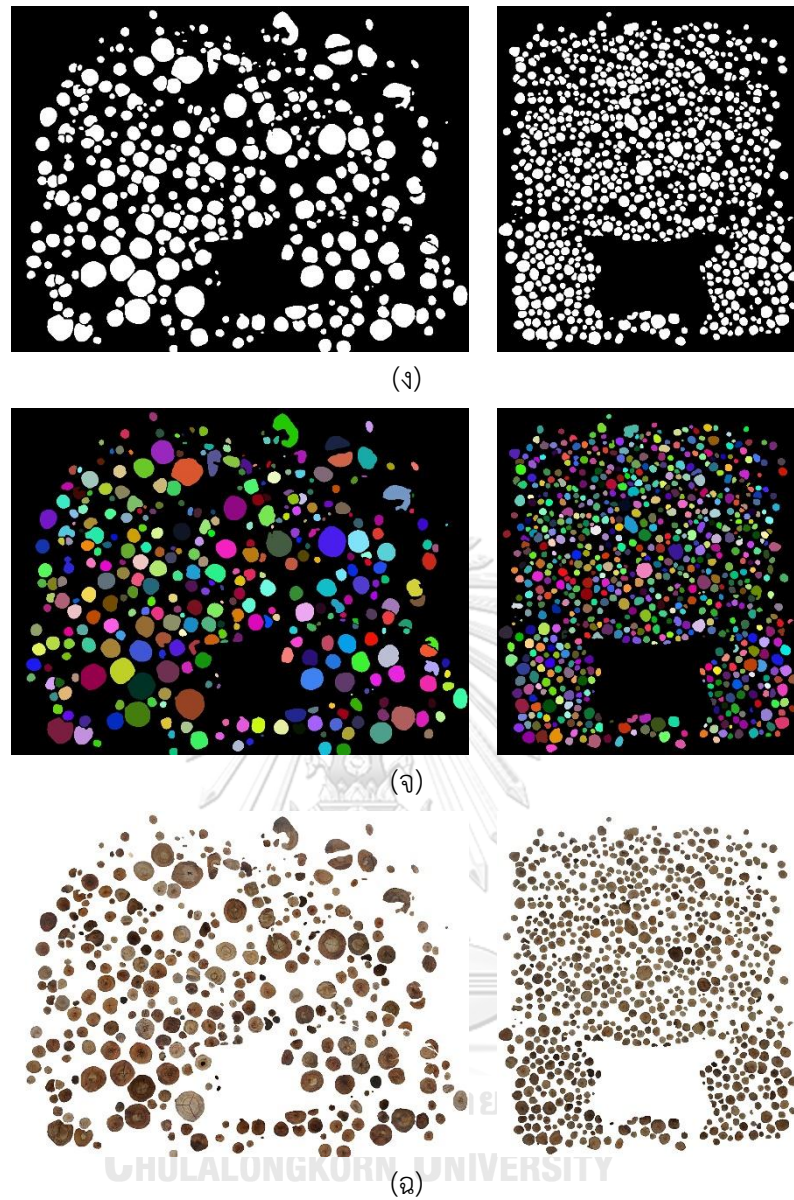


(ข)



(ค)

รูปที่ ก-2 ตัวอย่างผลลัพธ์ของขั้นตอนที่ได้จากการทดลอง (ก) ภาพนำเข้า (ข) ผลลัพธ์ของขั้นตอนการตรวจจับรัถบรทุกไม้ (ค) ผลลัพธ์ของการแยกส่วนหน้าตัดของปลายท่อนซุง



รูปที่ ก-2 ตัวอย่างผลลัพธ์ของขั้นตอนที่ได้จากการทดลอง (ต่อ)

(ง) ผลลัพธ์ของการเปลี่ยนแปลงลักษณะโครงร่างของภาพ (จ) ผลลัพธ์ของการนับปลายท่อนซุง

(ฉ) การซ้อนทับระหว่างภาพนำเข้าและภาพผลลัพธ์



(ก)

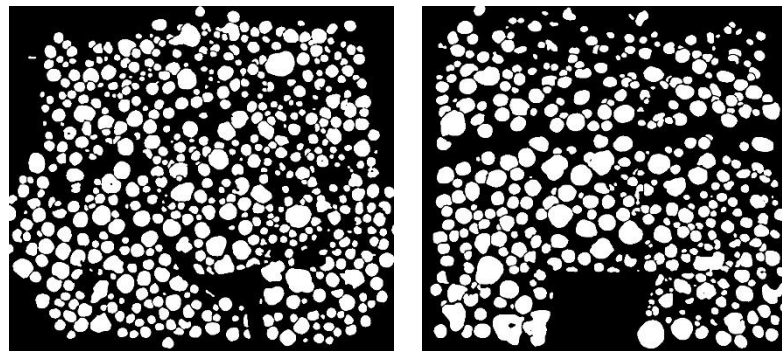


(ข)

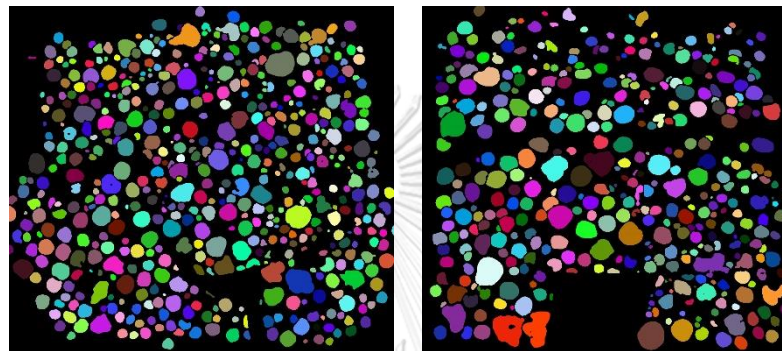


(ค)

รูปที่ ก-3 ตัวอย่างผลลัพธ์ของขั้นตอนที่ได้จากการทดลอง (ก) ภาพนำเข้า (ข) ผลลัพธ์ของขั้นตอนการตรวจจําับรณรรทุกไม้ (ค) ผลลัพธ์ของการแยกส่วนหน้าตัดของปลายท่อนซุง



(ง)



(จ)



(ฉ)

รูปที่ ก-3 ตัวอย่างผลลัพธ์ของขั้นตอนที่ได้จากการทดลอง (ต่อ)

(ง) ผลลัพธ์ของการเปลี่ยนแปลงลักษณะโครงร่างของภาพ (จ) ผลลัพธ์ของการนับปลายท่อนซุง

(ฉ) การซ้อนทับระหว่างภาพนำเข้าและภาพผลลัพธ์



(ก)

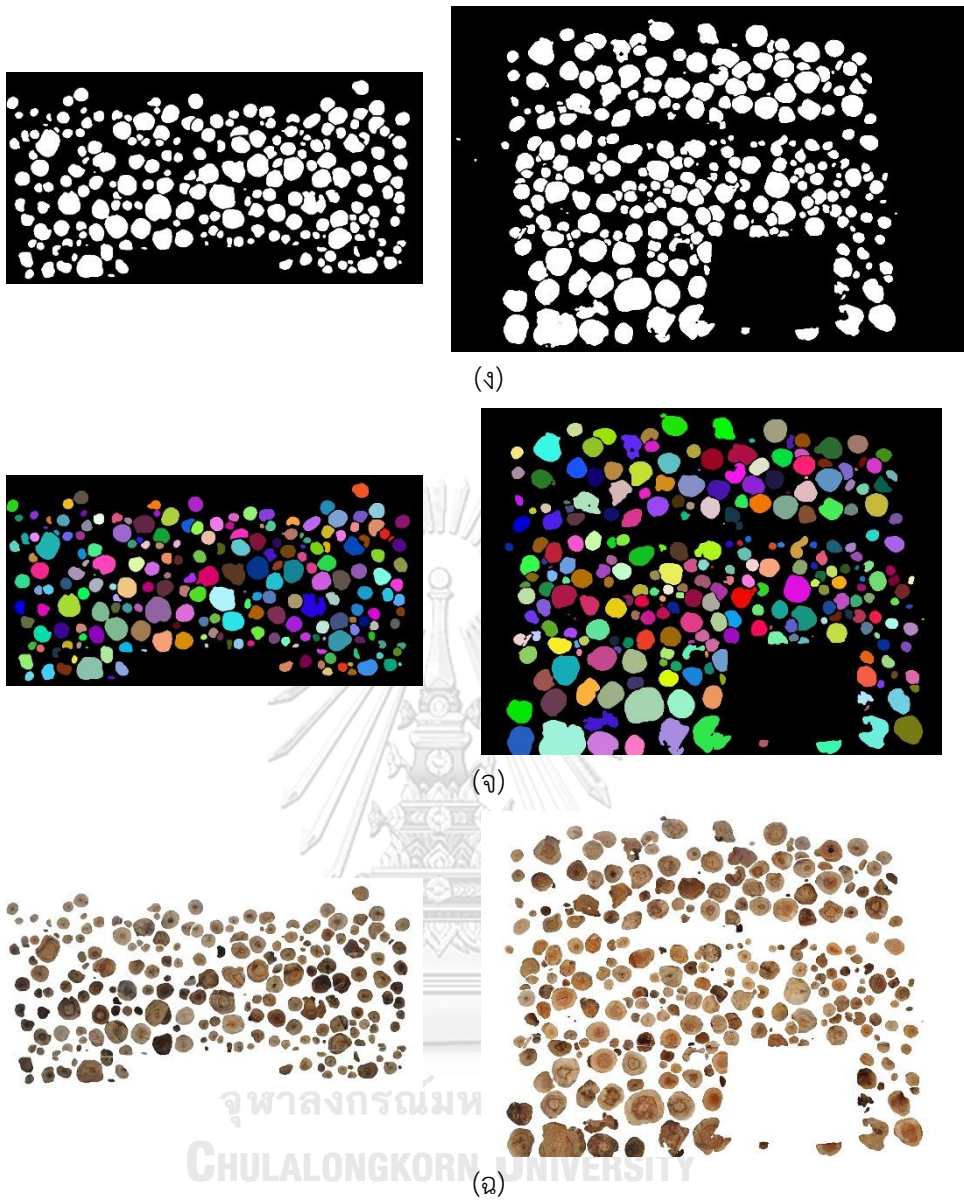


(ข)



(ค)

รูปที่ ก-4 ตัวอย่างผลลัพธ์ของขั้นตอนที่ได้จากการทดลอง (ก) ภาพนำเข้า (ข) ผลลัพธ์ของขั้นตอนการตรวจจําับรถบรรทุกไม้ (ค) ผลลัพธ์ของการแยกส่วนหน้าตัดของปลายท่อนซุง



รูปที่ ก-4 ตัวอย่างผลลัพธ์ของขั้นตอนที่ได้จากการทดลอง (ต่อ)

(ง) ผลลัพธ์ของการเปลี่ยนแปลงลักษณะโครงร่างของภาพ (จ) ผลลัพธ์ของการนับปลายท่อนซุง

(ฉ) การซ้อนทับระหว่างภาพนำเข้าและภาพผลลัพธ์



(ก)

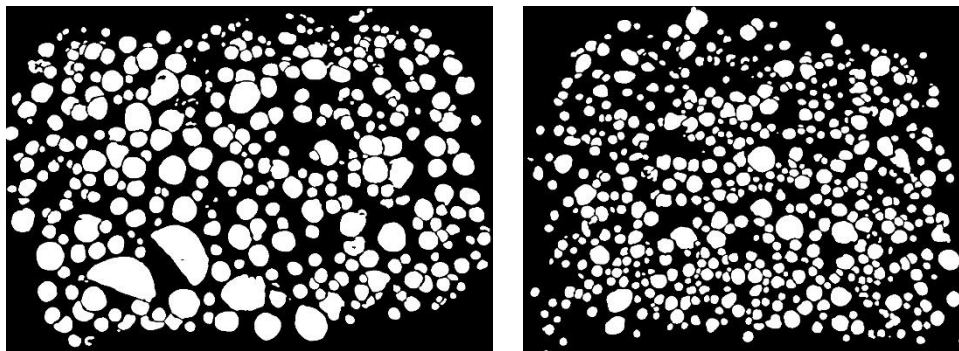


(ข)

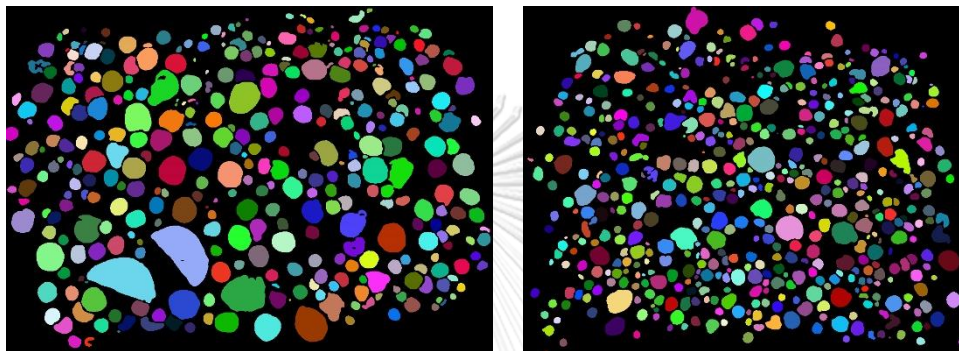


(ค)

รูปที่ ก-5 ตัวอย่างผลลัพธ์ของขั้นตอนที่ได้จากการทดลอง (ก) ภาพนำเข้า (ข) ผลลัพธ์ของขั้นตอนการตรวจจําับรรทุกไม้ (ค) ผลลัพธ์ของการแยกส่วนหน้าตัดของปลายท่อนซุง



(ง)



(จ)



(ฉ)

รูปที่ ก-5 ตัวอย่างผลลัพธ์ของขั้นตอนที่ได้จากการทดลอง (ต่อ)

(ง) ผลลัพธ์ของการเปลี่ยนแปลงลักษณะโครงร่างของภาพ (จ) ผลลัพธ์ของการนับปลายท่อนซุง

(ฉ) การซ้อนทับระหว่างภาพนำเข้าและภาพผลลัพธ์



บทความเรื่อง “Log-End Cut-Area Detection in Images Taken from Rear End of Eucalyptus Timber Trucks” เสนอในงานประชุมวิชาการ 2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE2018) ที่ Faculty of ICT, Mahidol University, Salaya Campus, Nakhon Pathom, Thailand จัดระหว่างวันที่ 11-13 กรกฎาคม พ.ศ.2561 บทความนี้ได้รับการตอบรับให้ตีพิมพ์ไว้ใน IEEEExplore® Digital Library



Log-End Cut-Area Detection in Images Taken from Rear End of Eucalyptus Timber Trucks

Noppawat Samdangdech, and Suebskul Phiphobmongkol
 Computer Graphic and Computer Imaging Laboratory
 Department of Computer Engineering, Faculty of Engineering,
 Chulalongkorn University, Bangkok, Thailand.
 E-mail: noppawat.sa@student.chula.ac.th, suebskul.p@chula.ac.th

Abstract—The visual estimation of log volume and size distribution of eucalyptus logs on a truck is a challenging task. In Thailand, inspectors at paper mills typically perform this task. The information is used to determine whether the logs pass the criteria for the mill and to find the appropriate price. This method is far from accurate and not efficient. This paper presents a new approach to automatically detects eucalyptus log-end cut area from rear-end images of eucalyptus timber trucks. The method used machine learning and image processing techniques. It consists of three parts: timber truck detection, log segmentation, and log counting. The proposed system was tested with 300 images of timber truck dataset and achieved an average accuracy of 94.45% in log segmentation and 2.71% of false negative.

Keywords—log detection; log counting; eucalyptus log; image processing; image segmentation; machine vision; machine learning;

I. INTRODUCTION

Eucalyptus provides suitable fiber to make pulp for paper manufacturing. Eucalyptus has many useful characteristics such as high pulp yield, high basic wood density, good fiber properties, and fast growing. The trees can be cut for use within 4-5 years. For these reasons, eucalyptus are commonly used in pulp and papermaking industries over other trees. Therefore, the industrial demand of eucalyptus trees is rising fast. In timber transportation, logs are loaded onto trucks and transported to paper mills. When a truck arrives at a paper mill, an inspector will visually estimate log diameters by using a template. The information will be used to select the logs that are appropriate for the mills and to determine the log price for each truck.

In Thailand, log diameter estimation procedure is normally performed by placing a reference diameter template on the log ends. Then, an inspector visually estimates the log diameters at the rear end of a timber truck using a template as shown in Figure 1 (b). Due to the inaccuracy of human estimation, sellers and buyers may settle at the prices that are too high or too low. Moreover, the log size determination process is time consuming and causes extra labor cost.

Nowadays, advances in digital image processing, machine learning, and other computing techniques provide methods that may support the log-end cut-area detection and diameter estimation. Log-end detection and measurement of eucalyptus logs on a truck is quite a challenging task for a computer.

Most log-end cut area of eucalyptus have some similar features such as common range of colors and textures. However, some log-end images may be affected by environmental and external factors such as light condition, shade, humidity, irregular shapes of log cuts, thus, make it more difficult for a computer for detection and measurement.

This paper proposed a log-end cut-area detection method using Single Shot Multi-Box Detector (SSD) [10] based object detection to detect the log ends on a truck in an image. Then semantic segmentation using Fully Convolutional Networks (FCNs) [11] was applied to extract each log end cut area from its background objects and undesired objects. Finally, image processing techniques were applied to separate connected logs and also to count the number of log ends.



Fig 1. An example image of a timber truck and reference diameter template on log end.

Section II describes a survey of the current vision-base in log-end detection. Section III describes the proposed method overview. Section IV describes how log-end cut areas were detected by using image processing techniques combined with machine learning approach. Section V explains the experimental setup and shows the results. Finally, we discuss and conclude the research work in Section VI.

II. RELATED WORKS

The works on log-end detection using algorithms on images are relatively rare, especially, from timber truck images. Nevertheless, there are two main approaches to detect logs, i.e. laser-base approach and vision-base approach. For the laser-base approach, Nylander et al. [1] surveyed Logmeter system that performed round wood measurement using 3D laser scanning from Woodtech Company in Chile. The method was performed when timber trucks arrived at a station that had

an embedded Log meter system. Then, laser measuring devices installed at three positions at the station would scan timber trucks thoroughly. The system generated 3D representation of timber truck. It also removed unwanted parts from the 3D representation e.g. truck components. Next, the system identified all surrounding logs by referring to the diameters of each periphery. Finally, the system generated the results by computing parameters from previous output. The result contained log information such as log 3D representation, log length, log diameter, and log volume.

For image processing approach, Rahman et al. [2] proposed a method to find log volume on timber trucks and collect log transportation data before arrival at the sawmill. The method used images taken from the rear end of timber trucks to process through the three main subsystems. The pre-processing is applied by separating Value component from HSV (Hue, Saturation, and Value) color space. The Value component correspond to the brightness of color [3]. Conversely, Hue and Saturation were defined as dominant color and purity of color [3]. The purpose of this method was to enhance and adjust the brightness and the contrast of an image. Then, Value properties were combined with Hue and Saturation properties, and converted an image into an RGB image. Additionally, the system detected edges using canny edge operation technique and dilated output edges to make edges thicker. Then, the main subsystem of this approach is the image segmentation. The RGB image was split into each channel and normalized using each pixel value in each channel divided by sum of pixel values from all channels, then converted the normalized RGB image into an HSV image. The system segmented logs by defining Hue range whose pixels correspond to logs color and kept only these pixels. Afterwards, they used edges from a second subsystem to superimpose on segmented log images. This method could eliminate pixels that were not log-end cut area and separated connected logs. Finally, they used connected component labeling to count logs in an image and get an average of about 97% success rate in log counting.

Knyaz et al. [4] proposed an image processing approach which improved Knyaz et al. [5] method. The aim was to estimate the volume of log stack using a calibration tool. The system consisted of two DSLR cameras which were located from 2 to 20 metres from the log stack. Then, two cameras took images and calibrated the same reference point of images through a photogrammetric system. Gradient vector was used to find edge in an image and set a threshold value to eliminate weak gradient values which was less than a threshold. If the gradient value was greater than the threshold, a binary image was then created from gradient values. Finally, the possible circles from contour points were created using iterative search of local maxima. These circles were then used to find log volume and stacking coefficient.

Mehrentsev and Kruglov [6] proposed an algorithm to

measure round timber. The system was consisted of three subsystems: object detection, log clustering, and log segmentation. In object detection, Sobel operation was used to create gradient vector components including magnitude and direction which calculated the center of circle. Then, the search of local maximum algorithm was applied to detect logs in an image. This method can solve overlapping-log problem and can detect connected logs using mean shift clustering [7]. In case of groups of logs that were separated from each other, a graph clustering algorithm was applied on the set of detected logs to find connected groups of logs. In log segmentation, watershed algorithm [8] and minimum s-t graph [9] were used to segment a region of each log. Finally, there could be some logs which were undetected or detected incorrectly, this system allowed manual editing to add or delete logs.

For all the mentioned log detection methods, the 3D laser scanning was appropriate for large logs and the operation must be performed in the place where embedded devices were installed. This method was used in the countries that had many wood industries, such as Scandinavian countries. Besides, the rest of the three systems surveyed previously dealt with clean log-end cut area images which had consistent color. Most methods worked well with round log ends.

Unfortunately, there are many issues of timber truck images which can cause difficulties to the above methods. These issues include randomly placed logs on trucks, different colors of logs, log length, overlapping logs, cracks on logs, shadowed on logs, and other undesired objects covering on some logs. In addition, logs vary in sizes and shapes which are irregular and may not be round. With these issues, the detection of timber trucks in Thailand are typically quite complex, especially on eucalyptus logs.



Fig. 2. Example images of variations of eucalyptus timber trucks in Thailand

III. PROPOSED METHOD

In this section, a new method was proposed to detect eucalyptus log-end cross section. The system overview is shown in Fig. 3. Initially, the input image was taken by a camera during daytime. The object detection was applied in the area of interest of the timber truck. An image was cropped using coordinates of detected boundary. Then, log segmentation was applied to segment each log from a background which operates by Fully Convolutional Networks

for semantic segmentation. Furthermore, an image processing technique was applied using binary conversion, morphological operation, and connected component labeling to separate and count each log. The output of log counting was the already segmented image whose log-end cut-area segments were counted.

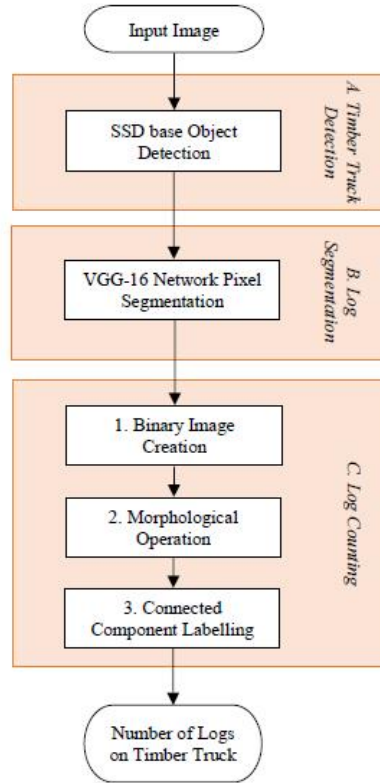


Fig. 3. The flowchart diagram shows the overview system of log end cut-area detection.

The proposed system can be described in three major steps as follows: timber truck detection, log segmentation, and log counting.

A. Timber Truck Detection

Single Shot Multi-Box Detector (SSD) network architecture proposed by Liu et al. [10] was applied to detect a bounding boxes of timber trucks in an image. We labeled a new class on timber truck images of training dataset and train our training dataset on existing pre-trained SSD Mobilenet v1 COCO. Then, the SSD network was used to find possible areas of bounding box which contained timber trucks at that time. When training process finished, the next step was to test

outcome on testing dataset and resulted in bounding boxes as shown in Fig. 4. (a2) and (b2).

However, there might be some input images with multiple timber trucks in an image. The system eliminated other trucks not of interest by calculating the area of the bounded boxes. The one with the biggest area was chosen as show in Fig. 4 (b2) and (b3). With this elimination, the system can reduce computation cost for semantic segmentation and eliminated undesired objects in an image, except the timber truck of interest.

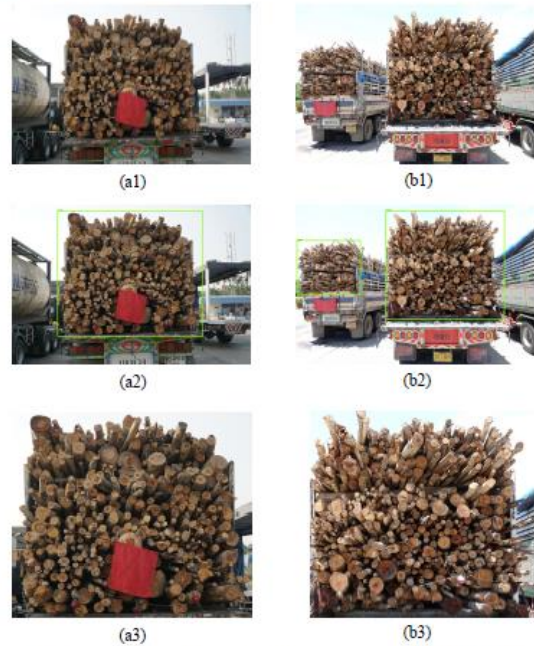


Fig. 4. (a1) normal input image and (b1) input image with multiple timber trucks, (a2) and (b2) detected region of timber trucks, (a3) and (b3) cropped timber trucks.

B. Log Segmentation

The system performed the Fully Convolutional Networks (FCNs) for semantic segmentation which was widely used to implement the end-to-end convolutional neural network [11]. From the timber truck detection step, a detected area of timber truck was obtained. Semantic segmentation approach was then used to extract only the pixel area of log-end cut area. Pixels were labeled in images of training dataset with a class of log end and non-log end cut area. To improve the accuracy rate, the data augmentation technique of OpenCV library was applied to generate more quantities and varieties of training dataset.

After the dataset preparation process, the transfer learning approach was used to fine-tune network parameters of existing pre-trained VGG-16 network which was proposed by

Simonyan and Zisserman [12]. Then, the FCNs was trained base on VGG-16 with our training dataset to classify pixel area of log-end cut area in an image. The results of testing dataset which were extracted pixels on log-end cut area were shown in Fig. 5.



Fig. 5. (a) and (b) are output images of log segmentation.

C. Log Counting

In this step, a log counting method was proposed using image processing techniques mostly used in OpenCV library. Segmented images from log segmentation step might have some log-end cut areas touching other logs and non-part of log end cut-area appeared. Therefore, to solve the problems, log counting was used as shown in Fig. 3.

1) Binary Image Creation

A binary image was created from segmented log end cut-area image by performing grayscale conversion of OpenCV library on segmented pixels. Then, global thresholding by Otsu's algorithm [13] was applied on a grayscale image to convert segmented image into binary image.

2) Morphological Operation

The results of segmented image have some log-end cut area that touched other logs. These touched or connected log-end cut areas would be separated by using opening operation of morphological [3] performed on binary image. The operation created a gap between touching log ends and reduced some noise in an image as shown in Fig. 6. (a1) and (b1).

3) Connected Component Labeling

The log-end cut areas were counted by using connected component labelling function of OpenCV. This procedure found different log-end cut areas and labeled the pixels that were of the same log-end cut area. The result of the labeling (counting) was shown in Fig 6. (a2), (b2), (a3), and (b3).

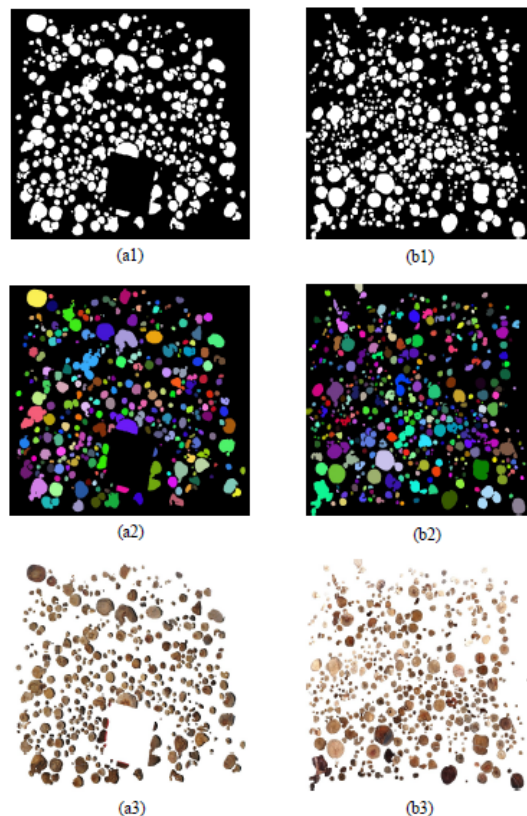


Fig. 6. (a1) and (b1) are results of opening operation applied on binary images, (a2) and (b2) are results of log counting, (a3) and (b3) are superimposed segmented on input images.

IV. EXPERIMENTS AND RESULTS

In this section, the experiments and results are described. The proposed method was implemented by using the following computing resources to process on the dataset

- PyCharm Community Edition IDE
- Tensorflow deep learning framework
- Ubuntu 16.04 LTS Operating System
- Intel® Core™ i7-6700 CPU @ 3.40 GHz
- NVIDIA GeForce GTX 1080Ti
- 32GB of Memory

A. Timber Trucks Dataset

Timber truck dataset were collected at the entrance of a paper mill (See Acknowledgement). The Sony A6000 digital camera with 18-105 mm lens was used to take images of timber trucks with variations of angles during the daytime. The dataset was divided into a training set and a testing set for

timber truck detection and log segmentation. Dataset included various cases of log-end cut area as mention in Section II.

For timber truck detection, 815 images were used for training with their annotations on existing pre-trained SSD Mobilenet v1 COCO model and 300 images of testing dataset were tested. For log segmentation, data augmentation on 50 images were applied with training dataset and achieved 550 images for training proposed. Then, we trained on pre-trained VGG16 model for 100 epochs and took an average of 7.44 minutes per epoch. Training dataset was tested on 300 images of timber trucks for log segmentation.

B. Results and Discussion

The proposed method was to detect log-end cut area in images taken from rear end of timber trucks. In the first part, the process found bounded area (bounding box) of timbers on the target truck. 300 images were tested. The calculated precision, recall, and F-Measure are shown in Table I.

TABLE I. TIMBER TRUCK AREA OF INTEREST DETECTION

Timber Truck Area of Interest (bounding box) Detection		
Precision	Recall	F-Measure
1	0.98	0.99

In the second part of the process, log segmentation and log counting were performed. The result of log counting was examined by comparing between the proposed system and the manual count. 300 images were tested and achieved the average accuracy of 94.45% as shown in Table II. The calculated precision, recall, and F-Measure are shown in Table III.

TABLE II. LOG END COUNTING RESULT

Log End Counting (Segmentation)	
Number of images segmented	Accuracy
300	94.45%

TABLE III. LOG END AREA OF INTEREST SEGMENTATION

Log End Area of Interest (labeling pixel area) Segmentation		
Precision	Recall	F-Measure
0.97	0.97	0.97

The segmented result from sample images in Table III was very satisfied when compared with result of manual counting. However, there were some segmented log ends which were affected. For example, some connected log ends cannot be separated because of continuous color and texture. Overlapping log and crack on log end were another problem of log end counting which both of log problems were external factors that cannot be handled. In addition, there were some segmented logs that were not an actual log end cut area which caused false positive. Also, some false negative occurred when a log end cut area was cover by some object for the cut area was far from perpendicular to the camera lens axis. Fig. 7. shows some of the problems mentioned above.

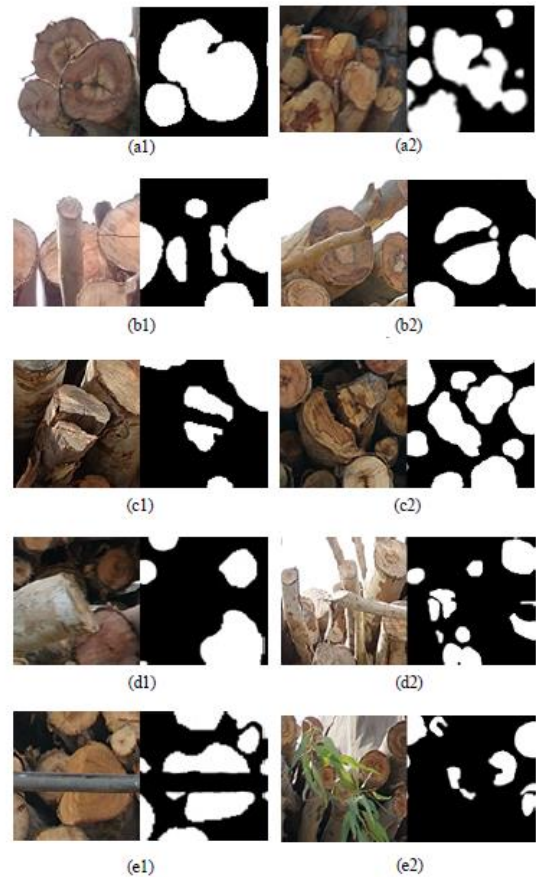


Fig. 9. Samples of original and segmented images of log problems (a1) and (a2) connected logs, (b1) and (b2) overlapping logs, (c1) and (c2) crack on log, (d1) and (d2) direction of log, (e1) and (e2) undesired object covered on logs

ภาคผนวก ค
ผลงานที่เป็นส่วนหนึ่งของวิทยานิพนธ์

- บทความเรื่อง “Log-End Cut-Area Detection in Images Taken from Rear End of Eucalyptus Timber Trucks” ได้รับรางวัล Best Paper ในหัวข้อ Application in Computer Science จากงานประชุมวิชาการ 2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE2018) ที่ Faculty of ICT, Mahidol University, Salaya Campus, Nakhon Pathom, Thailand จัดระหว่างวันที่ 11-13 กรกฎาคม พ.ศ.2561



ประวัติผู้เขียนวิทยานิพนธ์

นายณพัทธ์ สำแดงเดช เกิดวันที่ 24 กรกฎาคม 2534 ที่จังหวัดนครปฐม สำเร็จ การศึกษาระดับปริญญาตรีวิทยาศาสตร์บัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ คณะเทคโนโลยี สารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี และเข้าศึกษาต่อในหลักสูตรวิทยาศาสตร มหาบัณฑิต สาขาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2557

