

การพัฒนาการตรวจจับอันตรายระหว่างล้อและรางของรถไฟฟ้าด้วย
Faster R-CNN

DEVELOPMENT OF VISUAL INSPECTION OF THE INTERACTION BETWEEN
WHEEL AND RAILWAY OF ELECTRIC TRAINS USING FASTER R-CNN



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรดุษฎีบัณฑิต
สาขาวิชาวิศวกรรมไฟฟ้า
คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ. 2564

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้ KMITL-2021-EN-D-018-023 ญาติให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

DEVELOPMENT OF VISUAL INSPECTION OF THE INTERACTION BETWEEN
WHEEL AND RAILWAY OF ELECTRIC TRAINS USING FASTER R-CNN



A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF
DOCTORAL OF ENGINEERING IN ELECTRICAL ENGINEERING
FACULTY OF ENGINEERING
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2021

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
KMITL-2021-EN-D-018-023
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



COPYRIGHT 2021

FACULTY OF ENGINEERING

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามเผยแพร่ต่อสาธารณะ และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

| | |
|-----------------------------|---|
| หัวข้อวิทยานิพนธ์ | การพัฒนาการตรวจจับอันตรายระหว่างล้อและรางของรถไฟฟ้ ด้วย Faster R-CNN |
| นักศึกษา | นางสาวสุภารัตน์ พุ่มย่อย |
| รหัสนักศึกษา | 57601046 |
| ปริญญา | วิศวกรรมศาสตรดุษฎีบัณฑิต |
| สาขาวิชา | วิศวกรรมไฟฟ้า |
| พ.ศ. | 2564 |
| อาจารย์ที่ปรึกษาวิทยานิพนธ์ | รศ.ดร. สมยศ เกียรติวนิชวิไล |

บทคัดย่อ

วิทยานิพนธ์นี้นำเสนอเทคนิคปัญญาประดิษฐ์ (Artificial Intelligence) มาประยุกต์ใช้ร่วมกับระบบวิทัศน์อัตโนมัติเพื่อวิเคราะห์ความผิดปกติของอันตรายระหว่างล้อกับรางรถไฟฟ้ด้วย Faster R-CNN ซึ่งเทคนิคดังกล่าวสามารถระบุตำแหน่งของบริเวณขอบล้อและรางได้อย่างชัดเจน แสดงให้เห็นถึงความแม่นยำในการตรวจจับ สามารถเก็บข้อมูลและนำไปวิเคราะห์ผลรูปภาพได้ทันที ทำให้การตรวจจับเป็นแบบตามเวลาจริง (real-time) อีกทั้งยังนำเสนอการลดชั้นของโครงสร้าง VGG-16 ลง ผลจากการทดสอบแสดงให้เห็นว่าใช้เวลาในการประมวลผลรวดเร็วขึ้น

| | |
|-----------------------|--|
| Thesis Title | Development of visual inspection of the interaction between wheel and railway of electric using Faster R-CNN |
| Student | Ms. Suparat Pomyoy |
| Student ID. | 57601046 |
| Degree | Doctoral of Engineering |
| Program | Electrical Engineering |
| Year | 2021 |
| Thesis Advisor | Assot. Prof. Dr. Somyot Kaitwanidvilai |

ABSTRACT

This dissertation proposes an Artificial Intelligence (AI) technique applying with an automatic vision system for analyzing abnormality of the interaction between wheel and electrical-train railway using the Faster R-CNN. This technique can accurately and clearly detect positions of wheel rim and railway, showing the precise detection. This can collect the data and perform real-time picture analysis. It also presents a reduced of the VGG-16. The testing results show a faster processing detection time.

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้จะสำเร็จมิได้หากไม่ได้รับคำแนะนำ คำปรึกษา การชี้แนะและการเอาใจใส่ เป็นอย่างดีมาโดยตลอดจาก รศ.ดร.สมยศ เกียรติวนิชวิไล ซึ่งเป็นอาจารย์ที่ปรึกษา ข้าพเจ้ารู้สึกซาบซึ้งในความกรุณาและความอนุเคราะห์จากอาจารย์ และขอขอบพระคุณเป็นอย่างสูง

ขอขอบพระคุณ ผศ.ดร.อนรรฆพล แสนทน หัวหน้าโครงการการพัฒนาระบบปัญญาประดิษฐ์ วิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมไฟฟ้า ที่ให้บริการเพื่อการประเมินความปลอดภัยและสภาพทางวิ่ง และ ดร.ภูมิ คงหัวรอบ ที่ช่วยเหลือและให้คำปรึกษาทุกอย่าง ขอขอบคุณ นายกษิตศ ทิพย์วัฒน์ ที่ช่วยเหลือด้านโปรแกรมเกี่ยวกับงานวิจัยนี้

ขอขอบพระคุณคณาจารย์ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ทุกท่านที่ได้อบรม สั่งสอนและประสิทธิ์ประสาทวิชาให้กับข้าพเจ้ามาโดยตลอด

ขอขอบคุณพี่ๆ น้องๆ ส่วนสนับสนุนวิชาการ ภาควิชาวิศวกรรมไฟฟ้า และรุ่นพี่ รุ่นน้อง ภาควิชาวิศวกรรมไฟฟ้า สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังทุกคนที่ให้การสนับสนุนช่วยเหลือเรื่องงานและเวลาในการปฏิบัติงาน ตลอดจนอำนวยความสะดวกเรื่องต่างๆ เป็นอย่างมาก

ขอขอบคุณ คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ที่ให้ความช่วยเหลือในเรื่องค่าธรรมเนียมการศึกษา เป็นอย่างยิ่ง

ขอขอบคุณเจ้าหน้าที่งานบัณฑิตศึกษา คณะวิศวกรรมศาสตร์ โดยเฉพาะอย่างยิ่ง คุณสิริทิณี กิ่งช้าง ที่ให้ความอนุเคราะห์ ช่วยเหลือในเรื่องต่างๆ

ขอขอบพระคุณ รศ.ดร.อุมา สีนุญเรื่อง และขอขอบคุณ นางสาวปานมุก นพวรรณ ที่ช่วยในการเขียนและตรวจบทความภาษาอังกฤษ

ขอขอบคุณพี่ๆ น้องๆ ในห้องปฏิบัติการ AIM Laboratory คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังทุกคนที่ให้ความช่วยเหลือ ให้คำแนะนำต่างๆ และให้กำลังใจเสมอมา

ขอขอบคุณคุณชัชสุนทร อนันตพงษ์ คุณสรวิฑู สวัสดิ์ และทีมงาน IAMBEST2020 ที่อำนวยความสะดวกและช่วยเหลือเรื่องบทความ

สุดท้ายนี้ข้าพเจ้าขอกราบขอบพระคุณมารดาเป็นอย่างยิ่ง รวมถึงบิดาและครอบครัวของข้าพเจ้าที่เป็นกำลังใจ และให้การสนับสนุนในทุกๆเรื่อง ทำให้ข้าพเจ้าสามารถทำวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงด้วยดี

คุณค่า และประโยชน์อันพึงมาจากวิทยานิพนธ์ฉบับนี้ ข้าพเจ้าขอบแต่ผู้มีพระคุณทุกท่าน

สารบัญ

| | หน้า |
|---|------|
| บทคัดย่อภาษาไทย..... | I |
| บทคัดย่อภาษาอังกฤษ..... | II |
| กิตติกรรมประกาศ..... | III |
| สารบัญ..... | IV |
| สารบัญตาราง..... | VIII |
| สารบัญรูป..... | IX |
| บทที่ 1 บทนำ..... | 1 |
| 1.1 ความเป็นมาและความสำคัญของปัญหา..... | 1 |
| 1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา..... | 4 |
| 1.3 ขอบเขตการวิจัย..... | 4 |
| 1.4 แผนการดำเนินงาน..... | 4 |
| 1.5 ขั้นตอนการศึกษา..... | 4 |
| บทที่ 2 ทฤษฎีพื้นฐานที่ใช้ในการวิจัย..... | 6 |
| 2.1 การประมวลผลภาพ..... | 7 |
| 2.1.1 การปรับปรุงคุณภาพของภาพระดับเทา..... | 7 |
| 2.1.1.1 การทำงานของตัวกรองแบบคอนโวลูชัน (Convolution filter)..... | 7 |
| 2.1.1.2 ตัวกรองแบบเกาส์ (Gaussian's filter)..... | 8 |
| 2.1.1.3 ตัวกรองแบบโซเบล (Sobel's filter)..... | 9 |
| 2.1.2 การแบ่งส่วนภาพ (Image Segmentation)..... | 10 |
| 2.1.3 กระบวนการทางคณิตศาสตร์สำหรับปรับปรุงคุณภาพ (Morphological)..... | 10 |
| 2.1.4 การวิเคราะห์ข้อมูลภาพ..... | 12 |
| 2.1.4.1 การอ่านค่าจากลักษณะของอนุภาค..... | 12 |
| 2.1.4.2 บริเวณจุดสนใจของภาพ (Region of interest: ROI)..... | 12 |
| 2.2 การจำแนกรูปภาพ..... | 13 |
| 2.2.1 ปัญญาประดิษฐ์ (Artificial Intelligence : AI)..... | 14 |
| 2.2.2 โครงข่ายประสาทเทียม (Neural Network: NN)..... | 14 |
| 2.2.3 หลักการทำงานของโครงข่ายประสาทเทียม..... | 16 |
| 2.2.3.1 การเรียนรู้แบบมีครูสอน (Supervised learning)..... | 17 |
| 2.2.3.1.1 การทำแมปปิงโมเดล (Generalization of mapping model)..... | 17 |

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

| | หน้า |
|--|------|
| 2.2.3.1.2 โครงข่ายประสาทเทียมแบบย้อนกลับ (Backpropagation neural network)..... | 18 |
| 2.2.3.2 การเรียนรู้แบบไม่มีครูสอน (Unsupervised learning)..... | 21 |
| 2.2.4 หลักการทำงานของโครงข่ายประสาทเทียม..... | 21 |
| 2.2.5 โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional neural network)..... | 26 |
| 2.2.6 โครงข่ายประสาทเทียมแบบแพร่ไปข้างหน้า (Feed forward neural network)..... | 26 |
| 2.2.7 การปรับค่าน้ำหนักของโครงข่ายประสาทเทียม (Weight modification)..... | 27 |
| 2.2.8 การทำงานของโครงข่ายประสาทแบบคอนโวลูชัน (Convolutional neural network)..... | 28 |
| 2.2.8.1 ชั้นคอนโวลูชัน (Convolutional layer)..... | 28 |
| 2.2.8.2 ชั้นแมกส์พูลลิ่ง (Max pooling layer)..... | 31 |
| 2.2.8.3 การจำกัดความเป็นเชิงเส้น (ReLU)..... | 32 |
| 2.2.8.4 การเชื่อมต่อกันของแต่ละชั้นอย่างสมบูรณ์ (Fully connected layer)..... | 34 |
| 2.3 โครงข่ายประสาทเทียมแบบคอนโวลูชันสำหรับการตรวจจับวัตถุ (Convolutional Neural Network model for object detection)..... | 35 |
| 2.3.1 R-CNN (Regions with Convolutional Neural Network feature)..... | 35 |
| 2.3.2 Fast R-CNN (Fast Recurrent Convolutional Neural Network)..... | 36 |
| 2.3.3 Faster R-CNN (Faster Recurrent Convolutional Neural Network)..... | 37 |
| บทที่ 3 การตรวจจับอันตรกิริยาระหว่างล้อและรางรถไฟฟ้าด้วยเทคนิค Faster R-CNN..... | 44 |
| 3.1 การตรวจจับอันตรกิริยาระหว่างล้อและรางรถไฟฟ้าด้วยเทคนิค Faster R-CNN..... | 44 |
| 3.1.1 กำหนดขอบเขต ส่วนที่พิจารณาเพื่อจัดเตรียมภาพ..... | 45 |
| 3.1.2 โครงสร้างแบบจำลองปัญญาประดิษฐ์..... | 49 |
| 3.1.3 การเรียนรู้จดจำเพื่อวิเคราะห์หาตำแหน่งของล้อและราง..... | 50 |
| 3.2 การหาระยะห่างของล้อและรางด้วย Faster R-CNN โดยกำหนดจุดขอบของล้อรถไฟฟ้าให้เป็นจุดอ้างอิง..... | 51 |
| 3.2.1 เตรียมภาพในการฝึกฝนให้แบบจำลองเรียนรู้..... | 52 |

สารบัญ (ต่อ)

| | หน้า |
|--|------|
| 3.2.1.1 การหาระยะห่างของล้อและรางด้วย Faster R-CNN ด้วยโครงสร้างของ VGG-16 โดยกำหนดจุดขอบของล้อรถไฟให้เป็นจุดอ้างอิง..... | 55 |
| 3.2.1.2 การหาระยะห่างของล้อและรางด้วย Faster R-CNN ด้วยโครงสร้างของ VGG- 16 แบบลดชั้นโดยกำหนดจุดขอบของล้อรถไฟให้เป็นจุดอ้างอิง..... | 56 |
| บทที่ 4 ผลการทดลอง..... | 59 |
| 4.1 ผลการตรวจจับอันตรกิริยาระหว่างล้อและรางรถไฟด้วยเทคนิค Faster R-CNN..... | 59 |
| 4.1.1 ผลจากการเรียนรู้และการทดสอบของปัญญาประดิษฐ์ด้วยเทคนิค Faster R-CNN..... | 59 |
| 4.1.2 ผลจากการตรวจจับตำแหน่งและระยะห่างของล้อและรางด้วย Faster R-CNN..... | 62 |
| 4.2 ผลการตรวจจับอันตรกิริยาระหว่างล้อและรางรถไฟด้วยเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 โดยกำหนดให้ขอบล้อเป็นจุดอ้างอิง..... | 65 |
| 4.2.1 ผลจากการเรียนรู้และการทดสอบของปัญญาประดิษฐ์ด้วยเทคนิค Faster R-CNN ด้วย VGG-16 โดยกำหนดให้ขอบล้อเป็นจุดอ้างอิง..... | 65 |
| 4.2.2 ผลจากการตรวจจับตำแหน่งและระยะห่างของล้อและรางด้วยเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 โดยกำหนดให้ขอบล้อเป็นจุดอ้างอิง..... | 67 |
| 4.3 ผลการตรวจจับอันตรกิริยาระหว่างล้อและรางรถไฟด้วยเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 แบบลดชั้นโดยกำหนดให้ขอบล้อเป็นจุดอ้างอิง..... | 68 |
| 4.3.1 ผลจากการเรียนรู้และการทดสอบของปัญญาประดิษฐ์ด้วยเทคนิค Faster R-CNN ด้วย VGG-16 แบบลดชั้นโดยกำหนดให้ขอบล้อเป็นจุดอ้างอิง..... | 68 |
| 4.3.2 ผลจากการตรวจจับตำแหน่งและระยะห่างของล้อและรางด้วยเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 แบบลดชั้นโดยกำหนดให้ขอบล้อเป็นจุดอ้างอิง..... | 69 |
| บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ..... | 73 |
| 5.1 สรุปผลการวิจัย..... | 73 |
| 5.1.1 สรุปผลการตรวจจับอันตรกิริยาระหว่างล้อและรางรถไฟด้วยเทคนิค Faster R-CNN..... | 73 |

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

| | หน้า |
|--|------|
| 5.1.2 สรุปผลการตรวจจับอันตรายระหว่างล้อและรางรถไฟด้วยเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16..... | 74 |
| 5.1.3 สรุปผลการตรวจจับอันตรายระหว่างล้อและรางรถไฟด้วยเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 แบบลดชั้น..... | 74 |
| 5.2 ข้อเสนอแนะ..... | 74 |
| บรรณานุกรม..... | 76 |
| ภาคผนวก ผลงานวิจัยที่ได้รับการตีพิมพ์เผยแพร่..... | 80 |
| ประวัติผู้เขียน..... | 115 |



สารบัญตาราง

| ตารางที่ | หน้า |
|--|------|
| 2.1 การวิเคราะห์ข้อมูลจากอนุภาค..... | 12 |
| 2.2 คำอธิบายเซลล์ประสาทและเซลล์ประสาทเทียม..... | 15 |
| 3.1 ภาพและค่า Label (x, y) ของตำแหน่งขอบร่าง..... | 52 |
| 3.2 ตารางสรุปผลการวิเคราะห์..... | 52 |
| 4.1 ค่าต่างๆที่วัดได้จากเทคนิค Faster R-CNN ด้วย VGG-16 กับ VGG-16 แบบลดชั้น โดยกำหนดให้ขอบลือเป็นจุดอ้างอิง..... | 73 |



สารบัญรูป

| ภาพที่ | หน้า |
|--------|---|
| 1.1 | แผนภาพแสดงร้อยละของสาเหตุของการเกิดอุบัติเหตุทางรถไฟในประเทศสหรัฐอเมริกา.....2 |
| 1.2 | แผนภาพการลดลงอย่างมีนัยสำคัญของอุบัติเหตุจากการขนส่งทางรางของประเทศ สหรัฐอเมริกา.....2 |
| 2.1 | การทำงานของตัวกรองด้วยวิธีคอนโวลูชัน.....8 |
| 2.2 | สมาชิกในตัวกรองแบบเกาส์.....9 |
| 2.3 | สมาชิกของตัวกรอง (ก) สมาชิกสำหรับตัวกรองแนวนอน (ข) สมาชิกสำหรับตัวกรองแนวตั้ง.....9 |
| 2.4 | หน้าต่างโครงร่าง.....10 |
| 2.5 | (ก) ภาพต้นฉบับ (ข) ภาพหลังจากการทำกรรอน (ค) ภาพหลังจากการทำกรขยาย.....11 |
| 2.6 | (ก) ภาพหลังจากกระบวนการเปิด (ข) ภาพหลังจากกระบวนการปิด.....11 |
| 2.7 | ภาพที่ถูกกำหนดจุดสนใจ.....13 |
| 2.8 | ส่วนประกอบของโครงข่ายประสาทเทียม.....15 |
| 2.9 | การเรียนรู้ของแบบจำลองโครงข่ายประสาทเทียม.....16 |
| 2.10 | กราฟของฟังก์ชันซิกมอยด์.....18 |
| 2.11 | ชั้นโครงสร้างโครงข่ายประสาทเทียม.....22 |
| 2.12 | ไบแอส (Bias) และน้ำหนัก (Weight) ที่อยู่ในชั้นโครงสร้างโครงข่ายประสาทเทียม.....23 |
| 2.13 | โครงข่ายประสาทเทียมแบบเชื่อมต่อกันอย่างสมบูรณ์.....24 |
| 2.14 | ส่วนประกอบของเซลล์ประสาทและทางเดินประสาท.....25 |
| 2.15 | โครงข่ายประสาทเทียมแบบคอนโวลูชัน.....26 |
| 2.16 | ทิศทางการคำนวณของโครงข่ายประสาทเทียม.....27 |
| 2.17 | ตัวอย่างการแทนค่าพิกเซลลงบนรูปที่รับเข้ามา.....29 |
| 2.18 | จำลองเมทริกซ์ (ก) เมทริกซ์ที่ได้จากรูปที่รับเข้า (ข) เมทริกซ์ตัวกรองค่า.....29 |
| 2.19 | การทำงานของตัวกรองค่าและการกำเนิดเมทริกซ์ชุดใหม่หรือฟังก์ชันลักษณะ.....30 |
| 2.20 | ภาพขาวดำดั้งเดิมเมื่อผ่านการทำคอนโวลูชันกลายเป็นฟังก์ชันลักษณะ (Feature Map).....30 |
| 2.21 | การทำงานของชั้นคอนโวลูชัน.....31 |
| 2.22 | การทำงานของแมกซ์พูลลิง.....32 |
| 2.23 | ตัวอย่างผลลัพธ์หลังการทำ ReLU.....33 |
| 2.24 | กราฟฟังก์ชัน ReLU.....33 |
| 2.25 | การเชื่อมต่อกันของแต่ละชั้นอย่างสมบูรณ์ (Fully connected layer).....34 |
| 2.26 | โครงสร้างแบบจำลอง R-CNN.....36 |
| 2.27 | โครงสร้างแบบจำลอง Fast R-CNN.....37 |

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนำไปเผยแพร่โดยไม่ได้รับอนุญาตถือว่าผิดกฎหมาย

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

| ภาพที่ | หน้า |
|--|------|
| 2.28 โครงสร้างแบบจำลอง Faster R-CNN..... | 38 |
| 2.29 โครงสร้าง Faster R-CNN จำแนกกลุ่มและ RPN ช่วยหาตำแหน่งที่น่าจะเป็นวัตถุ..... | 39 |
| 2.30 โครงสร้างของ VGG-16 ที่นำมาใช้เป็นพื้นฐานของโมเดล..... | 39 |
| 2.31 รายละเอียดการทำงานในแต่ละชั้นของโครงสร้าง VGG-16..... | 41 |
| 2.32 ขอบเขตที่ Region proposal network นำมาพิจารณา..... | 43 |
| 3.1 พื้นที่บริเวณล้อและรางเพื่อกำหนดจุดพิจารณา..... | 45 |
| 3.2 ภาพล้อและรางแสดงภาพความเสี่ยงที่อาจเกิดความเสียหายต่อล้อ..... | 46 |
| 3.3 การแปลงรูปภาพจากวิดีโอไปยังรูปภาพทั้งหมด..... | 47 |
| 3.4 ตัวอย่างรูปภาพของล้อและรางที่ใช้ทดสอบการประเมินประสิทธิภาพ..... | 47 |
| 3.5 การเลือกบริเวณที่สนใจซึ่งก็คือบริเวณขอบรางและบริเวณขอบล้อ..... | 48 |
| 3.6 รายละเอียดโครงสร้าง Faster R-CNN..... | 49 |
| 3.7 รูปภาพและป้ายกำกับ..... | 50 |
| 3.8 การเตรียมข้อมูลภาพสำหรับใช้ในขั้นตอนการเรียนรู้..... | 52 |
| 3.9 ขั้นตอนการหาระยะห่างของล้อและรางด้วย Faster R-CNN ด้วยโครงสร้าง VGG-16 โดยกำหนดจุดขอบของล้อรถไฟให้เป็นจุดอ้างอิง..... | 56 |
| 3.10 วิธีการนำเสนอการลดชั้นของโครงสร้าง VGG-16 ที่นำมาใช้เป็นพื้นฐานของโมเดล..... | 57 |
| 3.11 ขั้นตอนการหาระยะห่างของล้อและรางด้วย Faster R-CNN ด้วยโครงสร้างของ VGG-16 แบบลดชั้นโดยกำหนดจุดขอบของล้อรถไฟให้เป็นจุดอ้างอิง..... | 58 |
| 4.1 กราฟ TensorBoard..... | 60 |
| 4.2 ผลลัพธ์ค่าสูญเสีย TensorBoard ที่ทำการรันจำนวน 10,000 รอบ..... | 61 |
| 4.3 ผลลัพธ์การตรวจหาตำแหน่งของขอบล้อและราง..... | 62 |
| 4.4 ผลลัพธ์การวิเคราะห์หาตำแหน่งล้อและรางจากเทคนิคปัญญาประดิษฐ์..... | 63 |
| 4.5 ผลลัพธ์การวัดระยะห่างของตำแหน่งล้อและรางช่วง 1-1,000..... | 63 |
| 4.6 ผลลัพธ์การวัดระยะห่างของตำแหน่งล้อและรางช่วง 1,001-2000..... | 64 |
| 4.7 ผลลัพธ์การวัดระยะห่างของตำแหน่งล้อและรางช่วง 2,001-2,699..... | 65 |
| 4.8 ค่าความผิดพลาดของน้ำหนักที่ทดสอบด้วยเทคนิคปัญญาประดิษฐ์โดยใช้โครงข่ายประสาทเทียมด้วยโครงสร้างของ VGG-16..... | 67 |
| 4.9 ผลจากการวัดระยะห่างของล้อกับรางด้วยการตรวจจับตำแหน่งของรางด้วย Faster R-CNN ด้วยโครงสร้างของ VGG-16..... | 68 |

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และแจ้งอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

| ภาพที่ | หน้า |
|---|------|
| 4.10 การตรวจจับตำแหน่งของรางด้วย Faster R-CNN ด้วยโครงสร้าง VGG-16..... | 69 |
| 4.11 ค่าความผิดพลาดของน้ำหนักที่ทดสอบด้วยเทคนิคปัญญาประดิษฐ์โดยใช้โครงข่ายประสาทเทียมด้วยโครงสร้างของ VGG-16 แบบลดชั้น..... | 70 |
| 4.12 ผลจากการวัดระยะห่างของล้อกับรางด้วยการตรวจจับตำแหน่งของรางด้วย Faster R-CNN ด้วยโครงสร้าง VGG-16 แบบลดชั้น..... | 71 |
| 4.13 การตรวจจับตำแหน่งของรางด้วย Faster R-CNN โดยใช้โครงข่ายประสาทเทียมด้วยโครงสร้างของ VGG-16 แบบลดชั้น..... | 71 |
| 4.14 ระยะห่างระหว่างล้อและรางรถไฟจากการตรวจจับด้วยวิธีการ Faster R-CNN ด้วยโครงสร้างของ VGG-16 กับ VGG-16 แบบลดชั้น..... | 72 |



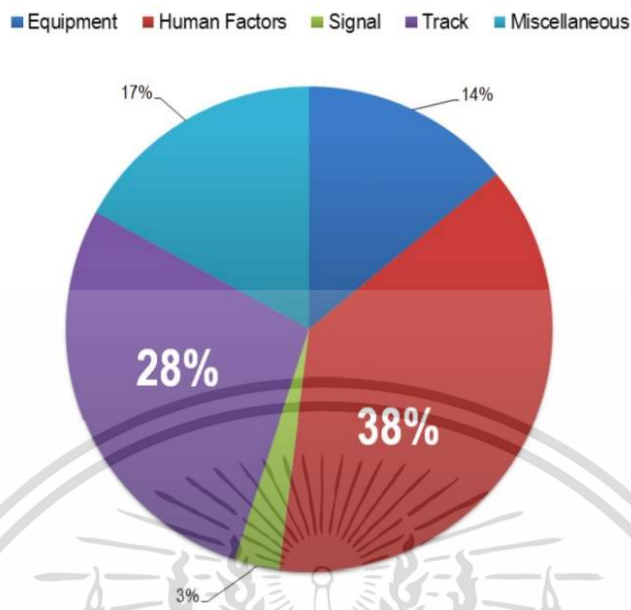
บทที่ 1

บทนำ

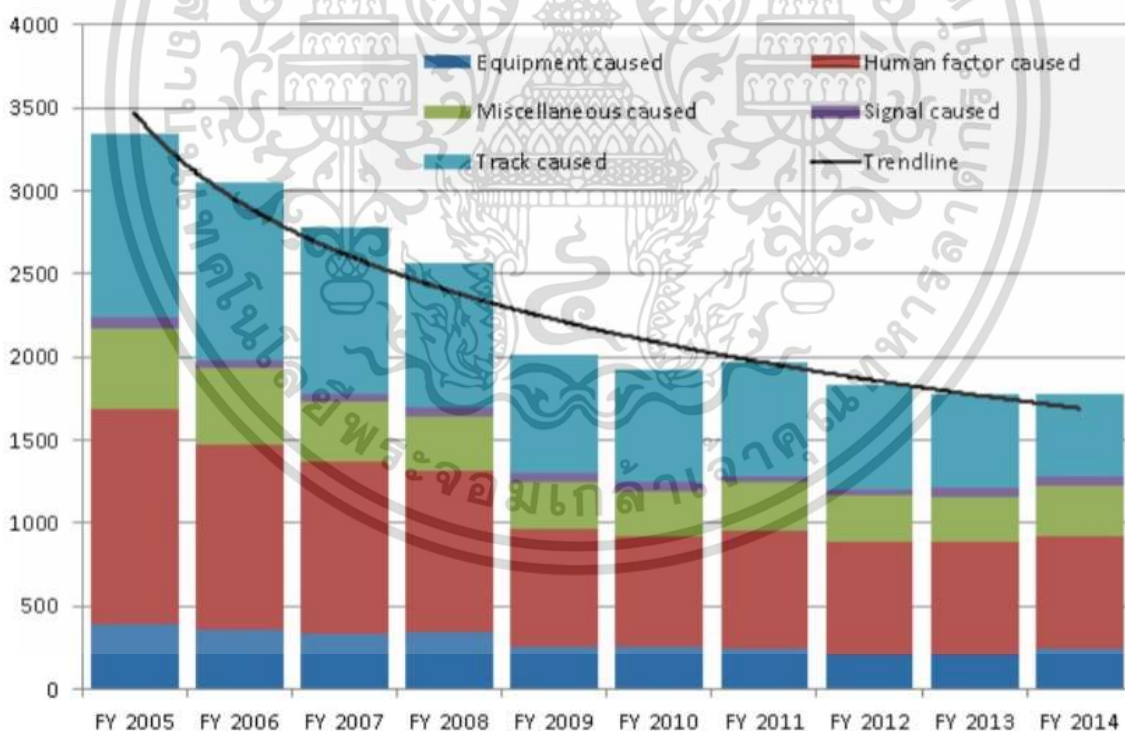
1.1 ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันประเทศไทยมีการพัฒนาระบบขนส่งทางรางโดยเฉพาะการคมนาคมด้วยรถไฟฟ้าทั้งในกรุงเทพฯ และปริมณฑลเพื่อลดปัญหาด้านการจราจรที่ติดขัด รวมทั้งมลพิษทางด้านอากาศ ซึ่งในปัจจุบันระบบการขนส่งทางรางมีการขยายเส้นทางออกหลายเส้นทางเป็นจำนวนมากเพื่อให้ครอบคลุมพื้นที่ทั้งในกรุงเทพฯ และปริมณฑลให้ทั่วถึง ซึ่งจากแผนพัฒนาเศรษฐกิจและสังคมแห่งชาติฉบับที่ 12 และยุทธศาสตร์ชาติระยะ 20 ปี นำไปสู่วิสัยทัศน์ของแผนพัฒนามาตรฐานระบบรางของประเทศไทย พ.ศ. 2560-2570 คือ “มุ่งสู่การยกระดับมาตรฐานระบบขนส่งทางรางของประเทศไทย ในด้านความปลอดภัยและคุณภาพในการให้บริการ ตลอดจนเพิ่มศักยภาพในการแข่งขันของอุตสาหกรรมที่เกี่ยวข้อง ให้เทียบเท่าระดับสากล” [1] จากแผนดังกล่าวจะเห็นได้ว่าการพัฒนาระบบรถไฟฟ้าจำเป็นต้องคำนึงถึงความปลอดภัยในการใช้งานเป็นหลักสำคัญส่งผลให้บริการทางด้านรถไฟฟ้าต้องอยู่ภายใต้มาตรฐานความปลอดภัย โดยจะต้องยึดหลัก RAMS (Reliability, Availability, Maintainability and Safety) กล่าวคือต้องมีมาตรฐานในการดูแลบำรุงรักษาให้ระบบของรถไฟฟ้ามีความพร้อมในการใช้งานทุกด้านทั้งหมดภายใต้มาตรฐานความปลอดภัยที่มีความน่าเชื่อถือ [2]

จากข้อมูลการเกิดอุบัติเหตุทางรางในประเทศสหรัฐอเมริกาที่เก็บรวบรวมโดย FRA (Federal of Railway Association) พบว่าปัจจัยที่ทำให้เกิดอุบัติเหตุทางรางประมาณร้อยละ 28 ดังแสดงในภาพที่ 1.1 ซึ่งมีสาเหตุจากความผิดพลาดที่เกิดจากทางรถไฟและอุปกรณ์ต่างๆบนรถไฟ [3] จากสาเหตุดังกล่าวสามารถทำให้ลดลงได้ภายใต้การตรวจวัด (Inspection) และการวางแผนซ่อมบำรุงที่มีประสิทธิภาพ จากการเก็บข้อมูลโดย FRA ในประเทศสหรัฐอเมริกา จะเห็นได้ว่าอุบัติเหตุภัยของการขนส่งทางรางมีการลดลงอย่างมีนัยสำคัญดังแสดงในภาพที่ 1.2 [3]



ภาพที่ 1.1 แผนภาพแสดงร้อยละของสาเหตุของการเกิดอุบัติเหตุทางรถไฟในประเทศสหรัฐอเมริกา [3]



ภาพที่ 1.2 แผนภาพการลดลงอย่างมีนัยสำคัญของอุบัติเหตุจากการขนส่งทางรางของประเทศสหรัฐอเมริกา [3]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังนั้นหากมีเครื่องมือที่สามารถตรวจวัดความผิดปกติของการทำงานของระบบรางได้อย่างมีประสิทธิภาพต่อการความปลอดภัยในการใช้งานรถไฟฟ้ า เนื่องจากช่วยลดโอกาสการเกิดอุบัติเหตุจากระบบราง โดยวิทยานิพนธ์นี้มีแนวคิดการประยุกต์องค์ความรู้ ได้แก่ ระบบวิทัศน์อัตโนมัติ (vision system) ปัญญาประดิษฐ์ (Artificial Intelligence) และเทคโนโลยีการตรวจวัดแบบตามเวลาจริง (real-time) มาใช้ในงานวิจัยเพื่อตอบโจทย์การช่วยพัฒนาระบบขนส่งทางรางในประเทศไทยให้มีความปลอดภัยต่อผู้ใช้งานมากยิ่งขึ้น

จากสาเหตุของการเกิดอุบัติเหตุของรถไฟฟ้ าซึ่งส่วนใหญ่เกิดจากการทำงานของล้อและรางที่ไม่สอดคล้องกันทำให้เกิดอุบัติเหตุรถไฟตกราง [3] เพื่อแก้ปัญหาดังกล่าววิทยานิพนธ์ฉบับนี้จึงได้นำเสนอการตรวจจับอันตรายกิริยาระหว่างล้อและรางเพื่อสังเกตและประเมินสภาพการทำงานของล้อและรางรถไฟฟ้ า เพื่อให้สามารถป้องกันอุบัติเหตุที่จะเกิดขึ้นได้โดยใช้ระบบวิทัศน์อัตโนมัติมาช่วยเนื่องจากระบบวิทัศน์อัตโนมัติเป็นระบบที่สามารถประมวลผลจากการเก็บข้อมูลรูปภาพซึ่งสามารถทดแทนการสังเกตจากสายตามนุษย์ที่มีข้อดีคือ ไม่มีข้อจำกัดทางเวลาในการใช้งาน ลดความผิดพลาดที่เกิดจากมนุษย์ [4] และสามารถเก็บข้อมูลได้เป็นจำนวนมาก นอกจากนี้ยังสามารถนำรูปภาพที่เก็บได้ไปเข้ากระบวนการอื่นๆเพื่อช่วยวิเคราะห์ในส่วนที่ต้องการได้อีก เช่น Xavier G. และคณะได้พัฒนาสมรรถนะของการตรวจจับอุปกรณ์ยึดรางรถไฟโดยใช้ระบบวิทัศน์อัตโนมัติร่วมกับตัวตรวจจับจำนวนหลายตัวภายในกรอบการเรียนรู้เดียวกันเพื่อลดเวลาในการวิเคราะห์ [4] Junbo L. และคณะนำเสนอ Vision based fastener inspection system (VFIS) ซึ่งเป็นระบบวิทัศน์อัตโนมัติในการตรวจจับอุปกรณ์ยึดรางรถไฟโดยใช้เวลาเพียงเล็กน้อยในขั้นตอนการเรียนรู้ของระบบ [5] Ohan Y. และคณะใช้เทคนิคการประมวลผลภาพถ่ายในการตรวจจับความผิดปกติบนพื้นผิวรางและจำแนกชนิดของความผิดปกติด้วย Fuzzy logic เพื่อป้องกันความเสียหายในการขนส่งทางราง [6] Vonani R. และคณะนำเสนอภาพ 3 มิติในการสร้างแบบจำลองของรางจากการวัดและข้อบกพร่องของรางที่ได้จากการตรวจจับ [7] ส่วน Cahan T. และคณะใช้การประมวลผลภาพในการตรวจจับความเสียหายของพื้นผิวราง [8] จากบทความดังกล่าวจะเห็นได้ว่าการประยุกต์ใช้ประโยชน์จากระบบวิทัศน์อัตโนมัติมีหลากหลายด้าน

ในวิทยานิพนธ์นี้ใช้เทคนิคทางปัญญาประดิษฐ์มาประยุกต์ใช้ร่วมกับระบบวิทัศน์อัตโนมัติ โดยการนำรูปถ่ายของล้อและรางรถไฟฟ้ าจากกล้องที่ติดอยู่กับล้อรถไฟฟ้ มาวัดเพื่อหาระยะห่างและนำข้อมูลที่ได้ออกมาวิเคราะห์หาความปลอดภัยในการเคลื่อนที่ของรถไฟฟ้ าโดยใช้เทคนิค Faster R-CNN เนื่องจากเป็นเทคนิคที่ใช้งานได้แม่นยำจากเทคนิคดังกล่าวสามารถเก็บข้อมูลและวิเคราะห์ผลได้ทันทีทำให้การตรวจจับเป็นแบบตามเวลาจริง (Real time) จากวิธีการดังกล่าวในงานวิจัยนี้ได้นำเสนอการลดชั้นโครงสร้างของ VGG-16 ใน Faster R-CNN เพื่อลดเวลาในการวิเคราะห์ผลให้เร็วขึ้นกว่าเดิม โดยผลที่ได้พบว่าวิธีการที่นำเสนอมีความแม่นยำใกล้เคียงกับเทคนิค Faster R-CNN แต่ใช้เวลาในการวิเคราะห์น้อยกว่า

1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

1.2.1 เพื่อประยุกต์ใช้เทคนิคปัญญาประดิษฐ์ในการวิเคราะห์ความผิดปกติของอันตรกิริยาระหว่างล้อยับราง

1.2.2 เพื่อทดสอบระบบที่นำเสนอเปรียบเทียบกับเทคนิคปัญญาประดิษฐ์ในการวิเคราะห์ความผิดปกติของอันตรกิริยาระหว่างล้อยับราง

1.2.3 สรุปผลเปรียบเทียบข้อดีของการใช้เทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 กับกระบวนการตรวจจับ

1.3 ขอบเขตการวิจัย

วิทยานิพนธ์ฉบับนี้ได้นำเสนอการสร้างระบบตรวจจับวิทัศน์อัตโนมัติที่สามารถตรวจสอบอันตรกิริยาระหว่างล้อยับรางแบบตามเวลาจริง โดยใช้เทคนิคทางปัญญาประดิษฐ์มาประมวลผลภาพเพื่อได้ประสิทธิภาพแม่นยำและรวดเร็ว และนำเสนอการลดชั้นโครงสร้างของ VGG-16 เพื่อลดเวลาในการวิเคราะห์ผลในการประมวลผลภาพ

1.4 ขั้นตอนการศึกษา

ในการศึกษาได้แบ่งขั้นตอนการศึกษาไว้โดยมีขั้นตอนดังนี้

- ก. ศึกษาและหาแนวทางในการในการวิเคราะห์ความความปลอดภัยในการเคลื่อนที่ของรถไฟฟ้
- ข. ใช้เทคนิคปัญญาประดิษฐ์ในการวิเคราะห์ความผิดปกติของอันตรกิริยาระหว่างล้อยับราง
- ค. ลดชั้นของโครงสร้าง VGG-16 ลงเพื่อลดเวลาในการวิเคราะห์ความผิดปกติของอันตรกิริยาระหว่างล้อยับรางด้วยเทคนิคปัญญาประดิษฐ์
- ง. ตรวจสอบผลและสรุปผลการเปรียบเทียบข้อดีของการใช้เทคนิคต่างๆกับกระบวนการตรวจจับ

1.5 โครงสร้างของวิทยานิพนธ์

วิทยานิพนธ์ฉบับนี้ได้แบ่งเนื้อหาออกเป็น 5 บทด้วยกันคือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

กล่าวถึงความเป็นมาของงานวิจัย ความมุ่งหมาย วัตถุประสงค์ ขอบเขตการวิจัย และขั้นตอนในการศึกษาของการประยุกต์ใช้เทคนิคปัญญาประดิษฐ์ในการวิเคราะห์ความผิดปกติของอันตรกิริยาระหว่างล้อยับราง

บทที่ 2

กล่าวถึงทฤษฎีพื้นฐานในการประมวลผลรูปภาพ ทฤษฎีทางด้านปัญญาประดิษฐ์ ได้แก่ โครงข่ายประสาทเทียมและ Faster R-CNN ที่ใช้ในการวิเคราะห์ความผิดปกติของอันตรกิริยาระหว่างล้อยับราง

บทที่ 3

กล่าวถึงการเตรียมภาพเพื่อใช้ในการวิเคราะห์ด้วย Faster R-CNN การเตรียมภาพเพื่อใช้ในการวิเคราะห์ด้วย Faster R-CNN โดยกำหนดให้ขอบลือเป็นจุดอ้างอิง รวมถึงวิธีการลดชั้นโครงสร้างของ VGG-16 ที่นำเสนอ

บทที่ 4

กล่าวถึงผลจากการวิเคราะห์ด้วย Faster R-CNN ผลจากการวิเคราะห์ด้วย Faster R-CNN โดยกำหนดให้ขอบลือเป็นจุดอ้างอิง และผลจากการวิเคราะห์ด้วย Faster R-CNN โดยทำการลดชั้นของ VGG-16 ที่นำเสนอ

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

บทที่ 2

ทฤษฎีพื้นฐานที่ใช้ในการวิจัย

ในบทนี้จะกล่าวถึงทฤษฎีที่เกี่ยวข้องกับงานในวิทยานิพนธ์นี้โดยกล่าวถึงระบบวิทัศน์อัตโนมัติ เทคนิคการประมวลผลภาพและการจำแนกคุณลักษณะควบคู่กับการใช้ปัญญาประดิษฐ์ รวมถึงการแปลงเวฟเล็ตแบบเต็มหน่วยที่ใช้ในงานวิจัย

การทำงานของระบบวิทัศน์อัตโนมัติ เป็นการนำภาพมาประมวลผลหรือคิดคำนวณด้วยคอมพิวเตอร์ เพื่อให้ได้ข้อมูลที่ต้องการทั้งในเชิงคุณภาพและเชิงปริมาณ มีขั้นตอนต่างๆที่สำคัญ คือ ต้องทำให้ภาพมีความคมชัดมากขึ้น จากนั้นทำการกำจัดสัญญาณรบกวนออกจากภาพ แบ่งส่วนของวัตถุที่เราสนใจออกมาจากภาพ แล้วจึงนำภาพวัตถุที่ได้ไปวิเคราะห์หาข้อมูลเชิงปริมาณ เช่น ขนาด รูปร่าง และทิศทางการเคลื่อนของวัตถุในภาพ ทำให้สามารถนำข้อมูลเหล่านี้ไปวิเคราะห์ และสามารถสร้างเป็นระบบเพื่อใช้ประโยชน์ในงานด้านต่างๆ เช่น Krishna และคณะนำเสนอระบบติดตามการจราจรอัตโนมัติด้วยเส้นอ้างอิงบนถนนจากภาพในวิดีโอโดยสามารถตรวจจับความเร็วของรถได้ด้วย [9] Bhaskar B. นำเสนอการขับเคลื่อนเองของรถยนต์โดยใช้ระบบวิทัศน์ในการตรวจจับตำแหน่งและนำทางการเคลื่อนที่ของรถยนต์ [10] Judith S. และคณะนำเสนอระบบจดจำป้ายทะเบียนรถยนต์ชั้นสูงสำหรับลานจอดรถ [11] Harikrishnan J. นำเสนอระบบติดตามการจดจำใบหน้าผู้เข้ามามีส่วนร่วมสำหรับฝ้าระวัง [12] S.M Amir K. นำเสนอระบบทางเข้าอัจฉริยะที่ใช้คอมพิวเตอร์วิทัศน์ในสภาพแวดล้อมองค์กร [13] เป็นต้น จะเห็นได้ว่าระบบเหล่านี้จำเป็นต้องมีการประมวลผลจากรูปภาพจำนวนมากและเป็นกระบวนการที่ต้องทำซ้ำๆกันในรูปแบบเดิม ซึ่งกระบวนการในลักษณะนี้หากใช้มนุษย์เป็นผู้วิเคราะห์คงต้องใช้เวลาและจำนวนคนปริมาณมาก อีกทั้งผู้วิเคราะห์ภาพอาจเกิดความเหนื่อยล้าจากการทำงานอาจส่งผลให้เกิดความผิดพลาดได้ ดังนั้นคอมพิวเตอร์จึงมีเข้ามามีบทบาทสำคัญในการทำหน้าที่เหล่านี้แทนมนุษย์นั่นเอง เนื่องจากคอมพิวเตอร์มีความสามารถในการคำนวณและประมวลผลข้อมูลจำนวนมากได้ในเวลาอันสั้น ทำให้สามารถเพิ่มประสิทธิภาพในการประมวลผลภาพและวิเคราะห์ข้อมูลต่างๆที่ต้องการจากภาพได้ ในบทนี้จะกล่าวถึงทฤษฎีโดยสังเขปของส่วนประมวลผลภาพและปัญญาประดิษฐ์ โดยใช้เทคนิคการประมวลผลภาพและการจำแนกคุณลักษณะมาพัฒนาควบคู่กับการใช้ปัญญาประดิษฐ์ โดยงานวิจัย [14] Dr.Ulagamuthalvi และคณะได้นำเสนอประสิทธิภาพแบบจำลองการตรวจจับวัตถุโดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network) Wei Z. และคณะนำเสนอ DECONV R-CNN สำหรับตรวจจับวัตถุขนาดเล็กบนการตรวจจับภาพระยะไกล Yan Z. และ Ying H. นำเสนอการตรวจจับควันในวิดีโอด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชัน [16] ในการพัฒนาโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network) จะเอาส่วนสำหรับการคัดแยกคุณลักษณะมาใช้ร่วมกับค่าน้ำหนัก (weight) ที่สามารถปรับตัวตามอินพุต (input) และ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอาต์พุต (output) ของโครงข่ายประสาทเทียม (Artificial Neural Network) ทำให้ได้ชุดคุณลักษณะที่ละเอียดและครบถ้วนในการจำแนกรูปภาพทั่วไป มีการเพิ่มขึ้นขั้นตอนของการทำแมกซ์พูลลิง (max pooling) เพื่อช่วยลดขนาดของข้อมูลที่ไม่จำเป็นหลังจากทำการคอนโวลูชันในขั้นตอนการหาคุณลักษณะของวัตถุที่จะทำการคัดแยกประเภทของวัตถุ ซึ่งส่งผลให้การคำนวณมีความรวดเร็วขึ้นอีกด้วย

2.1 การประมวลผลภาพ

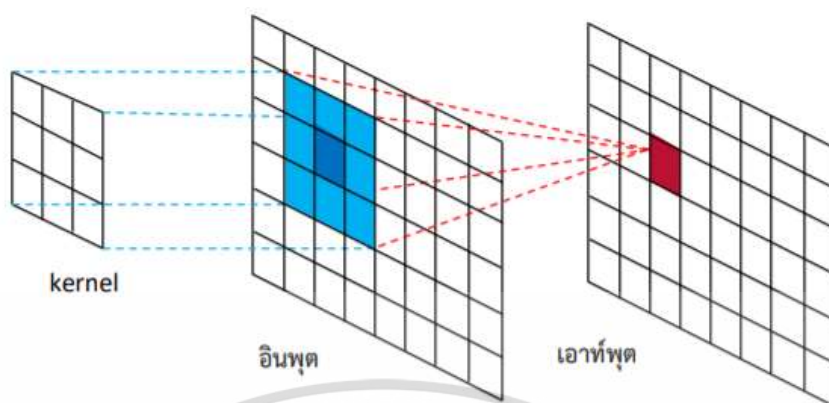
สำหรับการประมวลผลภาพนั้น ทัศนวิสัยที่เกี่ยวข้องจะถูกแบ่งออกเป็น 2 ส่วนหลักคือ ส่วนการประมวลผลภาพเบื้องต้นและส่วนของการจำแนกรูปภาพ ส่วนของการประมวลผลเบื้องต้นจะประกอบไปด้วย การปรับปรุงคุณภาพของภาพระดับเทา การแบ่งส่วนภาพ การปรับปรุงคุณภาพของภาพและการวิเคราะห์ข้อมูลภาพ

2.1.1 การปรับปรุงคุณภาพของภาพระดับเทา

ในการประมวลผลทางภาพนั้นการปรับปรุงคุณภาพของภาพมีเป้าหมายเพื่อต้องการลดสัญญาณรบกวนภายในภาพจากสภาพแวดล้อม ในการเก็บภาพนั้นคุณภาพของระบบตรวจจับจะถูกรบกวนด้วยสัญญาณรบกวนต่างๆ เช่น แสง อุณหภูมิจากกล้องที่สูงขึ้น การแปลงสัญญาณภาพจากตัวรับภาพเป็นข้อมูลดิจิทัล ดังนั้นการกำจัดสัญญาณรบกวนจึงมีความจำเป็นอย่างยิ่งเพื่อลดความผิดพลาดหรือข้อมูลที่ไม่จำเป็นต่อการวิเคราะห์ในขั้นตอนถัดไป ในการปรับปรุงคุณภาพของภาพสามารถทำได้โดยการใช้ตัวกรองชนิดต่างๆ ดังนี้

2.1.1.1 การทำงานของตัวกรองแบบคอนโวลูชัน (Convolutional filter)

ตัวกรองที่ใช้สำหรับการปรับปรุงภาพขั้นต้นนั้นจะใช้หน้าต่างหรือมาสก์ (mask) ที่กำหนดขึ้นมาแทนฟังก์ชันของตัวกรองและใช้กระบวนการคอนโวลูชัน (convolution operation) ในส่วนของมาสก์ที่นำไปซ้อนทับกับภาพอินพุต โดยสามารถอธิบายด้วยภาพที่ 2.1 กับสมการที่ (2.1) และ (2.2)



ภาพที่ 2.1 การทำงานของตัวกรองด้วยวิธีคอนโวลูชัน [17]

$$[x, y] = [x, y] * G[x, y] \quad (2.1)$$

$$= \sum \sum [n1, 2] \cdot G[x - n1, y - n2] \infty n2 \quad (2.2)$$

เมื่อ $[x, y]$ คือตัวกรอง $X[x, y]$ คือรูปต้นฉบับ $Y[x, y]$ คือภาพที่ผ่านตัวกรอง และ $n1, n2$ คือสมาชิกในตัวกรอง

2.1.1.2 ตัวกรองแบบเกาส์ (Gaussian's filter)

ตัวกรองแบบเกาส์ (Gaussian's filter) เป็นตัวกรองสำหรับปรับภาพให้นุ่มนวล (smoothing filter) ตัวกรองประเภทนี้ทำหน้าที่ในการลดปริมาณสัญญาณรบกวนภายในภาพหรือลดผลของการเปลี่ยนแปลงค่าความเข้มของแสงอย่างรวดเร็ว ทำให้ภาพที่ผ่านตัวกรองประเภทนี้มีสัญญาณรบกวนน้อยลงแต่สวดลายหรือขอบในภาพมีความเบลอ ซึ่งผลของความเบลอจะขึ้นอยู่กับค่าสัมประสิทธิ์ในตัวกรองและขนาดของตัวกรองโดยตัวกรองแบบเกาส์นั้นจะมีสมาชิกภายในตัวกรองดังภาพที่ 2.2

| | | | | |
|---|----|----|----|---|
| 1 | 4 | 7 | 4 | 1 |
| 4 | 16 | 26 | 16 | 4 |
| 7 | 26 | 41 | 26 | 7 |
| 4 | 16 | 26 | 16 | 4 |
| 1 | 4 | 7 | 4 | 1 |

ภาพที่ 2.2 สมาชิกในตัวกรองแบบเกาส์ [18]

2.1.1.3 ตัวกรองแบบโซเบล (Sobel's Filter)

ตัวกรองแบบโซเบลเป็นตัวกรองค่าความชันสำหรับตรวจจับแนวของการเปลี่ยนแปลงค่าความเข้มแสงในทิศทางต่างๆ โดยแนวการเปลี่ยนแปลงความเข้มแสงจะตั้งฉากกับแกนสมมาตรของตัวกรองสำหรับตัวกรองแบบโซเบลนั้นจะประกอบไปด้วยตัวกรอง 2 ตัวกรอง ขนาด 3×3 สำหรับตรวจจับค่าความเข้มแสงในแนวตั้งและแนวนอนโดยมีสมาชิกดังภาพที่ 2.3

| | | | | | |
|----|----|----|----|---|----|
| +1 | +2 | +1 | +1 | 0 | +1 |
| 0 | 0 | 0 | +2 | 0 | -2 |
| -1 | -2 | -1 | -1 | 0 | -1 |

(ก)

(ข)

ภาพที่ 2.3 สมาชิกของตัวกรอง (ก) สมาชิกสำหรับตัวกรองแนวนอน (ข) สมาชิกสำหรับตัวกรองแนวตั้ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.2 การแบ่งส่วนภาพ (Image Segmentation)

การแบ่งส่วนภาพเป็นขั้นตอนที่สำคัญในการประมวลผลทางภาพ ขั้นตอนนี้จะช่วยแบ่งแยกส่วนของรูปภาพเฉพาะส่วนที่สนใจออกมาจากภาพส่วนที่เป็นพื้นหลังเพื่อให้สามารถวิเคราะห์ในส่วนที่แบ่งแยกออกมาได้ ส่งผลให้การประมวลผลมีความแม่นยำและถูกต้อง โดยการทำการเทรซโฮลด์ภาพนั้นจะทำให้ภาพระดับเทากลายเป็นภาพสองระดับซึ่งมีบริเวณที่สนใจมีค่าเป็น 1 และบริเวณอื่น ๆ จะมีค่าเป็น 0 สามารถเขียนแทนด้วยสมการที่ (2.3)

$$(x, y) = \{1, \text{if } f(x, y) \geq T\} \quad 0, \text{otherwise} \quad (2.3)$$

เมื่อ (x, y) คือภาพที่ผ่านการเทรซโฮลด์ของ $f(x, y)$ โดยกำหนดค่าเทรซโฮลด์เท่ากับ T

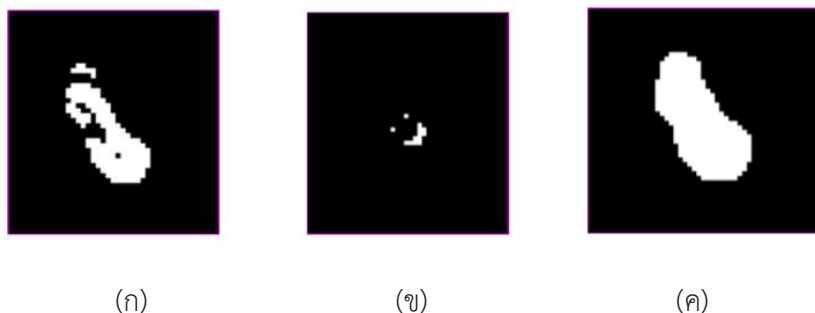
2.1.3 กระบวนการทางคณิตศาสตร์สำหรับปรับปรุงภาพ (Morphological)

กระบวนการสำหรับการปรับปรุงคุณภาพเป็นขั้นตอนสำหรับการเติมภาพหรือตัดภาพบางส่วนด้วยกระบวนการทางคณิตศาสตร์กับภาพสองระดับเมื่อต้องการปรับปรุงรูปทรงของอนุภาคภาพสองระดับ ภาพสองระดับประกอบด้วยบริเวณที่มีอนุภาคซึ่งกำหนดให้ค่าของพิกเซลเป็น 1 และค่าในส่วนอื่นๆเป็น 0 โดยได้มาจากกระบวนการแบ่งภาพด้วยการเทรซโฮลด์ กระบวนการพื้นฐานในการทำงานจะประกอบไปด้วย 2 ขั้นตอน ได้แก่ การกร่อน (erosion) คือกระบวนการที่ใช้กำจัดจุดพิกเซลในส่วนที่เป็นพื้นหลังของภาพหรือการลดขนาดของพิกเซลตามลักษณะของหน้าต่างโครงสร้างที่ใช้และการขยาย (dilation) คือกระบวนการที่ใช้กำจัดรูหรือพื้นหลังในกลุ่มของอนุภาคของรูปภาพหรือการขยายขนาดของพิกเซลให้มีลักษณะตามหน้าต่างโครงสร้างที่ใช้ โดยกระบวนการทั้งสองแบบที่จะดำเนินการด้วยหน้าต่างขนาด 3×3 มีสมาชิกภายในดังภาพที่ 2.4 และมีภาพต้นแบบเป็นภาพที่ 2.5 (ก) โดยได้ผลการดำเนินการดังภาพที่ 2.5 (ข) และ ภาพที่ 2.5 (ค) ตามลำดับ

| | | |
|---|---|---|
| 1 | 1 | 1 |
| 1 | 1 | 1 |
| 1 | 1 | 1 |

ภาพที่ 2.4 หน้าต่างโครงสร้าง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

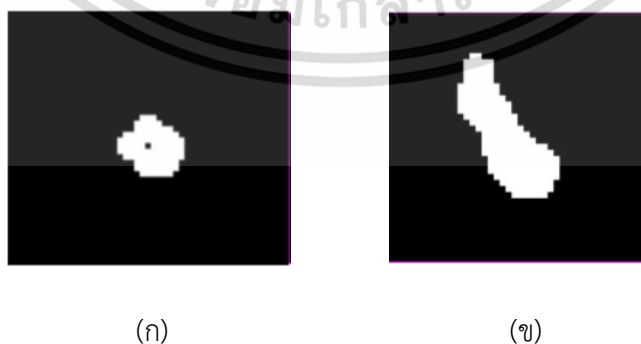


ภาพที่ 2.5 (ก) ภาพต้นฉบับ (ข) ภาพหลังจากการทำการกร่อน (ค) ภาพหลังจากการทำการขยาย

เมื่อรวมกระบวนการทั้งสองเข้าด้วยกันจะสามารถจำแนกลักษณะการทำงานตามลำดับการดำเนินการตามกระบวนการได้ดังนี้

1. การเปิด (opening) หรือการตัดการเชื่อมต่อของกลุ่มอนุภาคออกจากกันเพื่อให้สามารถดำเนินการวิเคราะห์อย่างละเอียดได้เนื่องจากมีจำนวนของอนุภาคหรือกลุ่มของพิกเซลที่เพิ่มมากขึ้น โดยลำดับการดำเนินการจะเป็นดังนี้ 1. ทำการกร่อน 2. ทำการขยาย เมื่อให้ภาพต้นฉบับเป็น ภาพที่ 2.6 (ก) และภาพที่ผ่านกระบวนการเป็นภาพที่ 2.6 (ข)

2. การปิด (closing) หรือการเชื่อมต่ออนุภาคสองส่วนเข้าด้วยกันเพื่อให้สามารถดำเนินการวิเคราะห์อย่างรวดเร็วได้จากจำนวนของพิกเซลส่วนอื่นได้ถูกรวมเข้าเป็นกลุ่มเดียวกันแล้วนั่นเอง โดยลำดับการดำเนินการจะเป็นดังนี้ 1. ทำการกร่อน 2. ทำการขยาย เมื่อให้ภาพต้นฉบับ เป็นภาพที่ 2.6 (ก) และภาพที่ผ่านกระบวนการเป็นภาพที่ 2.6 (ข)



ภาพที่ 2.6 (ก) ภาพหลังจากกระบวนการเปิด (ข) ภาพหลังจากกระบวนการปิด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.4 การวิเคราะห์ข้อมูลภาพ

หลังจากได้ภาพสองระดับแล้ว ขั้นตอนการวิเคราะห์ข้อมูลของภาพจะถูกนำมาใช้เพื่อเก็บข้อมูลจากภาพสองระดับแล้วไปทำกระบวนการในขั้นต่อไปในวิทยานิพนธ์นี้จะใช้ตำแหน่งของอนุภาคของภาพ (particle) เพื่อกำหนดขอบเขตที่สนใจ (Region of interest: ROI) สำหรับนำไปกำหนดจุดที่ต้องการนำไปวิเคราะห์ในขั้นตอนการจำแนกภาพ

2.1.4.1 การอ่านค่าจากลักษณะของอนุภาค

อนุภาคคือกลุ่มของพิกเซลของภาพสองระดับที่อยู่ติดกันโดยสามารถทำการตรวจวัดหรือดึงข้อมูลออกมาเพื่อบ่งบอกคุณสมบัติหรือแบ่งแยกชนิดของอนุภาค โดยการวัดตำแหน่งของอนุภาคเพื่อนำไปตีกรอบจุดสนใจของภาพโดยที่ข้อมูลที่น่าไปใช้จะถูกแสดงใน ตารางที่ 2.1

ตารางที่ 2.1 การวิเคราะห์ข้อมูลจากอนุภาค

| ข้อมูลที่วิเคราะห์ | ความหมายของข้อมูล |
|---------------------------|---------------------------------|
| Center of mass X | จุดศูนย์กลางมวลในแกน X |
| Center of mass Y | จุดศูนย์กลางมวลในแกน Y |
| Bounding rectangle left | กรอบสี่เหลี่ยมด้านซ้ายของอนุภาค |
| Bounding rectangle top | กรอบสี่เหลี่ยมด้านบนของอนุภาค |
| Bounding rectangle right | กรอบสี่เหลี่ยมด้านขวาของอนุภาค |
| Bounding rectangle bottom | กรอบสี่เหลี่ยมด้านล่างของอนุภาค |

2.1.4.2 ขอบเขตที่สนใจ (Region of interest: ROI)

ขอบเขตที่สนใจของภาพ (Region of interest: ROI) คือบริเวณเฉพาะส่วนของภาพที่สนใจสำหรับการประมวลผลเพื่อนำไปวิเคราะห์ข้อมูล การจำกัดจุดสนใจมีประโยชน์ในการช่วยจำกัดขอบเขตของการประมวลผล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับภาพที่มีขนาดใหญ่โดยการจำกัดการประมวลผลเฉพาะบริเวณที่กำหนดขึ้นและช่วยลดความผิดพลาดจากข้อมูลในส่วนที่ไม่เกี่ยวข้อง การกำหนดขอบเขตนั้นจะนิยมกำหนดเป็นรูปทรงเลขาคณิตอย่างง่าย เช่น สี่เหลี่ยมหรือวงกลม ภาพที่ 2.7 เลข 1 คือขอบเขตที่สนใจซึ่งเป็นกรอบสี่เหลี่ยมสีแดงจะถูกกำหนดขึ้นจากค่าที่อ่านได้จากการวิเคราะห์อนุภาคจากตารางที่ 2.1



ภาพที่ 2.7 ภาพที่ถูกกำหนดจุดสนใจ

2.2 การจำแนกรูปภาพ

ในงานการประมวลผลทางภาพ การจำแนกรูปภาพเป็นขั้นตอนในการจัดกลุ่มภาพที่ต้องการหรือในการแยกประเภทของรูปภาพที่มีลักษณะใกล้เคียงหรือเหมือนกับรูปภาพต้นฉบับเอาไว้ในกลุ่มเดียวกัน เช่น การแยกประเภทตัวอักษรภาษาอังกฤษ ตัวเลข รวมไปถึงการจำแนกชนิดของรูปภาพ ต่างๆ เช่น คน สัตว์ หรือสิ่งของ

วิธีการจำแนกรูปภาพแบบง่ายที่นิยมใช้ในงานประมวลผลทางภาพ [19]-[22] นั้น ส่วนใหญ่จะเป็นวิธีการที่ใช้ในการหาคุณลักษณะเด่นของภาพที่ตรวจจับหรือภาพที่ต้องการจำแนกเพื่อใช้เป็นตัวแทนในการจัดกลุ่มของรูปภาพด้วยการหาคุณลักษณะเด่นจากการประมวลผลภาพขั้นต้น จากนั้นนำไปใช้ในการคำนวณทางสถิติเพื่อจัดกลุ่มข้อมูล ซึ่งวิธีนี้เป็นวิธีที่ง่ายและรวดเร็วแต่ก็ยังมีข้อเสียที่ทำให้การจำแนกรูปภาพหรือการแยกประเภทผิดไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากที่ต้องการคือ การเลือกใช้คุณลักษณะเด่นของภาพแต่ละประเภทผิดไปจากความเป็นจริงหรือการเลือกคุณลักษณะเด่นของภาพไม่ครบถ้วนทำให้การจัดกลุ่มหรือแยกประเภทไม่ถูกต้องตามที่ต้องการ ด้วยเหตุนี้จึงต้องใช้วิธีการประมวลผลภาพขั้นสูงเข้ามาช่วยเพื่อการจัดกลุ่มโดยใช้เทคนิคปัญญาประดิษฐ์ด้วยโครงข่ายประสาทเทียม ซึ่งมีจุดเด่นในการใช้วิธีการทางสถิติสำหรับการประเมินความผิดพลาดจากเอาต์พุตที่เป็นคำตอบ หากภาพที่จำแนกโดยโครงข่ายประสาทเทียมมีความผิดพลาดเกินกว่าที่กำหนดเพื่อนำมาแก้ไขหรือปรับปรุงในส่วนของตัวกรองที่กำหนดขึ้นเพื่อเลือกคุณลักษณะและสามารถจำแนกประเภทของรูปภาพที่ต้องการได้อย่างถูกต้องและแม่นยำกว่า

2.2.1 ปัญญาประดิษฐ์

ปัญญาประดิษฐ์ (Artificial intelligence: AI) เป็นศาสตร์แขนงหนึ่งทางด้านวิทยาศาสตร์และเทคโนโลยีที่มีพื้นฐานมาจากวิชาวิทยาการคอมพิวเตอร์ ชีววิทยา จิตวิทยา ภาษาศาสตร์ คณิตศาสตร์และวิศวกรรมศาสตร์ที่เกี่ยวข้องกับวิธีการทำให้คอมพิวเตอร์มีความสามารถคล้ายมนุษย์หรือเลียนแบบพฤติกรรมมนุษย์โดยเฉพาะความสามารถในการคิดเองได้หรือมีปัญญา ปัญหานี้มนุษย์เป็นผู้สร้างให้คอมพิวเตอร์จึงเรียกว่าปัญญาประดิษฐ์

ปัญญาประดิษฐ์หรือเรียกอีกอย่างหนึ่งว่า AI คือวิธีการพัฒนาเครื่องคอมพิวเตอร์ในรูปแบบหนึ่งที่ทำให้เครื่องคอมพิวเตอร์สามารถคิด วิเคราะห์ แยกแยะ และตัดสินใจทำสิ่งต่างๆได้ใกล้เคียงกับมนุษย์เพื่อเป็นแนวทางในการพัฒนาเครื่องคอมพิวเตอร์ให้ต่างจากเครื่องจักรกลหรือโปรแกรมทั่วไป ส่งผลให้คอมพิวเตอร์สามารถทำงานได้ซับซ้อนและมีประสิทธิภาพมากขึ้น

มุมมองต่อปัญญาประดิษฐ์ที่แต่ละคนมองอาจไม่เหมือนกันขึ้นอยู่กับว่าการต้องการความฉลาดโดยคำนึงถึงพฤติกรรมที่มีต่อสิ่งแวดล้อมหรือค่านึงการคิดได้ของผลผลิตของปัญญาประดิษฐ์

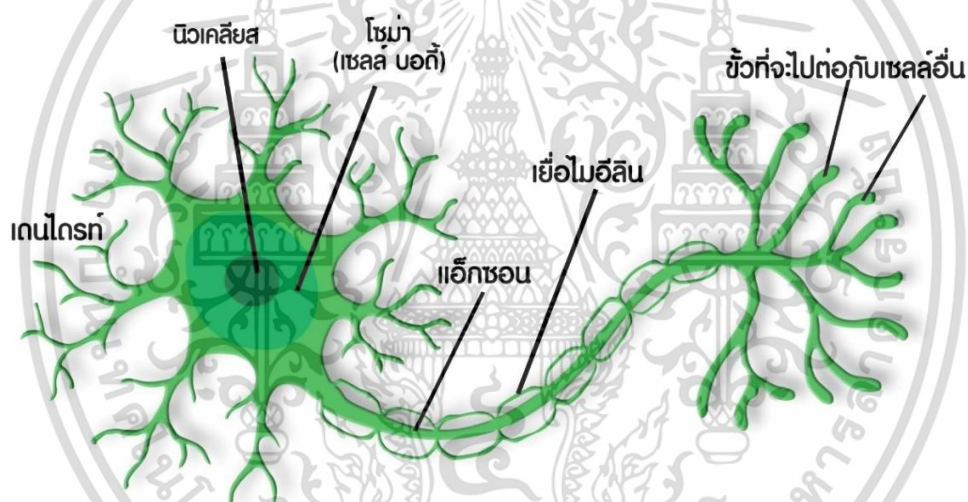
2.2.2 โครงข่ายประสาทเทียม

โครงข่ายประสาทเทียม (Neural Network) หรือ NN เป็นแนวคิดที่ถูกออกแบบให้ทำงานเช่นเดียวกับสมองของมนุษย์ซึ่งประกอบไปด้วยหน่วยประมวลผล (processing elements) ซึ่งมีเซลล์หลายๆตัวที่ทำหน้าที่คล้ายกับเซลล์สมองของมนุษย์ โดยแต่ละเซลล์จะโยงใยติดต่อกันโดยส่งสัญญาณออกเป็นเอาต์พุตออกมาในส่วนที่เรียกว่า แอ็กซอน (axon) ในแต่ละเซลล์จะรับรู้ข้อมูลจากหลายทางแล้วส่งต่อไปยังเซลล์อื่นๆโดยใช้หลักการกำลังของไซแนปส์ (synaptic strength) ของการเชื่อมโยงเซลล์สมอง ส่วนวิธีการประมวลผลภายในโดยเซลล์ประสาทแต่ละเซลล์จะมีจุดเชื่อมโยงระหว่างการทำงาน เป็น 2 ลักษณะ คือลักษณะการกระตุ้น (excitatory) เป็นการทำให้สัญญาณที่ส่งผ่านเข้ามามีความถี่ลดลงซึ่งแบบจำลองจะมีอัตราขยายหรือหดได้เมื่อถูกกำหนดด้วยค่าถ่วงน้ำหนัก (weights) สำหรับความสัมพันธ์ระหว่างเซลล์ประสาทกับเซลล์ประสาทเทียม แสดงไว้ในตารางที่ 2.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2.2 คำอธิบายเซลล์ประสาทและเซลล์ประสาทเทียม

| ลำดับที่ | เซลล์ประสาท | เซลล์ประสาทเทียม |
|----------|----------------------|-------------------------|
| 1 | ตัวเซลล์ (Cell body) | ยูนิต (Unit) |
| 2 | เดนไดรต์ (Dendrites) | ตัวแปรอินพุต (Input) |
| 3 | แอกซอน (Axon) | ตัวแปรเอาต์พุต (Output) |
| 4 | ไซแนปส์ (Synapse) | ค่าถ่วงน้ำหนัก (Weight) |



ภาพที่ 2.8 ส่วนประกอบของโครงข่ายประสาทเทียม [23]

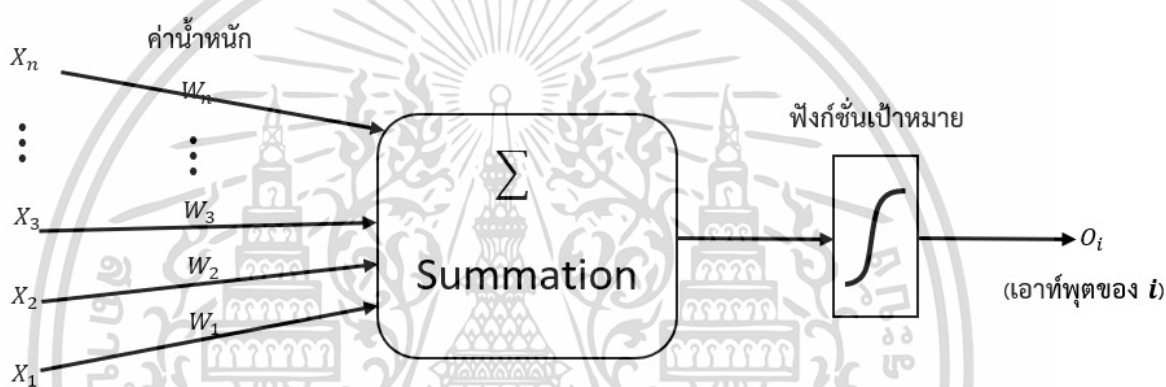
โครงข่ายประสาทเทียมเป็นแบบจำลองการทำงานของระบบประสาทส่วนกลางมีโครงสร้างเป็นลักษณะของโครงข่ายเชื่อมโยงกันระหว่างหน่วยสามารถที่จะรับรู้ข้อมูลและปรับตัวเข้ากับสถานการณ์หรือสิ่งแวดล้อมที่กำลังเผชิญอยู่ แบบจำลองโครงข่ายประสาทเทียมเป็นเครื่องมือชนิดหนึ่งที่ใช้ในการสร้างระบบคอมพิวเตอร์อัจฉริยะ (Intelligent computer system) อย่างได้ผล นอกจากนี้ ANNs (Artificial Neural Networks) ยังเป็นเครื่องมือที่มีประสิทธิภาพในการประยุกต์ใช้งานด้านต่างๆที่เกี่ยวข้องกับการคำนวณและการจดจำ เช่น การ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จำแนกข้อมูล (data classification) การทำนายเหตุการณ์ (forecasting) การบีบอัดข้อมูล (data Compression) การกรองสัญญาณรบกวน (noise filtering) เป็นต้น

2.2.3 การเรียนรู้ของโครงข่ายประสาทเทียม

การเรียนรู้ของโครงข่ายประสาทเทียมแบ่งเป็น 2 รูปแบบคือ การเรียนรู้แบบมีครูสอน (supervised learning) และการเรียนรู้แบบไม่มีครูสอน (unsupervised learning)



ภาพที่ 2.9 การเรียนรู้ของแบบจำลองโครงข่ายประสาทเทียม

การหาผลลัพธ์ค่าเอาต์พุตจากนิวรอน i

$$V_i = W_1X_1 + W_2X_2 + \cdots W_nX_n = \sum_n W_n X_n \quad (2.4)$$

$$O_i = f(V_i) \quad (2.5)$$

เมื่อ X_i คือ ค่าอินพุตที่ i

W_i คือ ค่าถ่วงน้ำหนักที่ i

V_i คือ ผลรวมของผลคูณค่าถ่วงน้ำหนักกับค่าอินพุต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

f คือ ฟังก์ชันการกระตุ้น

O_i คือ ค่าผลลัพธ์ค่าเอาต์พุตจากนิวรอน i เพื่อส่งผ่านไปนิวรอนถัดไป

2.2.3.1 การเรียนรู้แบบมีครูสอน (Supervised learning)

การเรียนรู้แบบมีครูสอนจะมีข้อมูลอินพุตจำเป็นต้องมีเป้าหมายเพื่อให้โครงข่ายประสาทเทียมเรียนรู้โดยการปรับค่าถ่วงน้ำหนักก่อนนำไปหาคำตอบ ซึ่งมีรายละเอียด ดังนี้

2.2.3.1.1 การทำแมปปิงโมเดล (Generalization of mapping model) ด้วย

ฟังก์ชันกระตุ้น

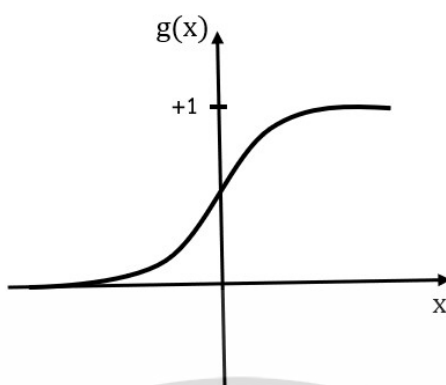
ในการทำแมปปิงโมเดลเพื่อหาคำตอบจำเป็นต้องอาศัยฟังก์ชันกระตุ้นในการดำเนินงาน โดยฟังก์ชันกระตุ้นมีหลายลักษณะ ในวิทยานิพนธ์นี้ใช้ฟังก์ชันซิกมอยด์ (Sigmoid function) แบบไบนารี ซิกมอยด์ในการแปลงค่าซึ่งมีค่าอยู่ในช่วง 0 ถึง 1 เท่านั้นในการนำไปแก้ปัญหา โดยฟังก์ชันซิกมอยด์แสดงดังสมการที่ 2.6 หรือสมการที่ 2.7

$$g(x) = \frac{1}{1+e^{-x}} \quad (2.6)$$

หรือ

$$g(x) = (1 + e^{-x})^{-1} \quad (2.7)$$

กราฟของฟังก์ชันซิกมอยด์แสดงดังภาพที่ 2.10



ภาพที่ 2.10 กราฟของฟังก์ชันซิกมอยด์

จากภาพที่ 2.10 แสดงกราฟของฟังก์ชันซิกมอยด์ ในการหาอนุพันธ์ของฟังก์ชันซิกมอยด์ แสดงดังสมการที่

2.8

$$\frac{d}{dx} g(x) = \frac{e^{-x}}{(1+e^{-x})^2} \cdot \frac{(1+e^{-x}-1)}{(1+e^{-x})^2} \quad (2.8)$$

$$\therefore g'(x) = g(x)(1-g(x)) \quad (2.9)$$

2.2.3.1.2 โครงข่ายประสาทเทียมแบบย้อนกลับ (Backpropagation neural network) แบบหลายชั้น

เป็นวิธีการเรียนรู้ของโครงข่ายประสาทเทียมโดยใช้ข้อมูลอินพุตคำนวณหาค่าเอาต์พุต ถ้าค่าเอาต์พุตใกล้เคียงกับเป้าหมายจนเป็นที่ยอมรับก็จะถือว่าการเรียนรู้ประสบความสำเร็จ หากค่าเอาต์พุตยังไม่ใกล้เคียงกับค่าเป้าหมายก็จำทำการปรับค่าถ่วงน้ำหนักแบบย้อนกลับไปยังจุดเริ่มต้นแล้วทำการเรียนรู้ใหม่จนกว่าจะได้ค่าเอาต์พุตเป็นที่น่าพอใจ

- การหาผลลัพธ์ค่าเอาต์พุตจากประสาทเทียม j ในชั้นซ่อน (output of hidden unit j)

$$V_j = \sum_i W_{ji} X_i \quad (2.10)$$

$$O_j = g(\sum_i W_{ji} X_i) \quad (2.11)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หรือ

$$O_j = g(V_j) \quad (2.12)$$

- การหาผลลัพธ์ค่าเอาต์พุตจากประสาทเทียม k ในชั้นเอาต์พุต (hidden-to-output)

$$V_k = \sum_j W_{kj} O_j \quad (2.13)$$

$$O_k = g(\sum_j W_{kj} O_j) \quad (2.14)$$

หรือ

$$O_k = g(V_k) \quad (2.15)$$

โดยมีค่าความผิดพลาดเฉลี่ยของประสาทเทียม $k = \frac{1}{2} e_k^2$ หรือ $E = \frac{1}{2} (t_k - O_k)^2$ เมื่อ t_k เป็นค่าเป้าหมายที่ประสาทเทียม k ดังนั้น ความผิดพลาดโดยรวมในชั้นนี้จึงเป็นดังนี้

$$E[W] = \frac{1}{2} \sum_m (t_m - O_m)^2 \quad (2.16)$$

หรือ

$$E[W_{mj}] = \frac{1}{2} \sum_m (t_m - g(W_{mj} O_j))^2 \quad (2.17)$$

- การปรับค่าถ่วงน้ำหนักแบบย้อนกลับในระดับชั้นเอาต์พุต

การปรับถ่วงน้ำหนักลักษณะนี้มาจากหลักการความสัมพันธ์ของค่าน้ำหนักที่เปลี่ยนแปลงเป็นสัดส่วนโดยตรงกับอนุพันธ์ของความผิดพลาดโดยรวมกับค่าถ่วงน้ำหนักเดิมในลักษณะลดลง (Gradient descent) ดังนั้น ค่าถ่วงน้ำหนักของ W_{kj} ที่เปลี่ยนไปจะมีความสัมพันธ์ ดังนี้

$$\Delta W_{kj} = -\eta \frac{\partial E}{\partial W_{kj}} \quad (2.18)$$

โดย η คืออัตราการเรียนรู้ไว้สำหรับควบคุมการปรับค่าถ่วงน้ำหนัก

จากสมการที่ 2.17 จะได้

$$\frac{\partial E}{\partial W_{kj}} = -(t_k - O_k) g''(\sum_j W_{kj} O_j) O_j \quad (2.19)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และจากสมการที่ 2.9 และสมการที่ 2.13 กับสมการที่ 2.14

$$\therefore \Delta W_{kj} = \eta \delta_k O_j \quad (2.20)$$

โดย

$$\delta_k = (t_k - O_k) O_k (1 - O_k) \quad (2.21)$$

ดังนั้น ค่าถ่วงน้ำหนักใหม่ของ W_{kj} จึงเป็นดังนี้

$$W_{kj}^{new} = W_{kj}^{old} + \Delta W_{kj} \quad (2.22)$$

- การปรับค่าถ่วงน้ำหนักแบบย้อนกลับในระดับชั้นซ่อน

การปรับค่าถ่วงน้ำหนักแบบย้อนกลับในระดับชั้นซ่อนมีลักษณะคล้ายการปรับในระดับชั้นเอาต์พุต การปรับค่าถ่วงน้ำหนักสำหรับในชั้นนี้ของ W_{ji} จะมีความสัมพันธ์ ดังนี้

$$\Delta W_{ji} = -\eta \frac{\partial E}{\partial W_{ji}} \quad (2.23)$$

โดย

$$\frac{\partial E}{\partial W_{ji}} = \frac{\partial E}{\partial O_j} \cdot \frac{\partial O_j}{\partial W_{ji}} \quad (2.24)$$

จากสมการที่ 2.16 และสมการที่ 2.10 กับสมการที่ 2.11 จะได้อนุพันธ์ทั้ง 2 ดังนี้

$$\frac{\partial E}{\partial O_j} = -\sum_m (t_m - O_m) g'(\sum_j W_{mj} O_{mj}) \quad (2.25)$$

และ

$$\frac{\partial E}{\partial O_j} = -g'(\sum_i W_{ji} X_i) X_i \quad (2.26)$$

$$\therefore \Delta W_{ji} = \eta \sum_m (t_m - O_m) O_m (1 - O_m) W_{mj} O_j (1 - O_j) X_i \quad (2.27)$$

หรือ

$$\Delta W_{ji} = \eta \delta_j X_i \quad (2.28)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดย

$$\delta_j = O_j(1 - O_j)X_i \quad (2.29)$$

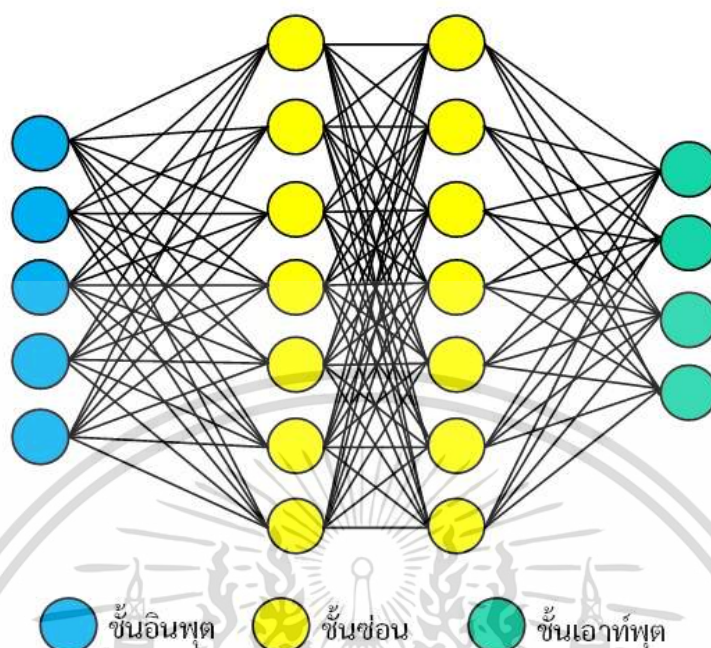
สำหรับการปรับค่าถ่วงน้ำหนักใหม่ของ W_{ji} ก็จะทำเนิการเช่นเดียวกับวิธีที่ผ่านมาในสมการที่ 2.22

2.2.3.2 การเรียนรู้แบบไม่มีครูสอน (Unsupervised learning)

การเรียนรู้แบบไม่มีครูสอนจะมีเพียงแค่ข้อมูลอินพุตไม่จำเป็นต้องมีเป้าหมาย โดยจะมุ่งเน้นการแบ่งกลุ่มชุดข้อมูลโดยข้อมูลนำเข้าที่ให้ผลลัพธ์ที่คล้ายๆกันจะถูกจัดเข้ากลุ่มเดียวกัน บางครั้งจึงเรียกว่าวิธีคลัสเตอร์ (cluster)

2.2.4 หลักการทำงานของโครงข่ายประสาทเทียม

การจำลองระบบโครงข่ายประสาทเทียมมีการจำลองเป็นชั้นๆ โดยโครงสร้างประกอบด้วยชั้นอินพุต (input layer) ชั้นซ่อน (hidden layer) และชั้นเอาต์พุต (output layer) ในแต่ละชั้นประกอบด้วยโหนด (node) ในชั้นซ่อนประกอบด้วยหน่วยที่ทำหน้าที่ส่งผ่านข้อมูลไปสู่ชั้นเอาต์พุต และในชั้นแต่ละชั้นเอาต์พุตประกอบด้วยหน่วยที่ทำหน้าที่ส่งตัวแปรด้านเอาต์พุต (output) ในระหว่างชั้นแต่ละชั้นจะมีการเชื่อมต่อ แต่การเชื่อมต่อจะมีค่าน้ำหนัก (weights) เฉพาะสำหรับทำหน้าที่แทนค่ากำลัง (strength) ของการเชื่อมต่อของเซลล์สมองมนุษย์ โครงสร้างของระบบโครงข่ายประสาทเทียมแสดงได้ดังภาพที่ 2.11



ภาพที่ 2.11 ชั้นโครงสร้างโครงข่ายประสาทเทียม [24]

จากภาพที่ 2.8 และภาพที่ 2.11 สามารถอธิบายส่วนต่างๆของโครงสร้างโครงข่ายประสาทเทียมได้ ดังนี้

1) ประสาทเทียม (neurons)

ในประสาทเทียม (neuron) จะต่างกันตามชั้น (layer) ที่อยู่ ถ้าเป็นอินพุต (input) ช่างในตัวจะมีข้อมูลที่ได้รับมา แต่ถ้าเป็นชั้นซ่อน (hidden Layer) จะมีสมการที่ช่วยในการคำนวณเพื่อทำนายว่าเป็นคลาสอะไรหรือคำนวณแบบถดถอย (regression) ก็ได้ แต่ถ้าเป็นเอาต์พุต (output) จะเป็นตัวที่บ่งบอกว่า เป็นคลาสอะไรนั่นเอง

2) ชั้นอินพุต (input layer)

ทำหน้าที่ในการรับข้อมูลเข้ามาในโครงข่ายประสาทโดยชั้นอินพุต (input Layer) จะมีเพียงชั้นเดียวและมีหน้าที่ส่งข้อมูลไปยังชั้นถัดไปหรือชั้นซ่อน (hidden Layer)

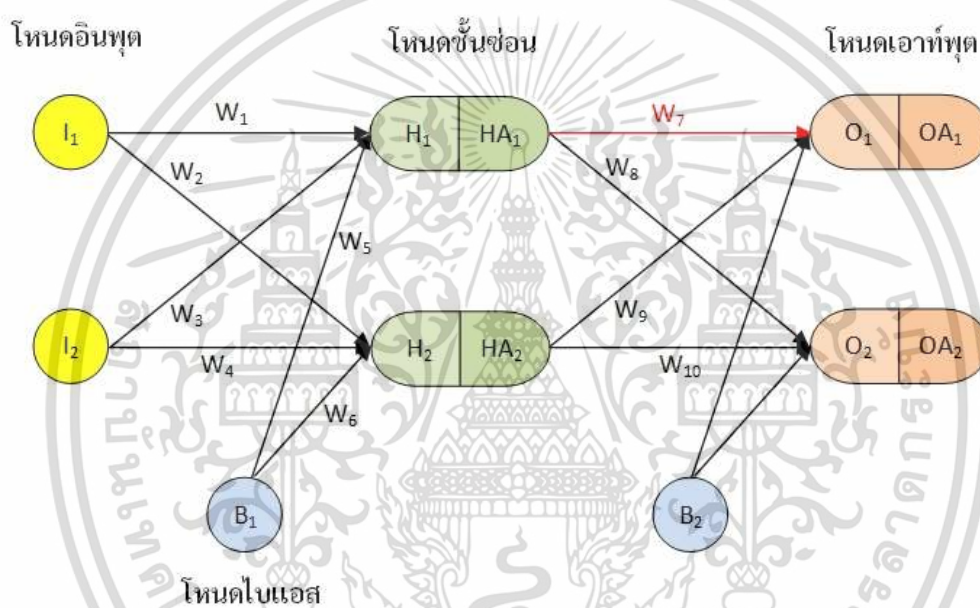
3) ชั้นซ่อน (hidden Layer)

ทำหน้าที่รับข้อมูลจากชั้น (layer) ก่อนหน้า ชั้นซ่อน (hidden Layer) มีจำนวนมากกว่า 1 ชั้นได้และหากต้องการความแม่นยำที่มากขึ้นสามารถเพิ่มจำนวนชั้นของชั้นซ่อน (hidden Layer) และจำนวนประสาทเทียม (Neurons) ให้มากขึ้นก็จะช่วยได้แต่ต้องเพิ่มให้เหมาะสม การเพิ่มชั้นมากเกินไปทำให้เกิดความหนาแน่นของข้อมูล

4) ชั้นเอาต์พุต (output layer)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชั้นเอาต์พุต (output layer) จะอยู่ท้ายสุดและรอรับค่าจากชั้นซ่อน (hidden layer) สุดท้าย ในชั้นเอาต์พุต (output) นั้นแต่ละโหนด (node) ของโครงข่ายประสาทเทียมจะมีค่าน้ำหนักของชนิดหรือคลาส (class) อยู่ เช่น หากมีประเภทของเอาต์พุต (output) ทั้งหมด 2 แบบ เช่น แมวกับสุนัข ฉะนั้นชั้นเอาต์พุต (output layer) จะมีประสาทเทียม (neurons) 2 ตัว ตัวแรกจะเป็นสุนัข ตัวที่สองจะเป็นแมว โดยเมื่อข้อมูลผ่านชั้นซ่อน (hidden layer) ไปสู่เอาต์พุต (output) แล้วประสาทเทียม (neurons) ทั้ง 2 ตัวจะมีค่าข้างในไม่เท่ากัน โดยประสาทเทียม (neuron) ตัวไหนมีน้ำหนักรวมกว่าแสดงว่าเป็นชนิดนั้นและในชั้นซ่อน (hidden layer) มีอีกสองสิ่งที่ถูกซ่อนไว้คือ ค่าไบแอส (bias) และค่าน้ำหนัก (weight)



ภาพที่ 2.12 ไบแอส (bias) และน้ำหนัก (weight) ที่อยู่ในชั้นโครงข่ายประสาทเทียม [25]

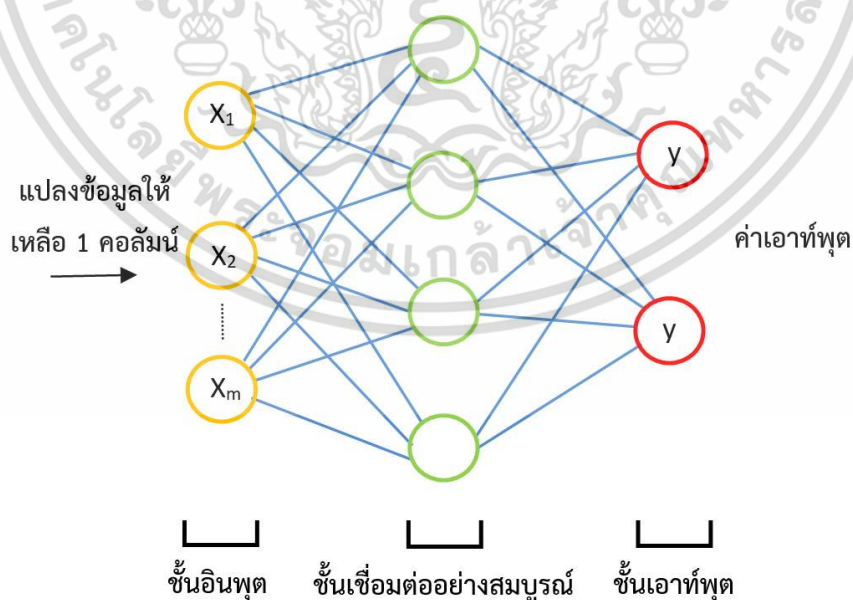
ทุกๆ ชั้นซ่อน (hidden layer) จะมีไบแอส (bias) เชื่อมต่ออยู่เพื่อให้ทุกๆ การคำนวณมีค่ายกระดับ ทุกๆ โครงข่ายในชั้นซ่อน (hidden layer) ต้องมีไบแอส (bias) เข้าไปคำนวณเพื่อให้ขอบเขตการตัดสินใจ (decision boundary) ไม่จำเป็นต้องผ่านจุดกำเนิด (origin) หรือเป็นการยกระดับค่าได้

ค่าน้ำหนัก (weight) จะส่งผลทุกๆ โครงข่ายที่มีค่าเอาต์พุต (output) ที่ไม่เท่ากันทำให้แต่ละชนิดมีน้ำหนักรวมไม่เท่ากันเวลาคำนวณว่าเป็นชนิดไหน ทำให้สามารถแยกได้ว่าข้อมูลนี้เป็นชนิดอะไรได้ด้วยการพิจารณาตัวเลขที่เอาต์พุต (output) ของการทำงานของโครงข่ายประสาทเทียม คือ เมื่อมีอินพุต (input) เข้ามายังโครงข่ายประสาทเทียมก็จะเอาอินพุต (input) มาคูณกับค่าน้ำหนัก (weight) ของแต่ละส่วน ผลที่ได้จากอินพุต (input) ทุกๆ ช่องของโครงข่ายจะเอามารวมกันแล้วก็เอามาเทียบกับค่าที่กำหนด (threshold) ไว้ ถ้าผลรวมมีค่ามากกว่าค่าที่

กำหนดแล้วโครงข่ายประสาทเทียมก็จะส่งเอาต์พุต (output) ออกไปโดยเอาต์พุต (output) นี้จะถูกส่งไปยังอินพุต (input) ของโครงข่ายอื่นๆที่เชื่อมกันในโครงข่าย ถ้าค่าน้อยกว่าเกณฑ์ที่กำหนดก็จะไม่เกิดค่าเอาต์พุต (output) นั้นๆ

สิ่งสำคัญคือต้องทราบค่าน้ำหนัก (weight) และค่าที่กำหนด (threshold) สำหรับสิ่งที่ต้องการเพื่อให้คอมพิวเตอร์รู้จักซึ่งเป็นค่าที่ไม่แน่นอนแต่สามารถกำหนดให้คอมพิวเตอร์ปรับค่าเหล่านั้นได้ โดยการสอนให้มันรู้จักแพทเทริน (pattern) ของสิ่งที่เราต้องการให้มันเรียนรู้ เรียกว่า “การแพร่ย้อนกลับ” หรือ “back propagation” นั่นเอง ในการฝึกโครงข่ายประสาทเทียมแบบป้อนไปข้างหน้า (Feed-forward neural network) จะมีการใช้วิธีการแบบการฝึกฝนแบบกระบวนการส่งค่าย้อนกลับ (back propagation) เพื่อใช้ในการปรับปรุงค่าน้ำหนักของเครือข่าย (network weight) หลังจากใส่รูปแบบข้อมูลสำหรับฝึกให้เครือข่ายในแต่ละครั้งแล้วค่าที่ได้รับหรือเอาต์พุต (output) จากเครือข่ายจะถูกนำไปเปรียบเทียบกับผลที่คาดหวังแล้วทำการคำนวณหาค่าความผิดพลาดซึ่งค่าความผิดพลาดนี้จะถูกส่งกลับเข้าสู่เครือข่ายเพื่อใช้แก้ไขค่าน้ำหนักคะแนนต่อไป ซึ่งเป็นการสอนโครงข่ายให้มีความผิดพลาดตามที่ต้องการ

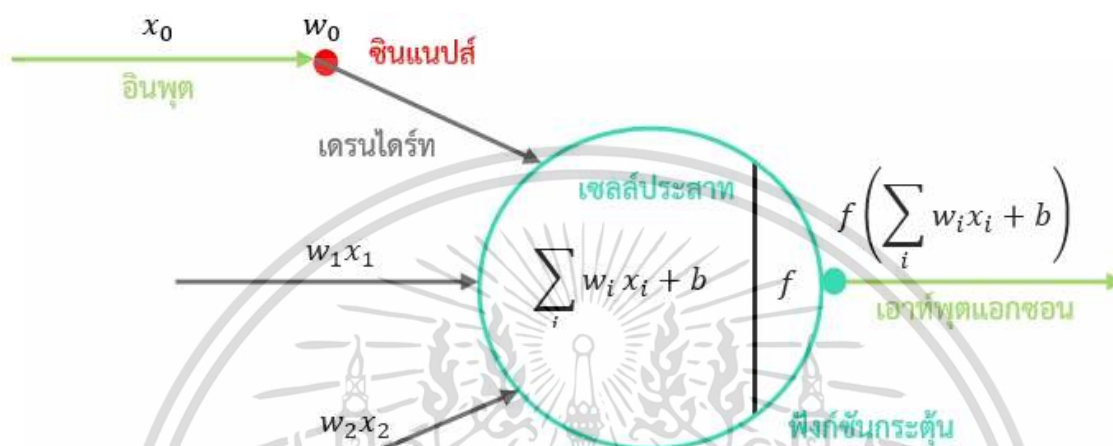
งานวิจัยนี้จะมุ่งเน้นไปที่การใช้ระบบโครงข่ายประสาทเทียมเพื่อใช้ในการแยกประเภทของรูปภาพ (classification) สำหรับการจำแนกบริเวณที่สนใจของล้อและรางรถไฟฟ้าโดยโครงข่ายประสาทเทียมที่สร้างขึ้นจะมีลักษณะของการเชื่อมต่อถึงกันทุกเส้นทางการเดินทางของเซลล์ประสาท โครงข่ายประสาทเทียมประเภทนี้จะถูกเรียกว่าโครงข่ายประสาทเทียมแบบเชื่อมต่อกันอย่างสมบูรณ์ (fully connected neural network) ดังภาพที่ 2.13



ภาพที่ 2.13 โครงข่ายประสาทเทียมแบบเชื่อมต่อกันอย่างสมบูรณ์ [26]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โครงข่ายประสาทเทียมแบบเชื่อมต่อกันอย่างสมบูรณ์ประกอบไปด้วยส่วนประกอบหลัก คือ นิวรอนหรือเซลล์ประสาทและส่วนของทางเดินประสาท ตามภาพที่ 2.14



ภาพที่ 2.14 ส่วนประกอบของเซลล์ประสาทและทางเดินประสาท [27]

ส่วนของเซลล์ประสาทจะมีความสามารถในการรวบรวมอินพุตทั้งหมดที่เข้ามาและส่งสัญญาณออกไปโดยผ่านฟังก์ชันกระตุ้น (activation function) เพื่อไปยังเซลล์ประสาทชั้นต่อไป ซึ่งขั้นตอนทั้งหมดสามารถเขียนแทนด้วยสมการที่ (2.30)

$$(\sum x_i w_i + b) \quad (2.30)$$

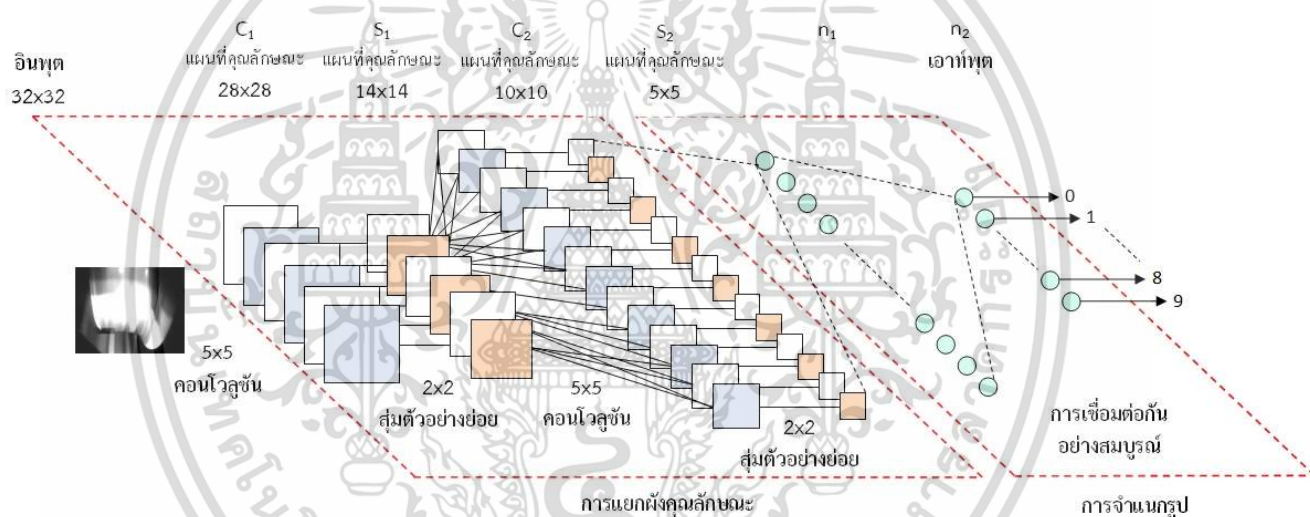
เมื่อ f คือฟังก์ชันกระตุ้น (activation function) x เป็นอินพุต w เป็นน้ำหนัก และ b เป็นค่าไบแอสที่กำหนดขึ้น

นอกจากนี้ยังมีวิธีการ (algorithm) ที่น่าสนใจที่มีโครงข่ายประสาทเทียมเป็นส่วนประกอบ โดยโครงสร้างของโครงข่ายประสาทเทียมที่งานวิจัยนี้ใช้คือโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution Neural Network: CNN) ซึ่งเป็นโครงสร้างที่เน้นใช้กับรูปภาพโดยจะดึงจุดเด่นของภาพนั้นๆออกมาได้เพื่อนำมาจำแนก (classification) ว่าเป็นชนิดอะไร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.5 โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution Neural Network: CNN)

ระบบโครงข่ายประสาทเทียมแบบคอนโวลูชันมีหลักการการทำงานผสมกันระหว่างตัวกรองแบบคอนโวลูชันทำงานร่วมกับระบบโครงข่ายประสาทที่สามารถแก้ไขค่าน้ำหนักได้หากคำตอบไม่เป็นไปตามที่ต้องการ โดยเป้าหมายของการใช้งานโครงข่ายประสาทเทียมแบบคอนโวลูชันคือการปรับแก้ไขค่าต่างๆในมาส์กของตัวกรองคอนโวลูชันส่งผลให้สามารถปรับปรุงตัวกรองเพื่อเลือกคุณลักษณะเฉพาะที่สอดคล้องกับอินพุตและเอาต์พุตที่ต้องการ โดยหลักการและทฤษฎีที่ใช้อ้างอิงจะแบ่งออกเป็น 2 ส่วนด้วยกัน คือ การทำงานของโครงข่ายประสาทเทียมและการทำงานของชั้นที่ใช้ในโครงข่ายประสาทเทียมแบบคอนโวลูชัน ดังภาพที่ 2.15

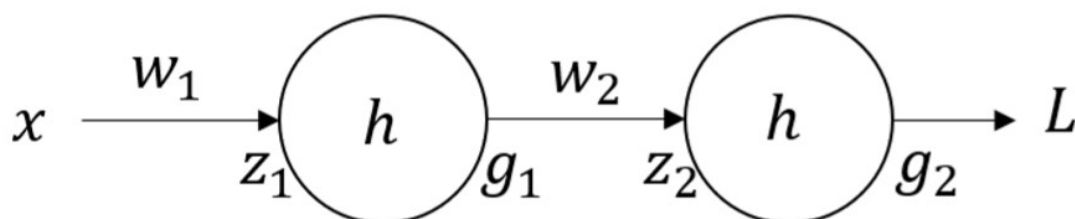


ภาพที่ 2.15 โครงข่ายประสาทเทียมแบบคอนโวลูชัน [28]

2.2.6 โครงข่ายประสาทเทียมแบบแพร่ไปข้างหน้า (Feed forward neural network)

การแพร่ไปข้างหน้าของโครงข่ายประสาทเทียมมีการทำงานโดยมีทิศทางการคำนวณจากอินพุต (input) ไปสู่เอาต์พุต (output) อย่างเดียวเท่านั้นโดยไม่มีการสลับหรือย้อนกลับของลำดับชั้นตอน เมื่อนำโครงสร้างมาพิจารณาจะได้ทิศทางคำนวณโดยสังเขปดังภาพที่ 2.16

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 2.16 ทิศทางการคำนวณของโครงข่ายประสาทเทียม

สามารถนำมาเขียนเป็นสมการสำหรับการดำเนินการไปข้างหน้า (forward pass) ได้ดังนี้

$$z_1 = xw_1 \quad (2.31)$$

$$g_1 = h(z_1) \quad (2.32)$$

$$z_2 = g_1w_2 \quad (2.33)$$

$$g_2 = h(z_2) \quad (2.34)$$

ให้ x เป็นอินพุต w_1, w_2 เป็นค่าน้ำหนักแต่ละชั้นตามลำดับ h เป็นฟังก์ชันกระตุ้น (activation function) z_1, z_2, g_1, g_2 เป็นกระบวนการคำนวณตามลำดับ และ L เป็นค่าความผิดพลาดของโมเดล

2.2.7 การปรับค่าน้ำหนักของโครงข่ายประสาทเทียม (Weight modification)

การปรับค่าน้ำหนักจะใช้วิธีการปรับตามการถดถอยของความลาดชันเกรเดียนต์ (Gradient descent) ซึ่งสามารถคำนวณได้จากการหาอนุพันธ์ของค่าความผิดพลาดเทียบกับน้ำหนักในตำแหน่งของเซลล์ประสาทนั้น เมื่อได้ค่าของเกรเดียนต์จากค่าความผิดพลาดเทียบกับน้ำหนักตำแหน่งใดๆแล้วการปรับปรุงค่าน้ำหนักจะเป็นไปตามสมการที่ (2.35)

$$w_{new} = w_{old} - \mu \Delta L \Delta w_{old} \quad (2.35)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อ w_{new} คือค่าน้ำหนักใหม่ที่ปรับแล้ว w_{old} คือค่าน้ำหนักเดิมในชั้นที่ i μ คืออัตราการเรียนรู้ที่มีค่า $1 > \mu > 0$ และ $L \partial w_{old}$ คือความลาดชันเกรเดียนต์ของค่าความผิดพลาดเทียบกับน้ำหนักในชั้นที่ i โดยความลาดชันเกรเดียนต์ของชั้นที่ i จะคำนวณจากการหาอนุพันธ์แบบลูกโซ่ได้ ดังสมการ (2.36) และ (2.37)

$$\partial L \partial w_2 = \partial L \partial g_2 \times \partial g_2 \partial z_2 \times \partial z_2 \partial w_2 \quad (2.36)$$

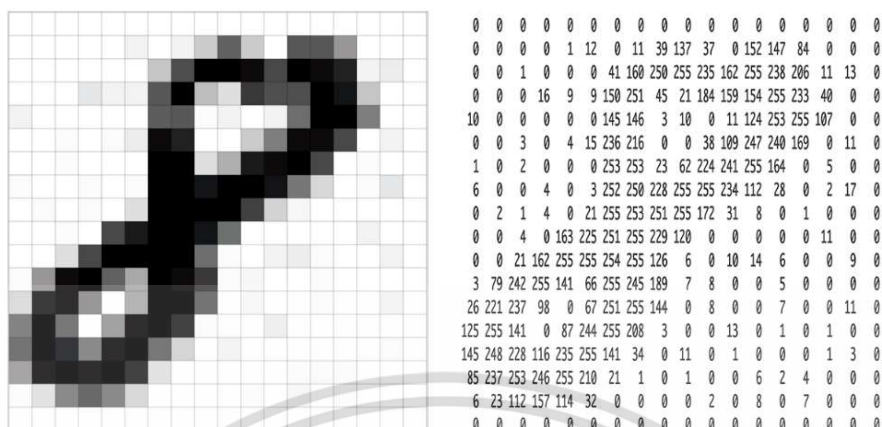
$$\partial L \partial w_1 = \partial L \partial g_2 \times \partial g_2 \partial z_2 \times \partial z_2 \partial g_1 \times \partial g_1 \partial z_1 \times \partial z_1 \partial w_1 \quad (2.37)$$

2.2.8 การทำงานของโครงข่ายประสาทแบบคอนโวลูชัน (Convolutional neural network)

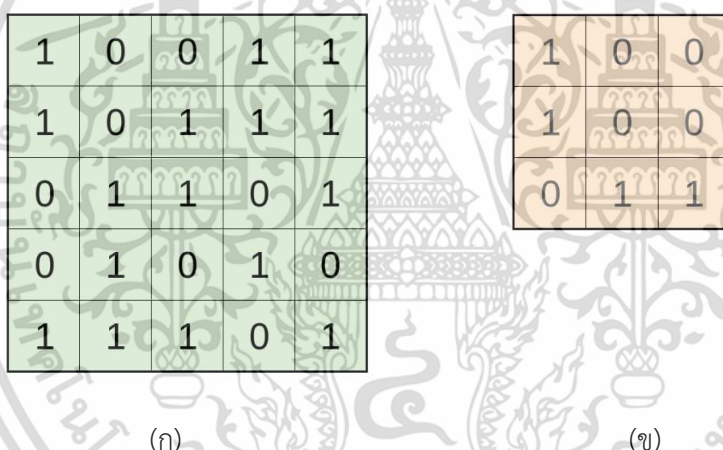
เมื่อกำหนดวิธีการปรับค่าน้ำหนักได้แล้วจะสามารถนำมาใช้งานร่วมกับการปรับตัวแปรในตัวกรองของคอนโวลูชันได้ โดยการสร้างโครงข่ายประสาทเทียมแบบคอนโวลูชันนั้นจะมีส่วนประกอบชั้นทั้งหมด 4 ส่วน โดยแบ่งเป็นชั้นคอนโวลูชัน (Convolutional layer) ชั้นแมกส์พูลลิง (Max pooling layer) ชั้นเรลู (Relu layer) และชั้นเชื่อมต่อกันอย่างสมบูรณ์ (fully connected layer)

2.2.8.1 ชั้นคอนโวลูชัน (Convolutional layer)

คอนโวลูชันเป็นกระบวนการที่ทำเพื่อสกัดเอาลักษณะสำคัญจากภาพออกมาโดยใช้ค่าพิกเซล (pixel) โดยค่าพิกเซลได้มาจากการมองของกล้องทั่วไปมีด้วยกัน 3 แชนแนล (channel) แบ่งเป็นสี ได้แก่ สีแดง สีน้ำเงิน และสีเขียว โดยแต่ละจุดสามารถแทนค่าด้วยตัวเลขเพื่อบอกความเข้มของสีนั้นๆ โดยมีค่าตั้งแต่ 0 ถึง 255 จากความเข้มน้อยไปหาความเข้มมาก ในการทำภาพขาวดำแชนแนลของภาพนั้นจะมีเพียงหนึ่งแชนแนลเท่านั้น คือ แชนแนลของสีดำ ซึ่งค่าตัวเลข 0 นั้นคือสีขาว ไล่ไปจนถึง 255 ซึ่งเป็นสีดำสนิทในแต่ละจุดจะทำการคำนวณแล้วเก็บไว้



ภาพที่ 2.17 ตัวอย่างการแทนค่าพิกเซลลงบนภาพที่รับเข้ามา [29]



ภาพที่ 2.18 จำลองเมทริกซ์ (ก) เมทริกซ์ที่ได้จากภาพที่รับเข้า (ข) เมทริกซ์ตัวกรองค่า [30]

ในแต่ละรูปจะมีชุดค่าเมทริกซ์ที่แตกต่างกัน จากภาพที่ 2.19 (ก) เป็นรูปภาพขนาด 5x5 พิกเซล ที่เป็นภาพขาวดำ โดยกำหนดให้ค่า 0 พิกเซลเป็นสีขาวและ 1 พิกเซลเป็นสีดำและมีการกำหนดเมทริกซ์อีกหนึ่งชุดขึ้นมา โดยใช้เมทริกซ์ชุดนี้เป็นตัวกรองค่าไปเก็บไว้ในเมทริกซ์ชุดที่เล็กกว่าและเรียกเมทริกซ์ชุดนี้ว่าตัวกรองค่า (filter) เคอเนล (kernel) หรือตัวตรวจจับลักษณะสำคัญ (feature detector) ดังที่ปรากฏในภาพที่ 2.16 (ข)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

| | | | | |
|-----|-----|-----|---|---|
| 1x1 | 1x0 | 1x1 | 0 | 0 |
| 0x0 | 1x1 | 1x0 | 1 | 0 |
| 0x1 | 0x0 | 1x1 | 1 | 1 |
| 0 | 0 | 1 | 1 | 0 |
| 0 | 1 | 1 | 0 | 0 |

อินพุต x ตัวกรอง

| | | |
|---|--|--|
| 4 | | |
| | | |
| | | |

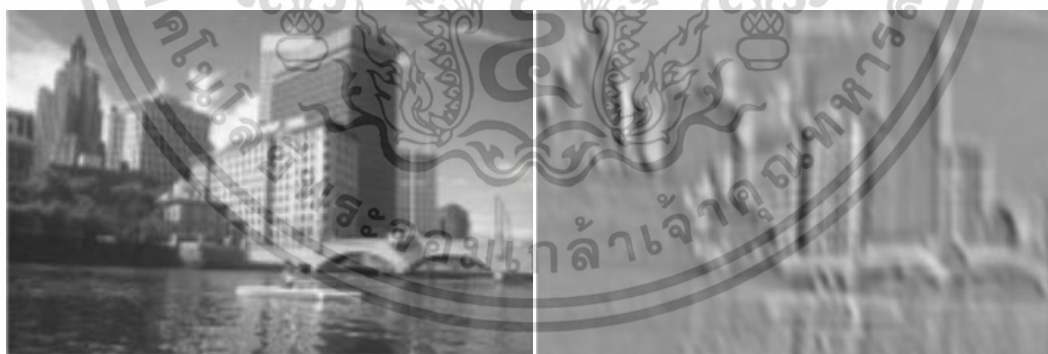
ฟังก์ชันลักษณะ

(ก)

(ข)

ภาพที่ 2.19 การทำงานของตัวกรองค่าและการกำเนิดเมทริกซ์ชุดใหม่หรือฟังก์ชันลักษณะ [31]

เมทริกซ์ที่ทำหน้าที่เป็นตัวกรองค่าจะเคลื่อนที่ไปทั่วภาพและคุณค่าเก็บไว้ในเมทริกซ์ชุดใหม่ดังภาพที่ 2.20 (ข) โดยผลลัพธ์ที่ได้เมื่อรูปผ่านการทำคอนโวลูชัน แสดงตัวอย่างดังภาพที่ 2.20

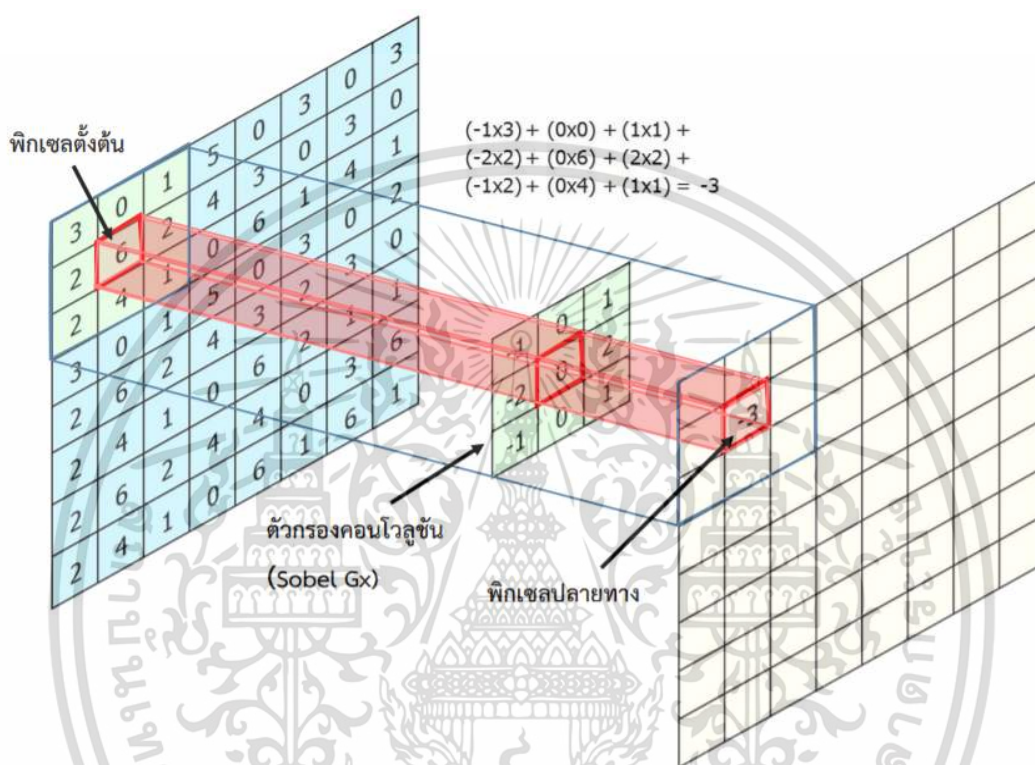


ภาพที่ 2.20 ภาพขาวดำดั้งเดิมเมื่อผ่านการทำคอนโวลูชันกลายเป็นฟังก์ชันลักษณะ (feature map) [32]

ชั้นคอนโวลูชัน (convolutional layer) คือ ชั้นที่ทำหน้าที่เป็นตัวกรองคุณลักษณะต่างๆด้วยวิธีการทางคอนโวลูชันคล้ายกับการประมวลผลภาพขั้นต้น โดยสมาชิกแต่ละตัวจะถูกแทนด้วยค่าน้ำหนัก w ที่สามารถปรับได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เพื่อให้การคัดเลือกคุณลักษณะสามารถปรับตัวตามอินพุตที่ป้อนเข้าไปยังโครงข่ายประสาท โดยหลักการทำงานจะอธิบายได้ดังภาพที่ 2.21



ภาพที่ 2.21 การทำงานของชั้นคอนโวลูชัน [33]

2.2.8.2 ชั้นแมกซ์พูลลิง (Max pooling layer)

การพูลลิง (pooling) ช่วยลดมิติของฟังก์ชันคุณลักษณะ (feature map) ลงแต่ยังคงรักษาข้อมูลสำคัญไว้ การพูลลิง (pooling) สามารถจำแนกเป็นประเภทต่างๆได้ เช่น พูลลิงด้วยค่าสูงสุด (max pooling) ค่าเฉลี่ย (average pooling) ผลรวมหรืออื่นๆ การพูลลิงทำให้ผลลัพธ์ที่ได้มีขนาดเล็กและจัดการได้ง่ายขึ้น นอกจากนี้ยังลดจำนวนพารามิเตอร์และการคำนวณที่เกินจำเป็นในโครงข่าย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชั้นแมกซ์พูลลิง (max pooling layer) คือชั้นที่ใช้สำหรับลดจำนวนตัวเลขที่ซ้ำกันสำหรับรักษาเลขน้อยๆ สำคัญเพื่อป้องกันไม่ให้เกิดข้อมูลที่ผ่านตัวกรองมีเยอะขึ้นเกินความจำเป็น มีหลักการทำงานคือเลือกค่าที่มากที่สุด เมื่อนำหน้าต่างตัวกรองไปเทียบกับภาพต้นแบบ ตัวอย่างในที่นี้สมมติให้หน้าต่างมีขนาด 2x2 และหน้าต่างนี้จะทำการเคลื่อนที่ทีละ 2 พิกเซลไปจนทั่วเมทริกซ์ของผังคุณลักษณะ (feature map) เพื่อทำการเก็บค่าที่สูงที่สุดในทุกๆ 2 พิกเซล ดังภาพที่ 2.22

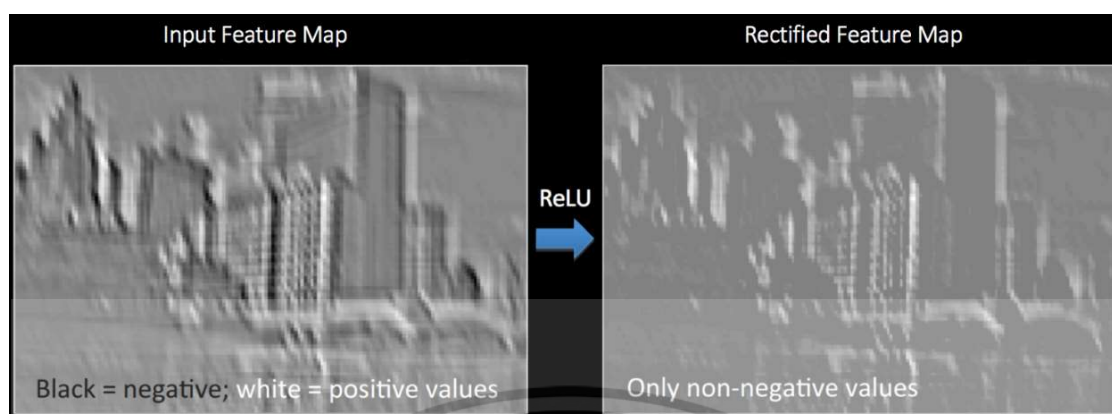


ภาพที่ 2.22 การทำงานของแมกซ์พูลลิง [33]

2.2.8.3 การขจัดความเป็นเชิงเส้น (ReLU)

หลังจากได้ผังคุณลักษณะมาแล้วจะทำการปรับแต่งให้ผังคุณลักษณะ (feature map) ไม่เป็นเชิงเส้นด้วยวิธีการเรลู (ReLU) โดยทำการแทนที่ผลของพิกเซลที่มีค่าเป็นเชิงลบในผังคุณลักษณะ (feature map) ด้วยค่า 0 จุดประสงค์ของการทำเรลู (ReLU) เพื่อให้โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network) เรียนรู้ข้อมูลที่ไม่เป็นเชิงเส้นจากภาพ เมื่อนำภาพเข้าสู่กระบวนการที่เป็นผังคุณลักษณะ (feature map) เข้ามาทำเรลู (ReLU) โดยสีดำในภาพเป็นค่าเชิงลบและสีขาวในภาพเป็นค่าเชิงบวก เมื่อทำการเรลู (ReLU) ค่าที่ได้จะมีเพียงค่าที่เป็นเชิงบวกเท่านั้นดังภาพที่ 2.23

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

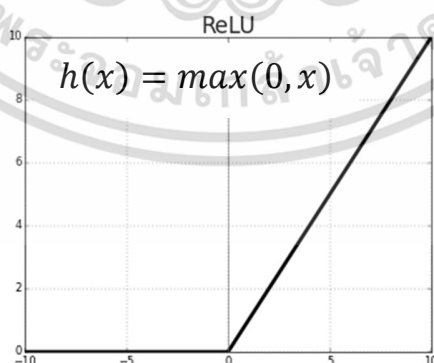


ภาพที่ 2.23 ตัวอย่างผลลัพธ์หลังการทำ ReLU

ชั้นเรลู (Relu layer) ในการใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network) นั้น แอคติเวชันฟังก์ชัน h ที่เหมาะสมคือ เรลูฟังก์ชัน (Rectifier linear unit: Relu) เนื่องจากตัวสมการมีความง่ายในการคำนวณสำหรับชุดคำสั่งของคอมพิวเตอร์และยังมีประโยชน์คล้ายกับการทำเทรชโฮลด์ (threshold) อีกด้วย เรลูฟังก์ชันจึงถูกเลือกมาเป็นส่วนประกอบสำหรับการทำหน้าที่ในส่วนของแอคติเวชันฟังก์ชัน โดยมีการทำงานดังสมการที่ (2.38)

$$h(x) = \max(0, x) = \begin{cases} 0 & \text{for } x \leq 0 \\ x & \text{for } x > 0 \end{cases} \quad (2.38)$$

เมื่อ $h(x)$ คือฟังก์ชันกระตุ้น (activation function), x คือค่าของแต่ละสมาชิกอินพุต

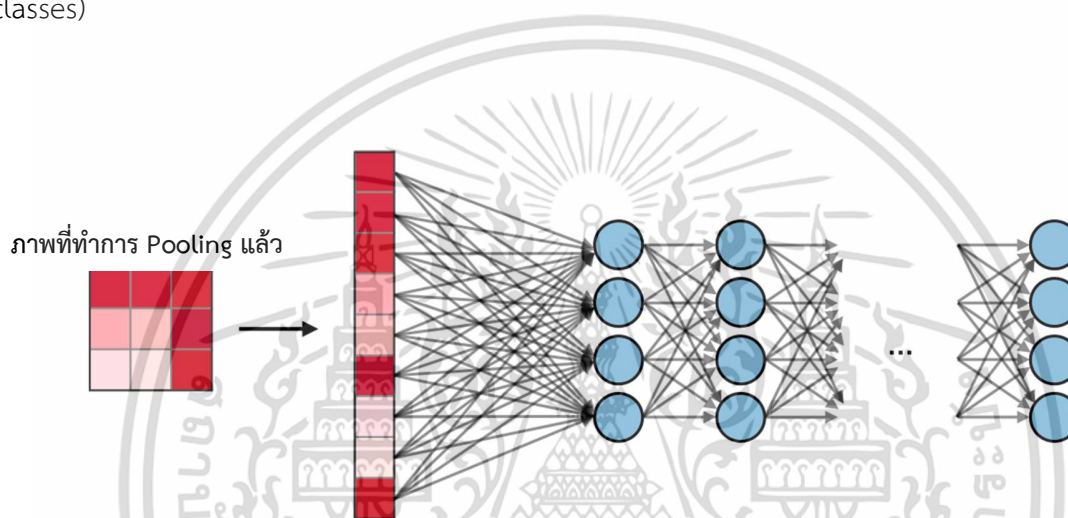


ภาพที่ 2.24 กราฟฟังก์ชัน ReLU

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.8.4 การเชื่อมต่อกันของแต่ละชั้นอย่างสมบูรณ์ (Fully connected layer)

กระบวนการคอนโวลูชัน (convolution) พูลลิง (pooling) และเรลยู (ReLU) กระบวนการทั้งสามต้องทำซ้ำจนกว่าจะมีการเชื่อมต่อกันของแต่ละชั้นอย่างสมบูรณ์ ผลลัพธ์จากคอนโวลูชัน (convolution) และพูลลิง (pooling) นั้นให้ลักษณะเด่นในระดับสูง (high-level feature) ของภาพที่รับเข้ามา จุดประสงค์ของการทำให้เชื่อมต่อกันแต่ละชั้นโดยสมบูรณ์นั้นเพื่อนำลักษณะเด่นไปทำการคัดกรองภาพที่รับเข้ามาให้อยู่ในรูปของคลาส (classes)



ภาพที่ 2.25 การเชื่อมต่อกันของแต่ละชั้นอย่างสมบูรณ์ (Fully connected layer) [34]

ผลลัพธ์ที่ได้จะแสดงค่าความมั่นใจ (confident) ออกมาจากการใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network) ผลลัพธ์ที่ได้คือค่าความมั่นใจและตำแหน่งของวัตถุที่อยู่ในภาพที่ป้อนเข้าสู่ระบบโดยข้อมูลเหล่านี้จะถูกนำไปวิเคราะห์ต่อไป

การเชื่อมต่อกันของแต่ละชั้นอย่างสมบูรณ์ (Fully connected layer) คือชั้นที่ช่วยรวมผลการคำนวณของชั้นขั้นต้นเพื่อนำไปจัดรูปแบบให้เหมือนกับโครงข่ายประสาทแบบเชื่อมต่อกันอย่างสมบูรณ์ (Fully connected) และในขั้นตอนการจำแนกประเภทของรูปจะใช้สมการซอฟแมกซ์-ครอสเอนโทรปี (Softmax-Cross entropy) ซึ่งจะถูกใช้แทนในส่วนของสมการค่าความผิดพลาดดังสมการที่ (2.39)

$$L = -\log(eSk \sum eSj j) \quad (2.39)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อ L คือ ค่าความผิดพลาดของเอาต์พุต, S_k คือเอาต์พุตของประเภทนั้น, S_j คือ ค่าเอาต์พุตของคำตอบประเภทอื่น

จะเห็นได้ว่าโครงข่ายประสาทเทียมเหมาะกับการจำแนกประเภทรูปภาพโดยการหาคุณลักษณะของรูปภาพจากการปรับเปลี่ยนตามอินพุตและเอาต์พุตทำให้ได้คุณลักษณะที่ครบถ้วนกว่าการเลือกคุณลักษณะเอง โดยบางคุณลักษณะเป็นสิ่งที่มนุษย์ไม่สามารถอธิบายออกมาเป็นสมการหรือเขียนโปรแกรมให้คอมพิวเตอร์คัดแยกออกมาได้ ดังนั้น วิทยานิพนธ์นี้จึงเลือกวิธีใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันมาช่วยวัดระยะห่างระหว่างล้อและรางรถไฟ

2.3 โครงข่ายประสาทเทียมแบบคอนโวลูชันสำหรับการตรวจจับวัตถุ (Convolutional Neural Network model for object detection)

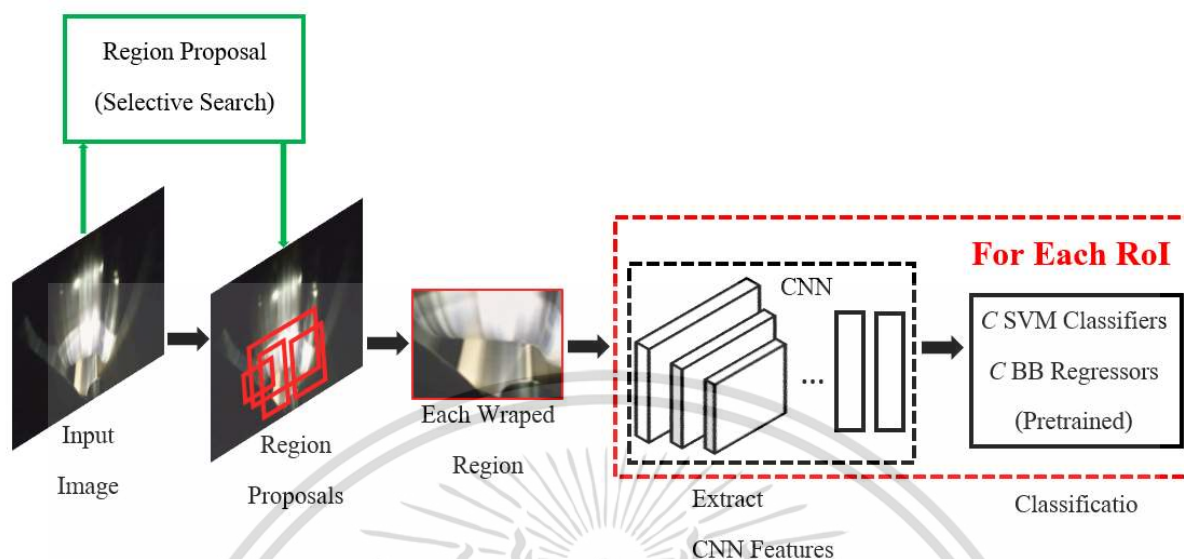
การทำงานที่กล่าวมานั้นเป็นพื้นฐานของโครงข่ายประสาทแบบคอนโวลูชัน (Convolutional Neural Network) ซึ่งมีโครงข่ายแบบจำลอง แต่ละตัวก็จะมีชั้นการทำงานต่างๆ ในจำนวนที่ต่างกันมีวิธีการเฉพาะของโครงข่ายนั้นๆ

2.3.1 R-CNN (Regions with Convolutional Neural Network feature) [35]

ในปี 2014 R. Girshick, J. Donahue, T. Darrell, and J. Malik ได้นำเสนอวิธี R-CNN (Region with CNN feature) ที่เอา CNN มาต่อยอดช่วยงานด้านการตรวจจับวัตถุ (object detection) แต่ปัญหาคืองานด้านการตรวจจับคือสิ่งที่เราต้องการหาบางที่อาจไม่มีปรากฏในภาพเลยก็ได้หรืออาจปรากฏขึ้นมาพร้อมกันทีเดียวหลายอัน แต่ละอันอยู่คนละมุมในภาพและมีขนาดเล็กบ้างใหญ่บ้างต่างๆกันไป ลักษณะนี้การจำแนกรูปภาพ (image classifier) อย่างเดียวไม่สามารถแก้ปัญหาได้ วิธีการ R-CNN จึงเพิ่มส่วนการทำ region proposals ที่ทำหน้าที่เสนอ RoI (Region of interest) เป็นกรอบสี่เหลี่ยมมุมฉากหลายชิ้นหลายตำแหน่งขึ้นมาจำนวน 2,000 อันต่อ 1 ภาพอินพุต จากนั้นภาพอินพุตต้นฉบับจะถูกตัด (crop) ด้วย RoI ก่อนจะถูกส่งต่อไปให้ CNN ทำ การแยกคุณลักษณะ (feature extraction) และ SVM (Support vector machine) ทำการจำแนก (classification)

วิธีการนี้ไม่ว่าภาพหนึ่งจะมีวัตถุ (object) อยู่กี่คลาส ไม่ว่าแต่ละวัตถุจะอยู่ตำแหน่งไหนของภาพหรือมีขนาดเท่าใดก็ตาม R-CNN ก็ยังสามารถตามไปสร้าง RoI ครอบและตัด (crop) ส่วนของวัตถุนั้นๆ เพื่อส่งภาพที่ crop มาให้ CNN และ SVM ช่วยกัน classifier ต่อได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 2.26 โครงสร้างแบบจำลอง R-CNN

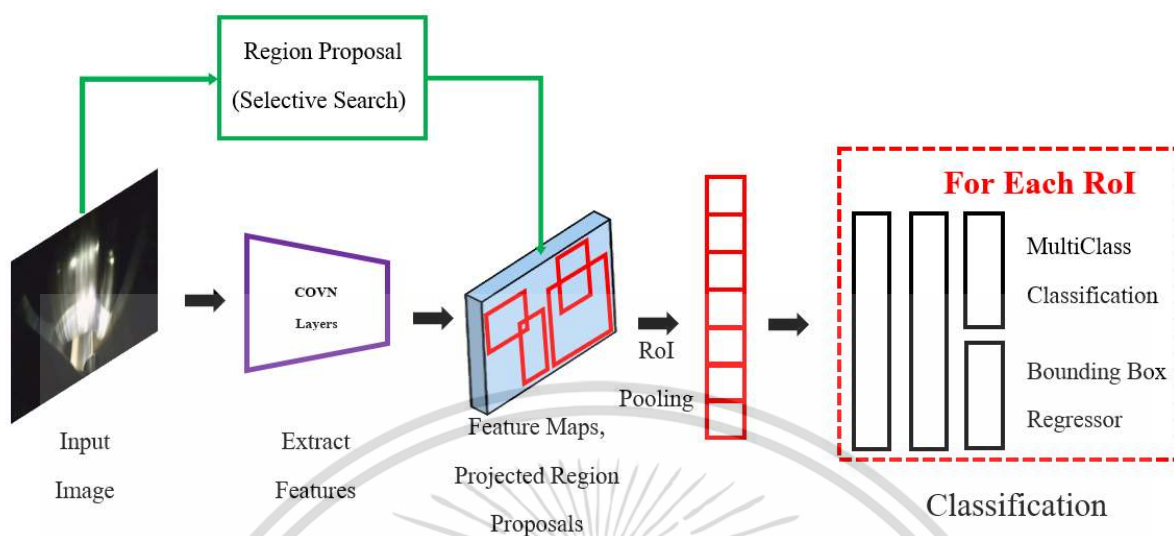
ข้อเสียของ R-CNN คือ

- 1) จำนวน RoI เยอะเกินไป (1 ภาพ มี 2,000 RoI)
- 2) ส่วน CNN feature extraction ของแต่ละ RoI ทำงานซ้ำซ้อนกัน
- 3) มีแบบจำลองแยกกันอยู่หลายส่วน ไม่สามารถทำ end-to-end training รวดเดียวได้

2.3.2 Fast R-CNN (Fast Recurrent Convolutional Neural Network) [36]

ในปี 2015 R. Girshick [36] ได้นำเสนอ Fast R-CNN (Fast region-based convolutional neural network) เพื่อแก้ปัญหา R-CNN โดยการรวมทุกส่วนให้เป็น 1 เน็ตเวิร์ค (ทำ End-to-end training ได้แล้ว) คือทำงานโดยนำรูปภาพทั้งรูปมาเป็นข้อมูลนำเข้าและประมวลผลทั้งรูปภาพโดยการทำคอนโวลูชันและพูลลิงที่ค่าสูงสุดหลายครั้ง ในการสร้างฟังก์ชันคุณลักษณะและใช้เทคนิคที่ชื่อว่า RoI pooling (Region of interest pooling) ระหว่างการทำงานของกรนำเสนอพื้นที่ที่สนใจ (region proposal) และได้ทำการสร้างแบบจำลองของตัวจำแนกหมวดหมู่ของโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional neural network classifier และ Bounding box regressor บนโมเดลเดียวกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

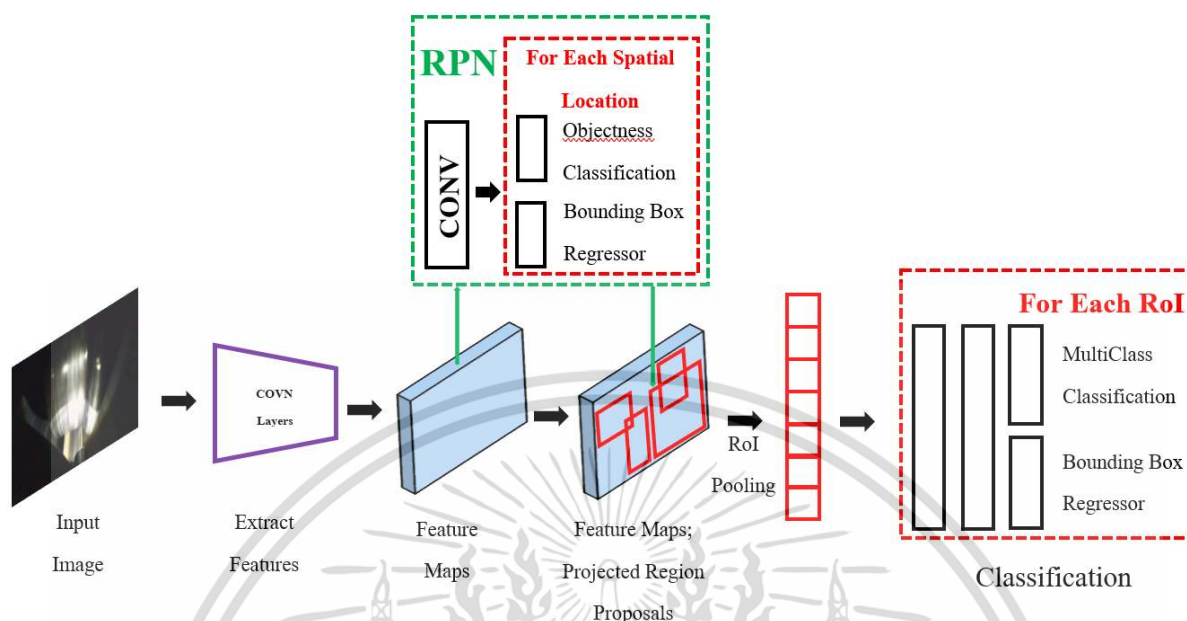


ภาพที่ 2.27 โครงสร้างแบบจำลอง Fast R-CNN

จาก [34] พบว่า Fast R-CNN ใช้เวลาทำงานต่อภาพแค่ 2.3 วินาที ในขณะที่ R-CNN ใช้เวลาทำงานต่อภาพถึง 49 วินาที

2.3.3 Faster R-CNN (Faster Recurrent Convolutional Neural Network) [37-39]

สำหรับโครงสร้างแบบจำลองนิวรอนที่วิทยานิพนธ์นี้เลือกมาใช้คือ Faster R-CNN ซึ่งเป็นแบบจำลองที่พัฒนาต่อเนืองมาจาก R-CNN [35] และ Fast R-CNN [36] มีความสามารถในการตรวจหาตำแหน่งของวัตถุในรูปภาพและสามารถจำแนกวัตถุตามชนิดได้ โดย Faster R-CNN มีแนวความคิดหลักคือการนำเอาการค้นหาแบบเลือก (Selective search) มารวมไว้ในแบบจำลองประสาทเป็นโครงสร้างเดียวกัน พบว่าในการทำงานต่อ 1 ภาพใช้เวลาเพียง 0.2 วินาที

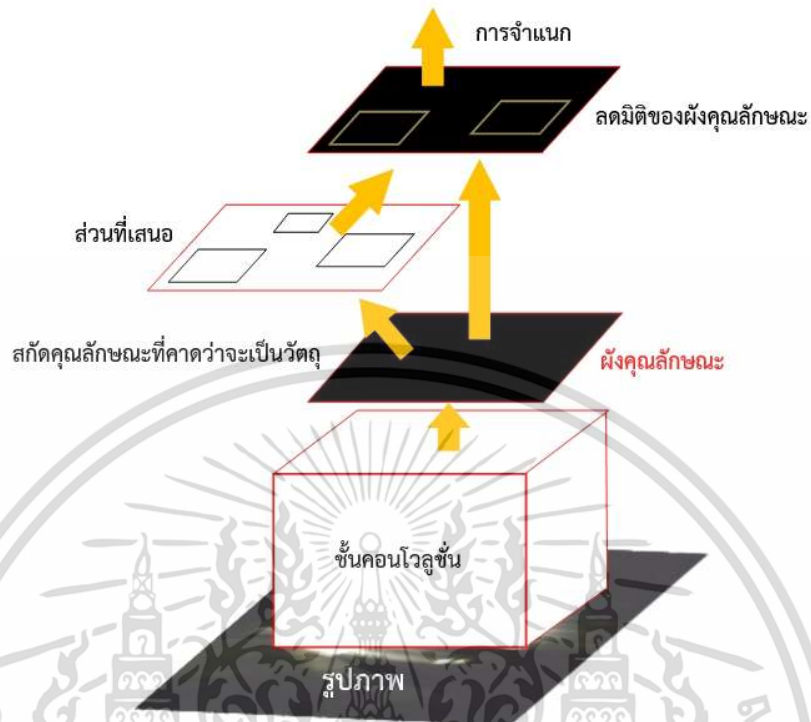


ภาพที่ 2.28 โครงสร้างแบบจำลอง Faster R-CNN

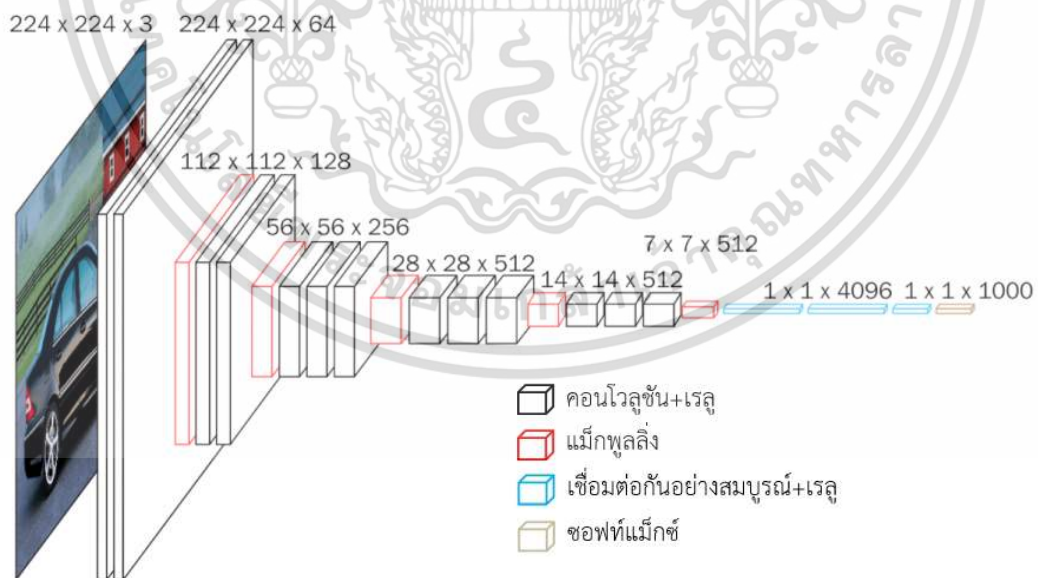
แบบจำลองนี้มีส่วนประกอบหลัก 3 ส่วนคือ

- 1) ส่วนฐานที่ทำหน้าที่สกัดคุณลักษณะ หรือ feature หรือพิจารณาลักษณะเด่นของวัตถุในภาพ
- 2) ส่วนชั้นจำแนกส่วนที่คาดว่าจะเป็วัตถุ (Region Proposal Network: RPN) ที่ทำหน้าที่แยกบริเวณที่ “คาดว่าจะหรืออาจจะ” เป็นวัตถุจากฟังก์ชันคุณลักษณะ (feature map)
- 3) ส่วนจำแนกประเภทที่นำเอาฟังก์ชันคุณลักษณะและขอบเขต (region) ที่ได้จาก RPN มาประมวลผลโดยทำ RoI pooling เพื่อตอบว่าบริเวณใดของภาพมีวัตถุอยู่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 2.29 โครงสร้าง Faster R-CNN จำแนกกลุ่มและ RPN ช่วยหาตำแหน่งที่น่าจะเป็นวัตถุ [37]



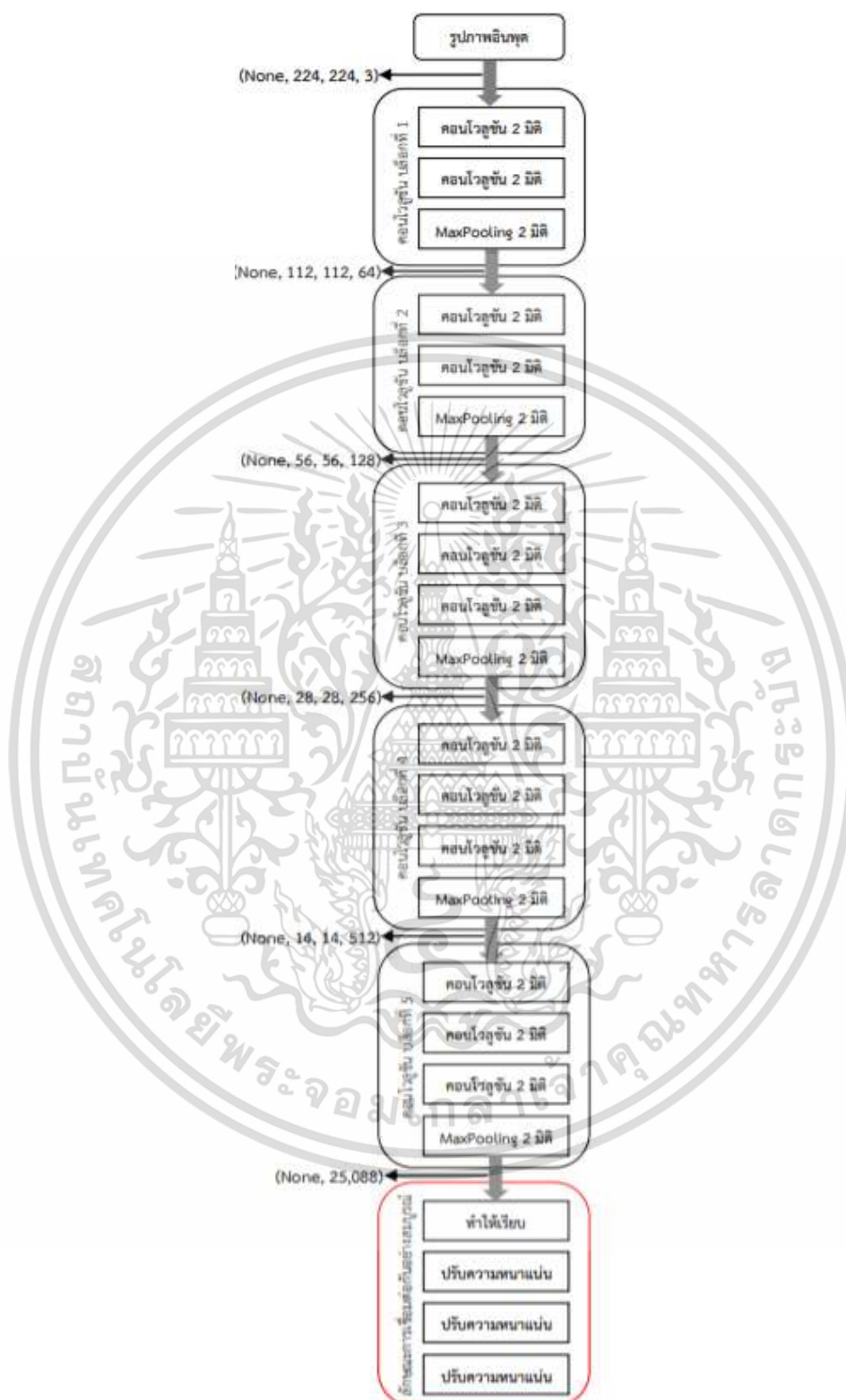
ภาพที่ 2.30 โครงสร้างของ VGG-16 ที่นำมาใช้เป็นพื้นฐานของโมเดล [40]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โครงสร้าง VGG-16 สามารถรองรับการประมวลผลกลับรูปภาพสีและขาวดำ โดยชั้นแรกสุดสามารถรองรับขนาดของรูปภาพขนาด 224x224 เปิดภาพแบบ RGB (Red Green Blue) โครงสร้างการทำงานจะแบ่งเป็นการคอนโวลูชันจำนวน 5 ครั้ง จากนั้นคอนโวลูชันแบบเชื่อมต่ออย่างสมบูรณ์ (full connection) โดยรายละเอียดสามารถแสดงได้ดังรูปภาพที่ 2.27



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 2.31 รายละเอียดการทำงานในแต่ละชั้นของโครงสร้าง VGG-16 [41]

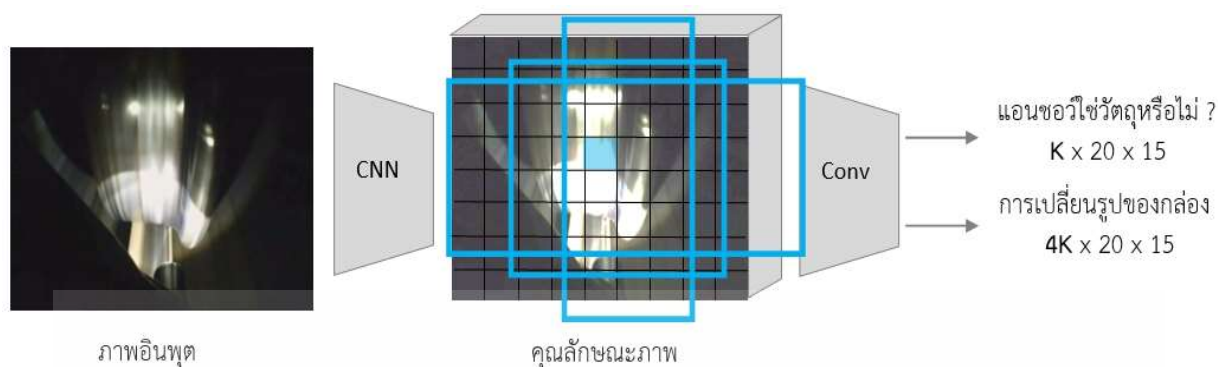
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทำงานของ Faster R-CNN มีการแบ่งการประมวลผลเป็นสองขั้นตอน ดังนี้

ขั้นตอนแรก เป็นการหาขอบเขต (region หรือ bounding box) ที่มีความน่าจะเป็นที่วัตถุอยู่เพื่อจะส่งเข้าสู่ขั้นตอนที่สองต่อไป ในขั้นตอนนี้จะมีส่วนประกอบที่สำคัญคือ

- **Backbone network** เป็นโครงข่ายประสาทเทียม (Neural network) ที่รับรูปภาพอินพุต (input) และประมวลผลจนได้เป็นผังคุณลักษณะ (Feature map) ออกมา โดยโครงข่ายแบบ Backbone (Backbone network)
- **Region proposal network** เป็นโครงข่ายประสาทเทียม (Neural network) ที่ทำการจำแนก (Classification) เพื่อจำแนกว่าขอบเขตไหนที่น่าจะมีวัตถุอยู่ รวมทั้งมีการทำการถดถอย (regression) เพื่อที่จะได้ขอบเขตที่แม่นยำขึ้น สำหรับ Region proposal network นี้เอง เป็นสิ่งที่ทำให้ Faster R-CNN แตกต่างจาก R-CNN ที่ผ่านมา ซึ่งใช้อัลกอริทึมแบบ Region proposal ภายนอกอย่าง Selective Search เข้ามาช่วย
- ขอบเขตที่ Region proposal network นำมาพิจารณา (อาจเรียกว่า Anchor box) คือกล่องสี่เหลี่ยมหลายๆขนาดที่ครอบคลุมผังคุณลักษณะ (feature map) ในตำแหน่งต่างๆ ดังเช่นรูปด้านล่างนี้แสดงตัวอย่างการทำงานของขั้นตอนแรก เมื่อรูปขนาด 640x480 ผ่าน CNN ที่เป็น Backbone network เข้ามาจะได้ผังคุณลักษณะเด่น (feature map) ขนาด 20x15 และมีจำนวนช่อง (channel) เท่ากับ 512
- กำหนดให้แต่ละจุดของผังคุณลักษณะมี anchor box จำนวน K ขนาด จะได้ว่ามีจำนวน anchor box ที่ต้องพิจารณาทั้งหมดเท่ากับ $K \times 20 \times 15$ และในจำนวน anchor box ทั้งหมดนี้ ในส่วนของ region proposal network ก็จะทำการศึกษาขอบเขตที่น่าจะเป็นวัตถุมากที่สุดมาจำนวนหนึ่ง เช่น 300 ขอบเขต เพื่อส่งเป็น Region proposal หรือพื้นที่ที่สนใจ (Region of Interest : RoI) สำหรับขั้นตอนต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 2.32 ขอบเขตที่ Region proposal network นำมาพิจารณา [42]

ขั้นตอนที่สอง ประกอบด้วยกระบวนการต่างๆ ดังนี้

- RoI Pool เป็นการทำ RoI ของฟังก์ชันคุณลักษณะ (feature map) ที่มีขนาดแตกต่างกันให้มีขนาดเดียวกัน เพื่อเป็นอินพุต (input) ให้กับโครงข่ายประสาทเทียม (Neural network) ในขั้นต่อไป โดย RoI Pool จะใช้ max pooling
- ส่วนสุดท้ายจะเป็นส่วนจำแนก (classification) ที่ทำการจำแนกว่าเป็นวัตถุประเภทใด (ส่วนแยกชนิด class) และส่วน regression ที่ทำการปรับขอบเขตของ RoI ให้แม่นยำขึ้น (ส่วน box)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

การตรวจจับอันตรายระหว่างล้อและรางรถไฟด้วยเทคนิค Faster R-CNN

จากบทที่ 2 ที่ได้กล่าวมาแล้วว่าเทคนิค Faster R-CNN เป็นเทคนิคที่มีความสามารถในการตรวจหาตำแหน่งของวัตถุในรูปภาพและสามารถจำแนกวัตถุตามชนิดได้ทำให้สามารถประมวลผลรูปภาพได้จำนวนมากและรวดเร็ว งานวิจัยนี้จึงเลือกใช้เทคนิค Faster R-CNN มาเป็นเครื่องมือในศึกษาการตรวจจับอันตรายระหว่างล้อและรางรถไฟเพื่อควบคุมความปลอดภัยระหว่างที่รถไฟมีการเคลื่อนที่จากกล้องวิดีโอที่ติดตั้งตรงล้อรถไฟบริเวณใต้ตู้โดยสารรถไฟบ้านที่ภาพการเคลื่อนที่ขณะที่รถไฟวิ่งแล้วทำการแปลงสัญญาณภาพวิดีโอให้เป็นรูปภาพแล้วนำรูปภาพดังกล่าวมาประมวลผลด้วย Faster R-CNN โดยขั้นตอนวิธีการในการตรวจจับอันตรายระหว่างล้อและรางรถไฟด้วยเทคนิค Faster R-CNN มีดังนี้

1. กำหนดขอบเขตเตรียมภาพ
2. เทคนิคปัญญาประดิษฐ์
3. การเรียนรู้การจดจำและหาตำแหน่งซึ่งจะแสดงการวิเคราะห์แบบตามเวลาจริง

3.1 การตรวจจับอันตรายระหว่างล้อและรางรถไฟด้วยเทคนิค Faster R-CNN

การนำเทคนิค Faster R-CNN มาตรวจจับอันตรายระหว่างล้อและรางรถไฟ คอมพิวเตอร์ที่ใช้ประมวลผลมีข้อมูลอุปกรณ์ดังนี้

Intel Core™ i5-8600

CPU @ 3.10GHz

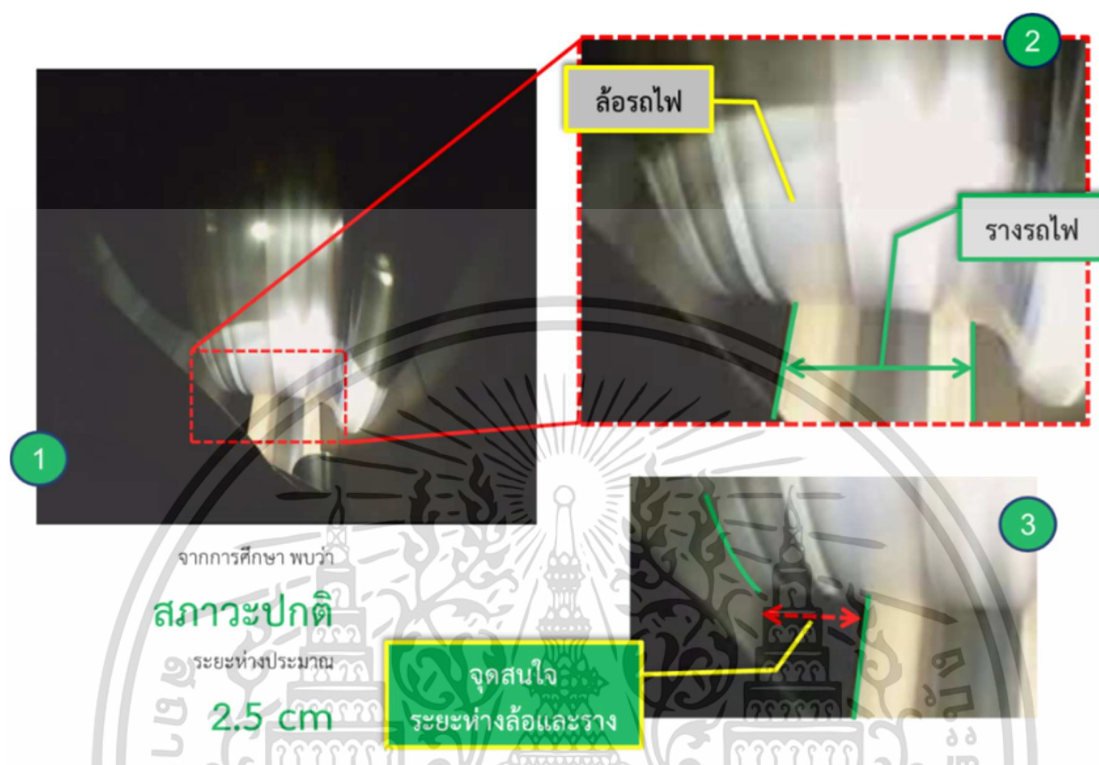
Ram: 16 GB.

GPU: GeForce GTX 1080

NVIDIA CUDA Core: 2560

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

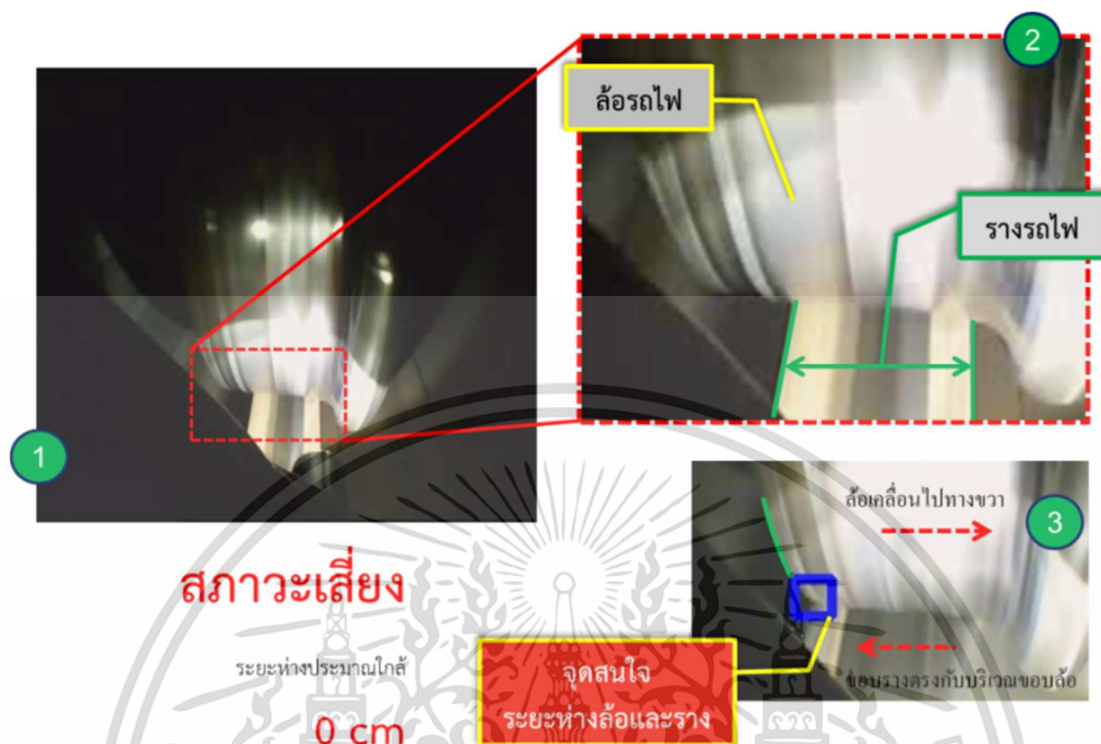
3.1.1 กำหนดขอบเขต ส่วนที่พิจารณาเพื่อจัดเตรียมภาพ



ภาพที่ 3.1 พื้นที่บริเวณล้อและรางเพื่อกำหนดจุดพิจารณา

พิจารณาภาพที่ 3.1 เป็นตัวอย่างสัญญาณภาพที่ได้จากการบันทึกจากกล้องวิดีโอที่ติดตั้งตรงล้อรถไฟฟ้า บริเวณใต้ตู้โดยสาร ในการเก็บข้อมูลรูปภาพโดยใช้กล้องบันทึกข้อมูลในรูปแบบวิดีโอโดยจากภาพจะเห็นล้อรถไฟ และรางรถไฟจากภาพที่ 3.1 หมายเลข 3 เมื่อพิจารณาตำแหน่งของล้อรถไฟและรางรถไฟจะเห็นระยะห่างระหว่าง ล้อและรางซึ่งเป็นสภาวะปกติของการเคลื่อนที่ของรถไฟฟ้าและจะใช้ระยะห่างล้อและรางรถไฟฟ้านี้ในการ วิเคราะห์การประเมินคุณภาพความปลอดภัยและสภาพทางวิ่ง ซึ่งเมื่อนำสัญญาณภาพมาพิจารณาลักษณะที่เกิด ความเสี่ยงพบว่าลักษณะของล้อและรางจะเป็นดังภาพที่ 3.2

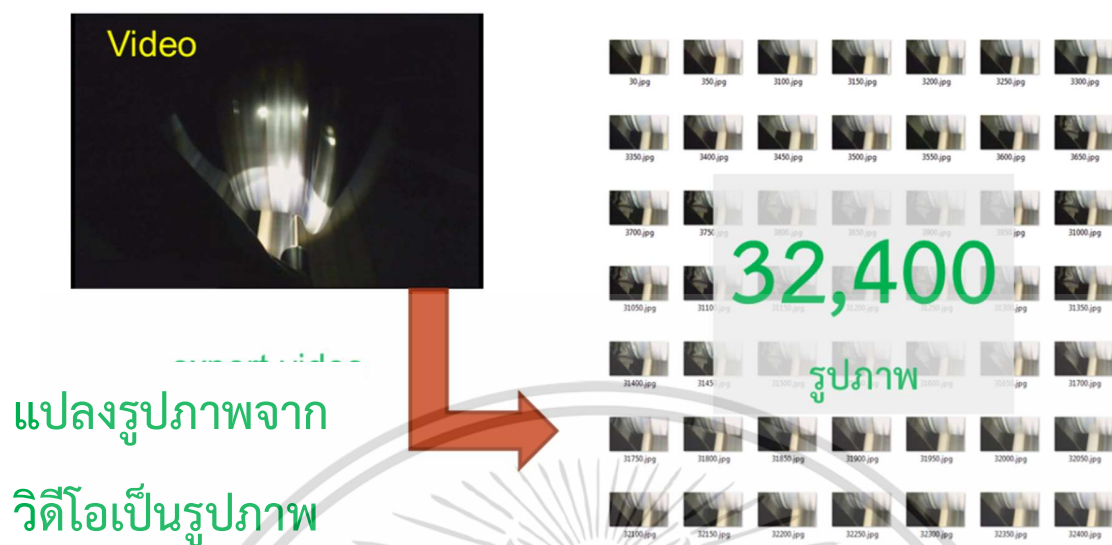
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



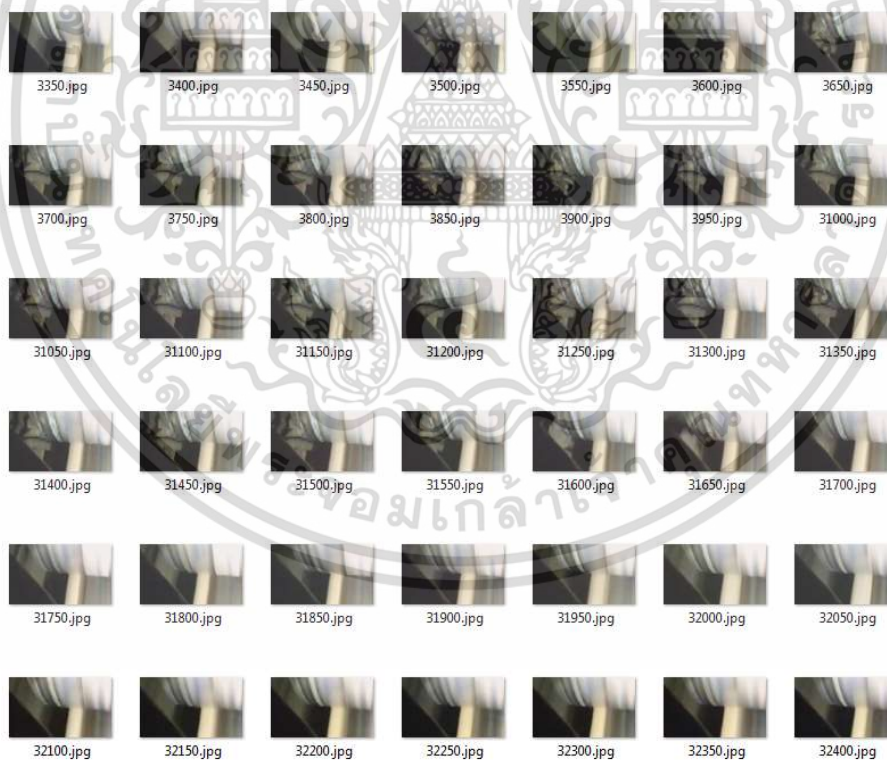
ภาพที่ 3.2 ภาพล่อและรางแสดงสภาพความเสี่ยงที่อาจเกิดความเสียหายต่อล่อ

จากภาพที่ 3.2 หมายเลข 3 จะเห็นได้ว่าบริเวณขอบล้อด้านนอกเคลื่อนที่เข้าใกล้ขอบรางเป็นอย่างมาก ซึ่งส่งผลให้ขอบของบังใบของล่ออีกข้างเกิดการกระทบกับรางซึ่งอาจส่งผลให้เกิดความเสียหายกับรางและล่อหรือ อาจส่งผลให้รถไฟขาดรางได้

เมื่อทำการกำหนดบริเวณที่สนใจแล้วจึงนำสัญญาณภาพวิดีโอมาทำการแปลงภาพเป็นรูปภาพจำนวน 32,400 ภาพ เพื่อใช้สำหรับการเรียนรู้และจดจำของเทคนิคปัญญาประดิษฐ์



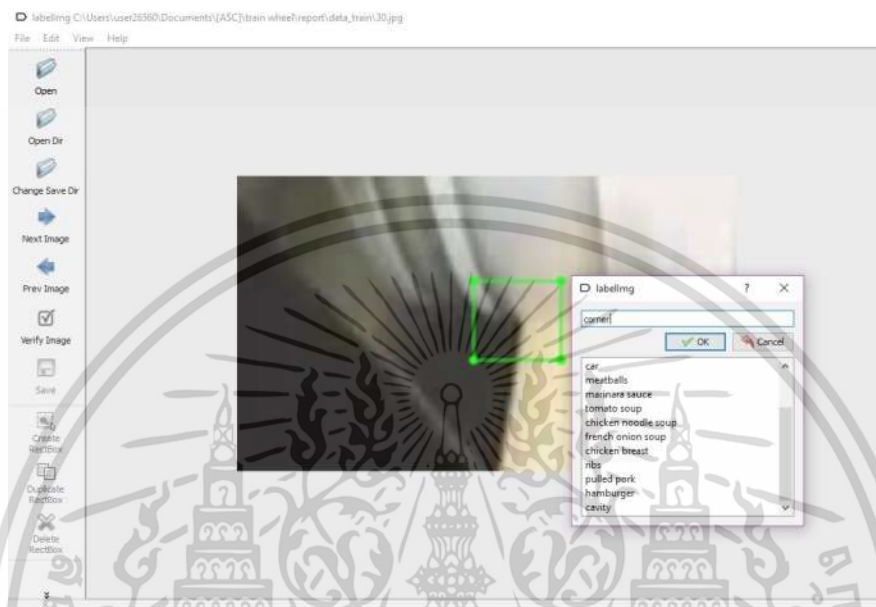
ภาพที่ 3.3 การแปลงรูปภาพจากวิดีโอไปเป็นรูปภาพทั้งหมด



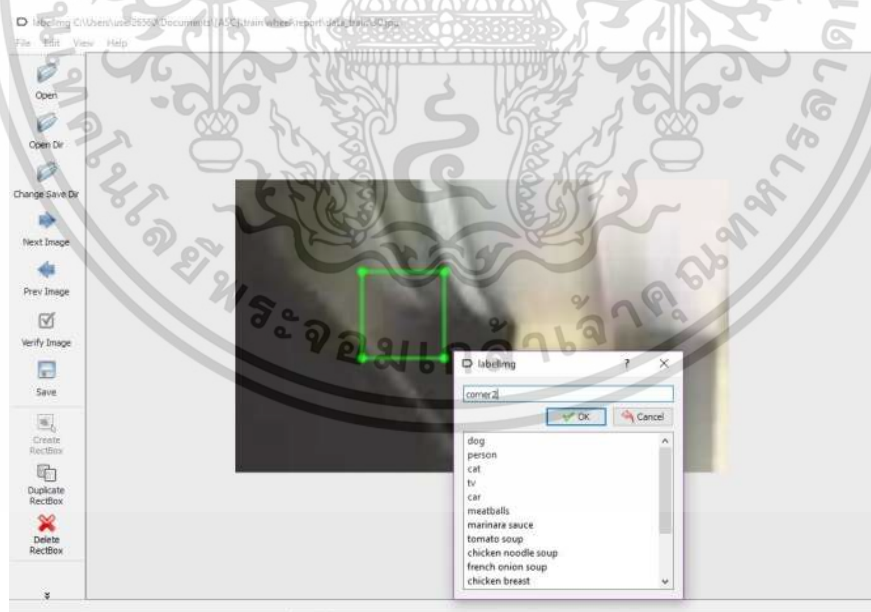
ภาพที่ 3.4 ตัวอย่างรูปภาพของล้อและรางที่ใช้ทดสอบการประเมินประสิทธิภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากข้อมูลจำนวนทั้งหมดได้แยกข้อมูลออกเป็น 2 ส่วนหลักคือ ข้อมูลที่ใช้สำหรับให้เทคนิค ปัญญาประดิษฐ์เรียนรู้และข้อมูลที่ใช้ในการทดสอบความสามารถของ AI โดยกำหนดพื้นที่ 2 บริเวณคือ บริเวณ ขอบรางและบริเวณขอบล้อ ซึ่งสามารถแสดงได้ดังตัวอย่างรูปภาพที่ 3.5 (ก) และ 3.5 (ข) ตามลำดับ



(ก) บริเวณขอบราง

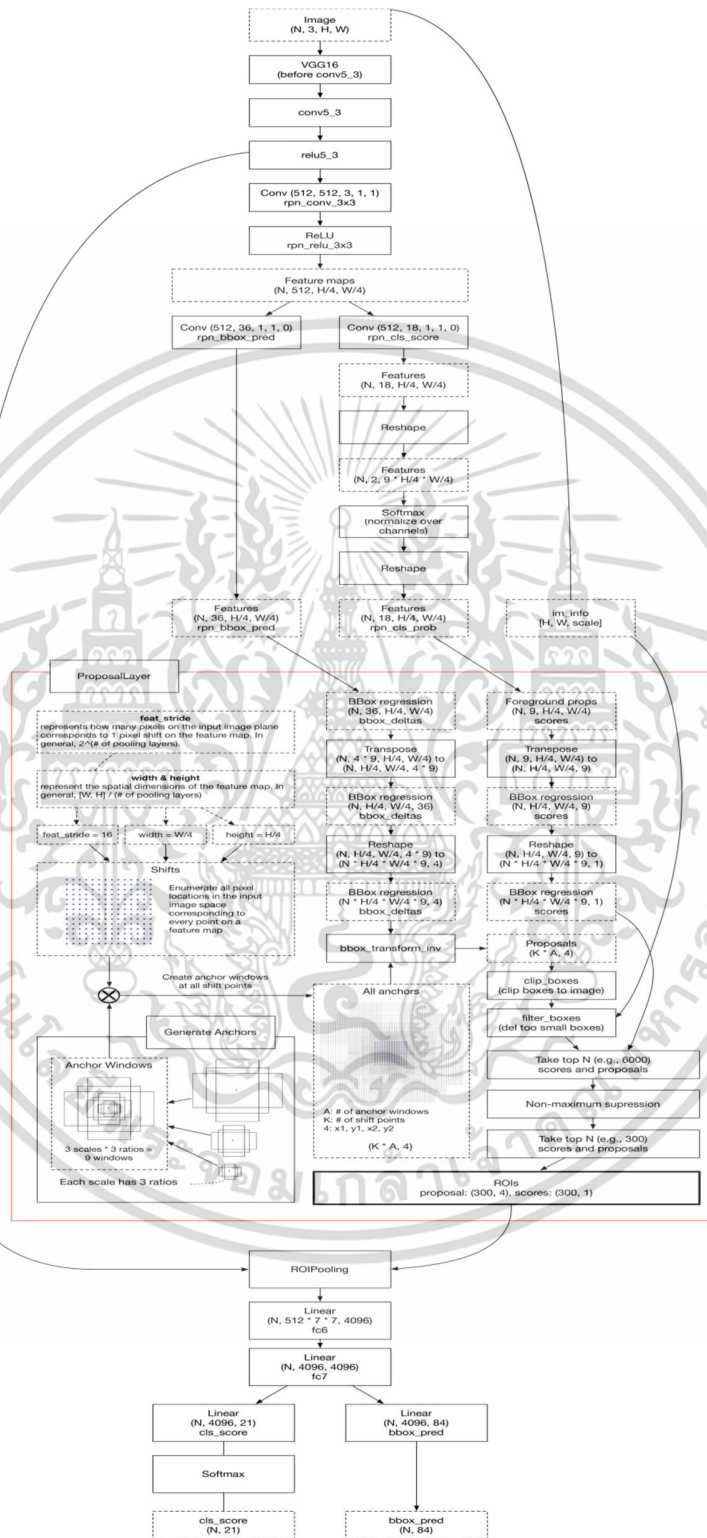


(ข) บริเวณขอบล้อ

ภาพที่ 3.5 การเลือกบริเวณที่สนใจซึ่งก็คือบริเวณขอบรางและบริเวณขอบล้อ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.2 โครงสร้างแบบจำลองปัญญาประดิษฐ์



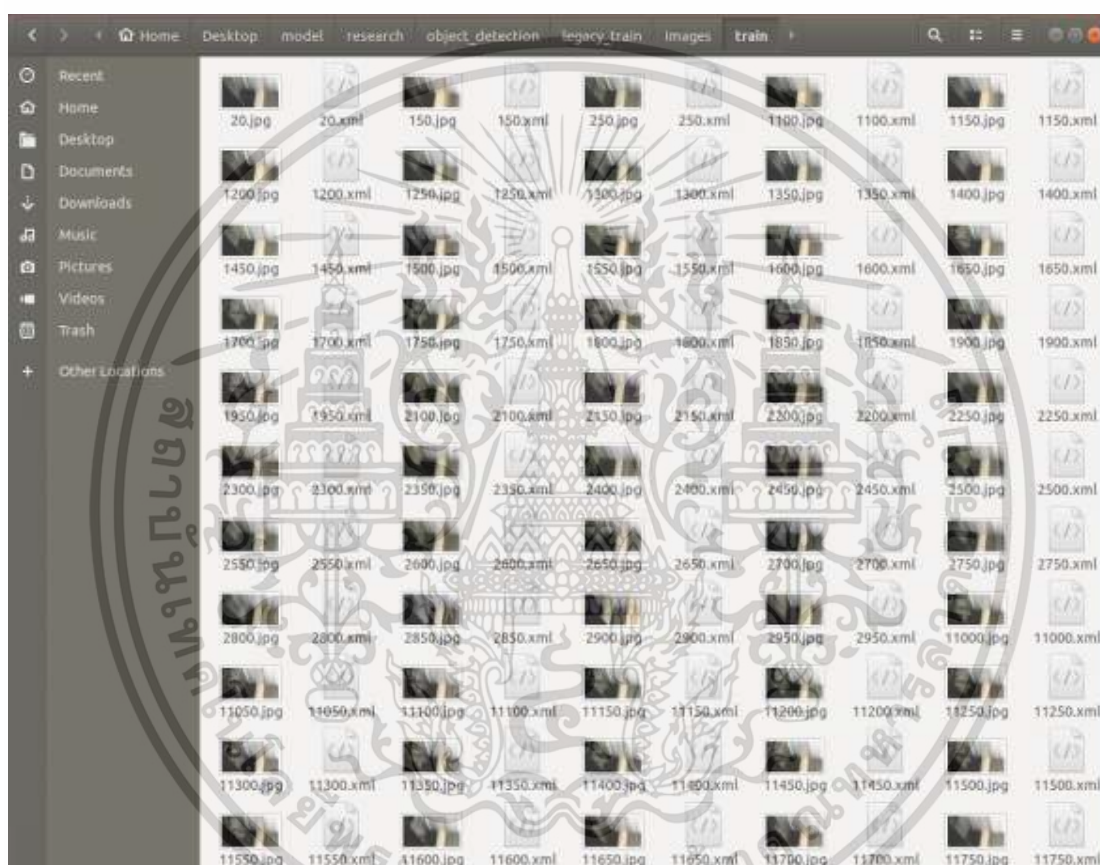
ภาพที่ 3.6 รายละเอียดโครงสร้าง Faster R-CNN

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากภาพที่ 3.6 รูปภาพจะผ่าน VGG-16 ทำการคอนโวลูชัน เรลูแล้วแยกคุณลักษณะเด่นของภาพออกมา จากนั้นเข้าสู่ชั้น Proposal จะได้พื้นที่ที่สนใจใน RO pooling คู่ชั้นเชิงเส้น (linear) และสรุปผล

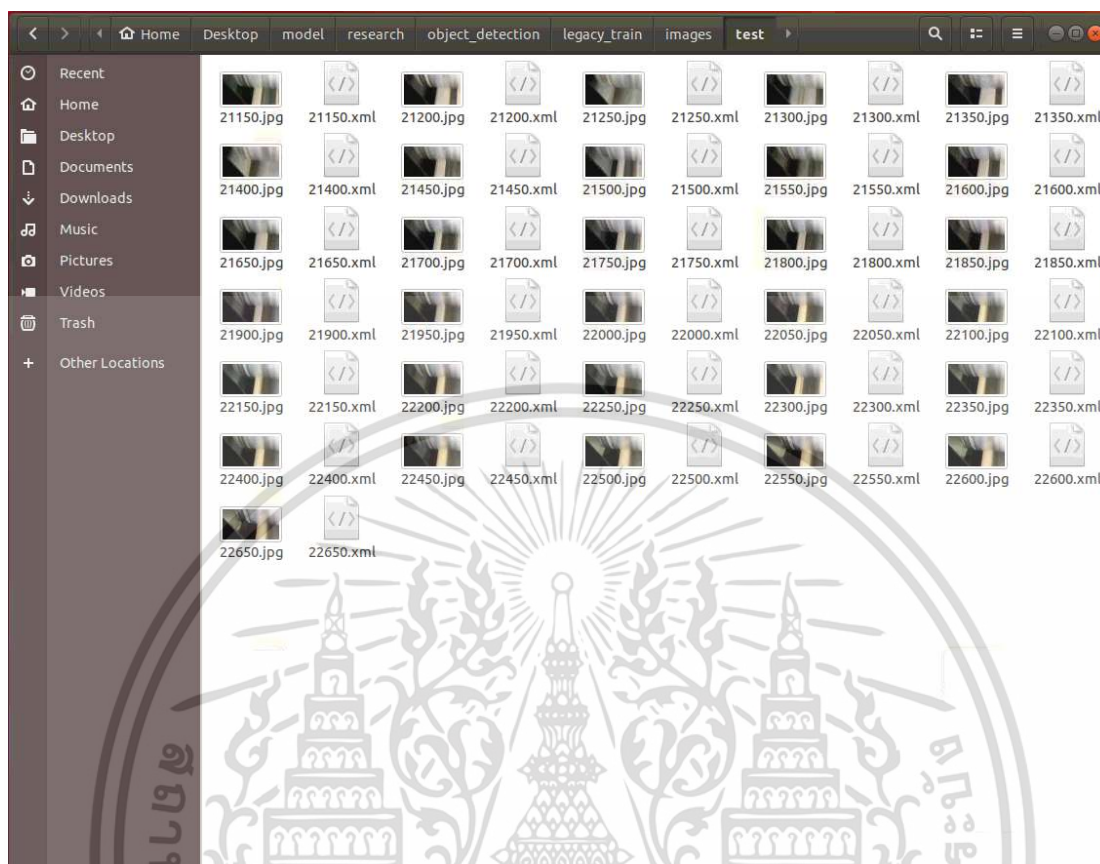
3.1.3 การเรียนรู้จดจำเพื่อวิเคราะห์หาตำแหน่งของล้อและราง

ขั้นตอนนี้เป็นการจัดเตรียมรูปภาพที่กำหนดพื้นที่บริเวณขอบล้อและบริเวณขอบราง (ดังรูปภาพ 3.7) ด้วยโปรแกรม LabelImg โดยกำหนดข้อมูลการเรียนรู้จำนวน 50 รูป และข้อมูลทดสอบจำนวน 31 รูปภาพ



(ก) ส่วนเรียนรู้ (training)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ข) ส่วนทดสอบความถูกต้อง (testing)

ภาพที่ 3.7 รูปภาพและป้ายกำกับ

เมื่อเตรียมข้อมูลภาพแล้วก็จะนำภาพเหล่านี้ไปให้ปัญญาประดิษฐ์เรียนรู้และทำการทดสอบด้วยโปรแกรม Tensorflow โดยโปรแกรม Tensorflow จะแสดงกราฟ Cross Entropy ออกมา ซึ่งกราฟ Cross Entropy จะเป็นตัวบอกผลที่ได้จากการเรียนรู้และการทดสอบของปัญญาประดิษฐ์ซึ่งก็คือค่า loss นั้นเอง

3.2 การหาระยะห่างของล้อและรางด้วย Faster R-CNN โดยกำหนดจุดขอบของล้อรถไฟให้เป็นจุดอ้างอิง

การนำเทคนิค Faster R-CNN มาตรวจจับอันตรกิริยาระหว่างล้อและรางรถไฟโดยกำหนดจุดขอบของล้อรถไฟให้เป็นจุดอ้างอิง คอมพิวเตอร์ที่ใช้ประมวลผลมีข้อมูลอุปกรณ์ดังนี้

Intel Core Xeon CPU @2.20 GHz

Ram:12 GB

GPU: Nvidia Tesla K80

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

NVIDIA CUDA Core:2496

3.2.1 เตรียมภาพในการฝึกฝนให้แบบจำลองเรียนรู้

การเตรียมข้อมูลภาพในการฝึกฝนให้แบบจำลองเรียนรู้ในแต่ละภาพจะต้องกำหนดค่า Label (x, y) ของบริเวณที่สนใจหรือตำแหน่งของขอบล้อ





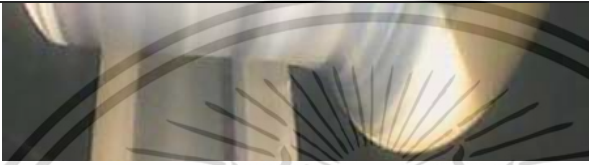
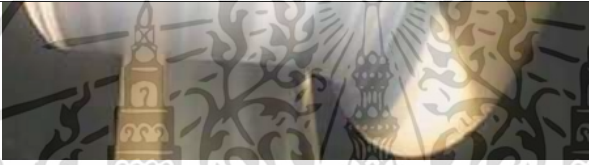
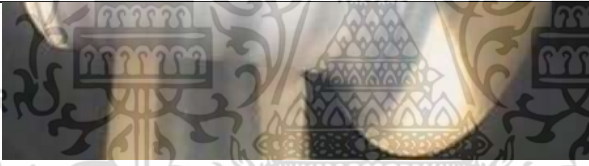
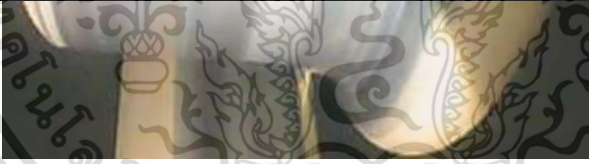
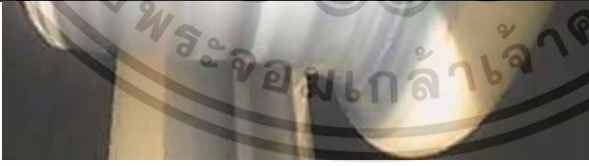

ภาพที่ 3.8 การเตรียมข้อมูลภาพสำหรับใช้ในขั้นตอนการเรียนรู้

จากวิดีโอความยาว 55 นาที ทำการแปลงมาเป็นรูปภาพจำนวน 3,339 ภาพ ซึ่งภาพที่ใช้สำหรับการเรียนรู้เตรียมไว้จำนวน 2,586 ภาพ โดยตัวอย่างของภาพที่เตรียมไว้บางส่วน แสดงในตารางที่ 3.2



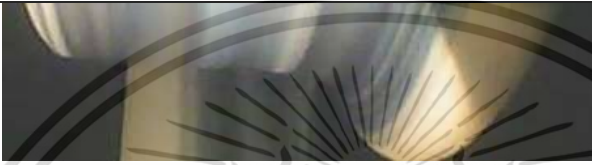
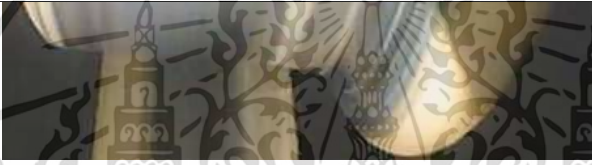
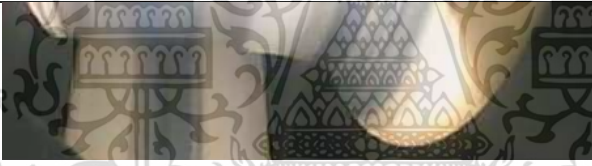
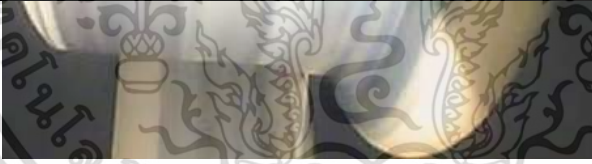


ตารางที่ 3.1 ภาพและค่า Label (x, y) ของตำแหน่งขอบล้อ

| ลำดับ ที่ | ภาพที่เตรียม | ค่า Label (x, y) |
|--------------|--|------------------|
| 1 |  | (33,32) |
| 2 |  | (35,31) |


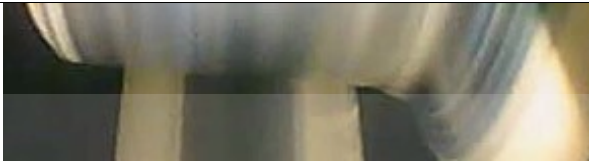
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

| | | |
|----|--|---------|
| 3 |  | (30,29) |
| 4 |  | (30,33) |
| 5 |  | (19,28) |
| 6 |  | (4,24) |
| 7 |  | (38,31) |
| 8 |  | (12,25) |
| 9 |  | (9,24) |
| 10 |  | (17,29) |

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

| | | |
|----|--|---------|
| 11 |  | (29,29) |
| 12 |  | (25,26) |
| 13 |  | (35,37) |
| 14 |  | (20,31) |
| 15 |  | (6,23) |
| 16 |  | (30,33) |
| 17 |  | (11,21) |
| 18 |  | (36,35) |

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

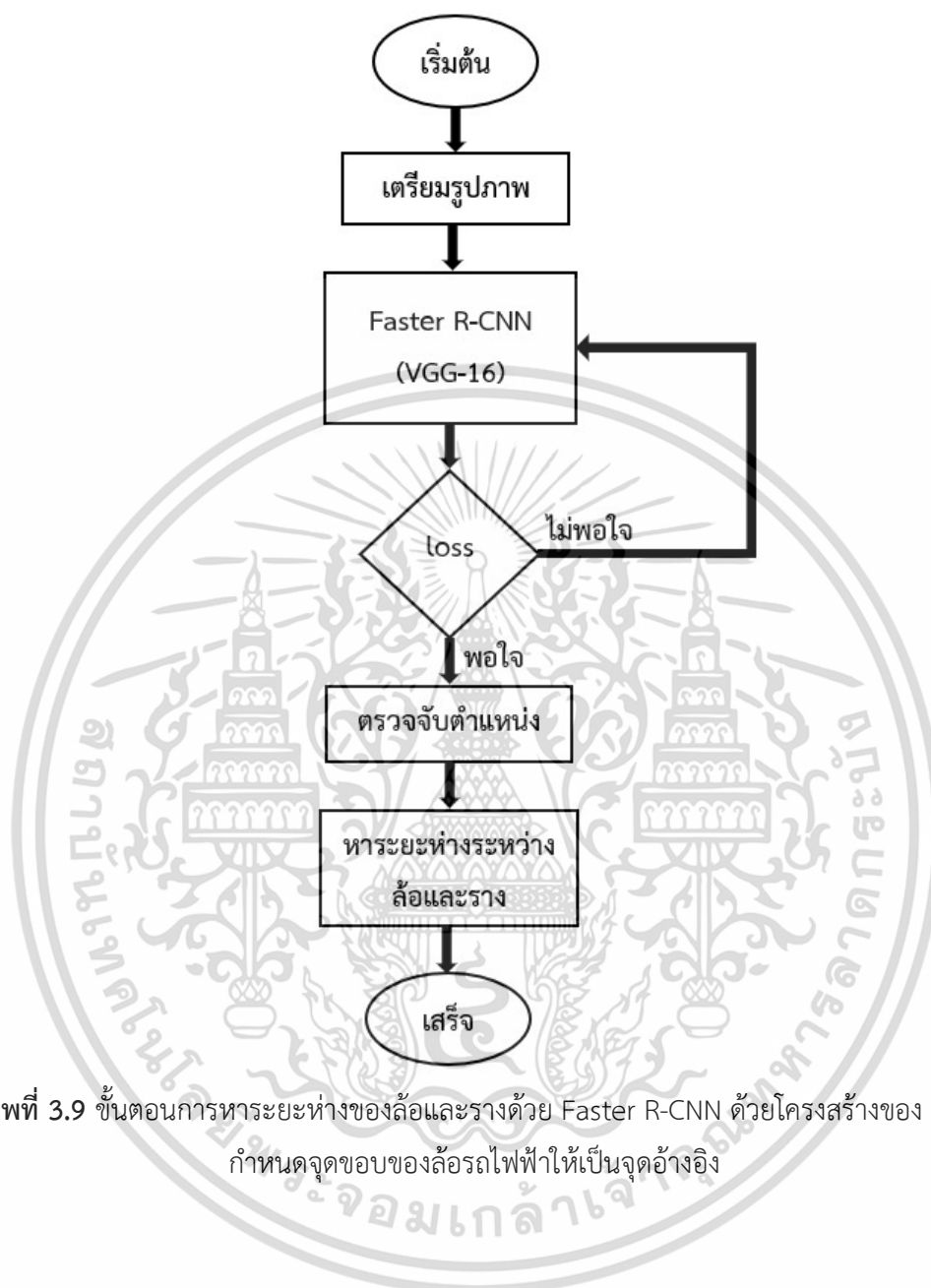
| | | |
|----|--|---------|
| 19 |  | (4,24) |
| 20 |  | (17,32) |

ตารางที่ 3.1 แสดงตัวอย่างภาพและค่า Label (x, y) ของตำแหน่งของขอบล้อเพื่อใช้ในการเตรียมข้อมูลภาพสำหรับใช้ในขั้นตอนการเรียนรู้เพื่อฝึกฝนแบบจำลอง

เมื่อเตรียมภาพแล้วก็นำภาพเหล่านี้ไปให้ปัญญาประดิษฐ์เรียนรู้และทำการทดสอบด้วย Open source Computer Vision ซึ่งเป็นไลบรารีฟังก์ชันการเขียนโปรแกรมมาใช้ในการหาผลที่ได้จากการเรียนรู้และการทดสอบของปัญญาประดิษฐ์ซึ่งหรือค่า loss โดยนำเสนอ 2 วิธีคือ การหาระยะห่างของล้อและรางด้วย Faster R-CNN โดยกำหนดจุดขอบของล้อรถไฟให้เป็นจุดอ้างอิงโดยใช้โครงสร้างของ VGG-16 และการหาระยะห่างของล้อและรางด้วย Faster R-CNN โดยกำหนดจุดขอบของล้อรถไฟให้เป็นจุดอ้างอิงโดยลดชั้นโครงสร้างของ VGG-16

3.2.1.1 การหาระยะห่างของล้อและรางด้วย Faster R-CNN ขึ้นด้วยโครงสร้างของ VGG-16 โดยกำหนดจุดขอบของล้อรถไฟให้เป็นจุดอ้างอิง

ในการหาระยะห่างของล้อและรางด้วย Faster R-CNN โดยกำหนดจุดขอบของล้อรถไฟให้เป็นจุดอ้างอิงด้วยโครงสร้างของ VGG-16 เมื่อเตรียมภาพสำหรับให้ปัญญาประดิษฐ์เรียนรู้และทดสอบแล้วจึงนำภาพที่เตรียมไปเข้า Faster R-CNN จนกว่าค่า loss จะมีค่าเข้าใกล้ศูนย์ โดยมีขั้นตอนตามภาพที่ 3.9



ภาพที่ 3.9 ขั้นตอนการหาระยะห่างของล้อและรางด้วย Faster R-CNN ด้วยโครงสร้างของ VGG-16 โดยกำหนดจุดขอบของล้อรถไฟให้เป็นจุดอ้างอิง

3.2.1.2 การหาระยะห่างของล้อและรางด้วย Faster R-CNN ด้วยโครงสร้างของ VGG-แบบลดชั้นโดยกำหนดจุดขอบของล้อรถไฟให้เป็นจุดอ้างอิง

เมื่อเตรียมภาพสำหรับให้ปัญญาประดิษฐ์เรียนรู้และทดสอบแล้วจึงนำภาพที่เตรียมไปเข้า Faster R-CNN จนกว่าค่า loss จะมีค่าเข้าใกล้ศูนย์

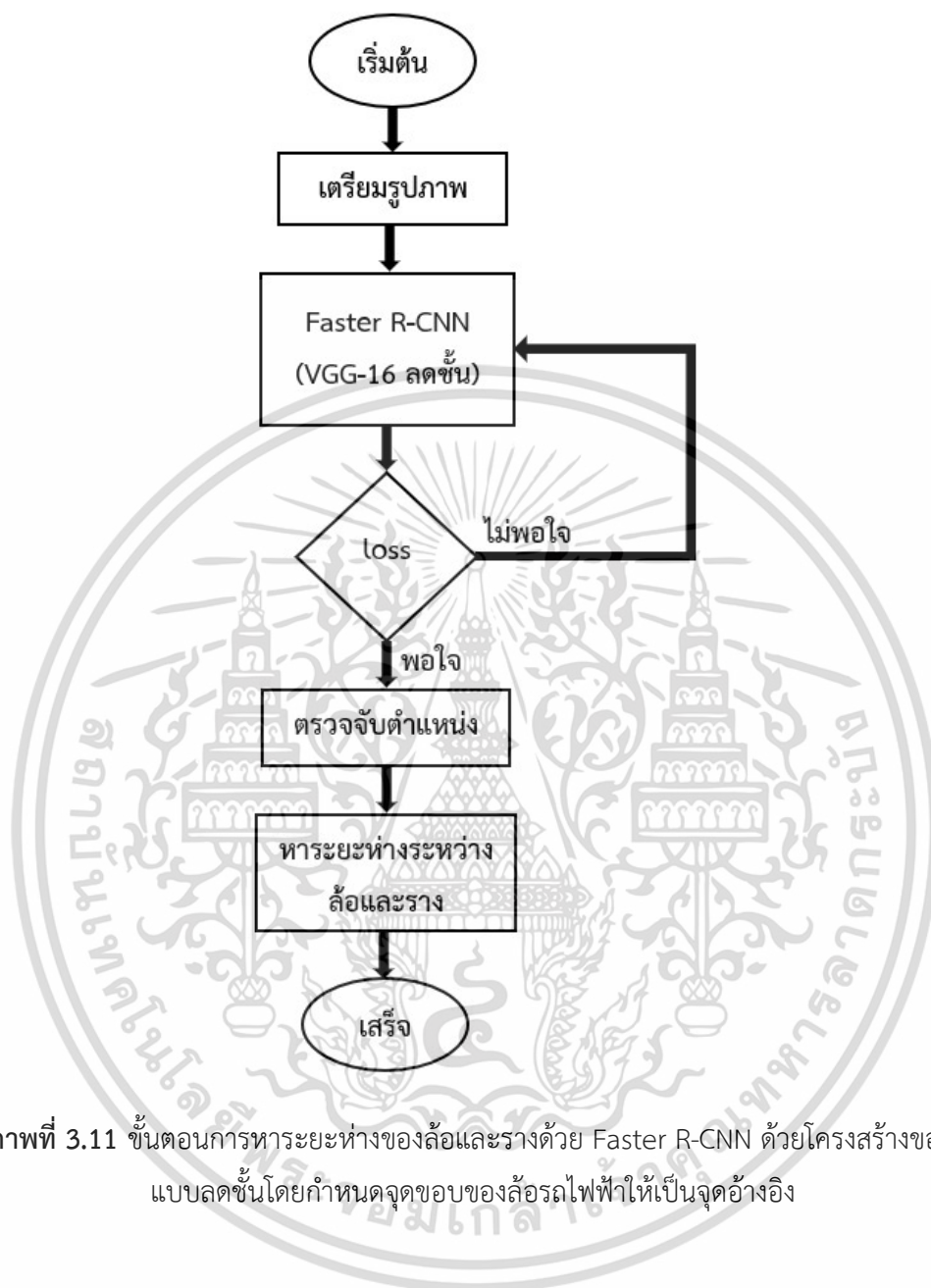
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในการหาระยะห่างของล้อและรางด้วย Faster R-CNN โดยกำหนดจุดขอบของล้อรถไฟให้เป็นจุดอ้างอิงด้วยโครงสร้างของ VGG-16 แบบลดชั้นที่นำเสนอ โครงสร้างของ VGG-16 เหลือเพียงการรับภาพอินพุตเข้ามาทำการคอนโวลูชันเพียงแค่ 2 ครั้ง แล้วพูลลิง 1 ครั้ง เพื่อลดขั้นตอนและข้อมูลในการวิเคราะห์ ดังรูปที่ 3.10



ภาพที่ 3.10 วิธีการนำเสนอการลดชั้นโครงสร้างของ VGG-16 ที่นำมาใช้เป็นพื้นฐานของโมเดล

โดยการหาระยะห่างของล้อและรางด้วย Faster R-CNN โดยกำหนดจุดขอบของล้อรถไฟให้เป็นจุดอ้างอิงด้วยโครงสร้างของ VGG-16 แบบลดชั้น มีขั้นตอนตามภาพที่ 3.11



ภาพที่ 3.11 ขั้นตอนการหาระยะห่างของล้อและรางด้วย Faster R-CNN ด้วยโครงสร้างของ VGG-16 แบบลดชั้นโดยกำหนดจุดขอบของล้อรถไฟให้เป็นจุดอ้างอิง

บทที่ 4

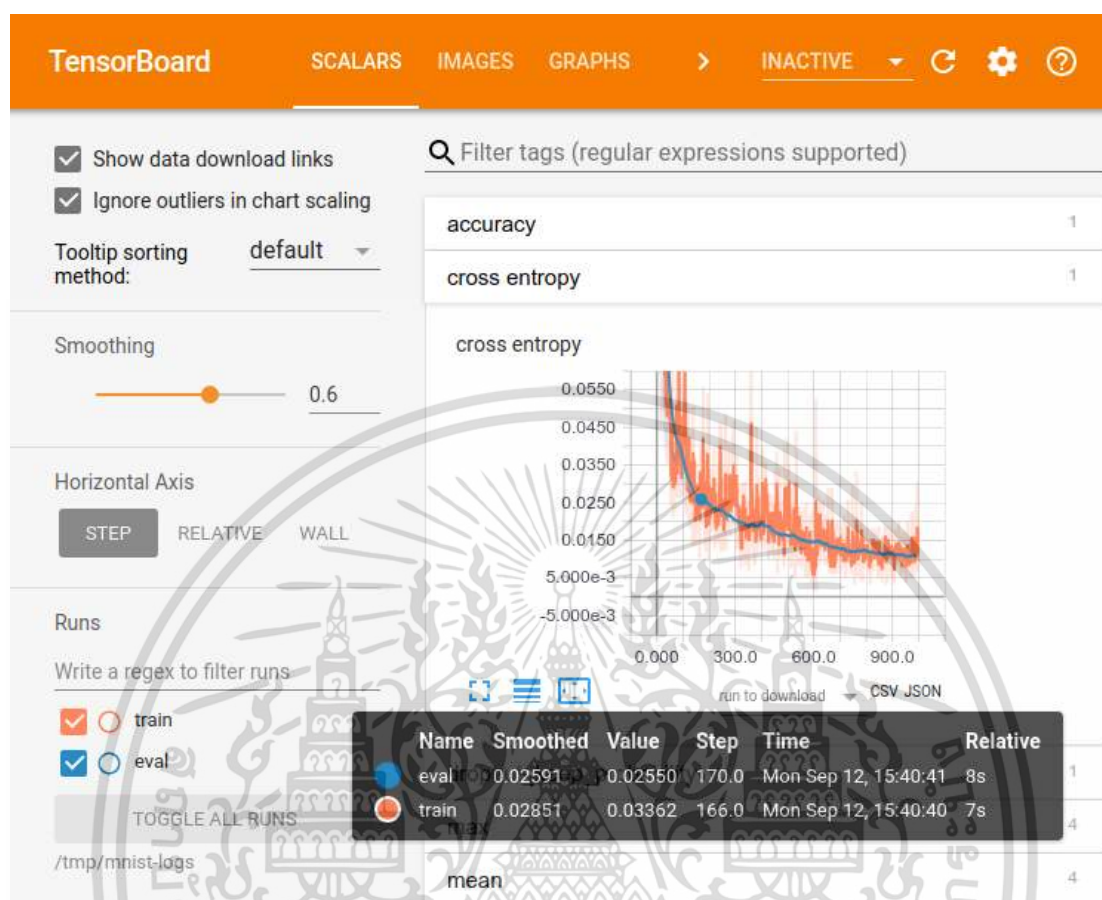
ผลการทดลอง

จากบทที่ 3 ได้ทำการเตรียมข้อมูลภาพสำหรับการเรียนรู้และการทดสอบของปัญญาประดิษฐ์ที่ใช้ในเทคนิค Faster R-CNN จะได้ผลการเรียนรู้ ผลการทดสอบ และผลการตรวจจับอันตรกิริยาระหว่างล้อและรางรถไฟฟ้า ดังนี้

4.1 ผลการตรวจจับอันตรกิริยาระหว่างล้อและรางรถไฟฟ้าด้วยเทคนิค Faster R-CNN

4.1.1 ผลจากการเรียนรู้และการทดสอบของปัญญาประดิษฐ์ด้วยเทคนิค Faster R-CNN

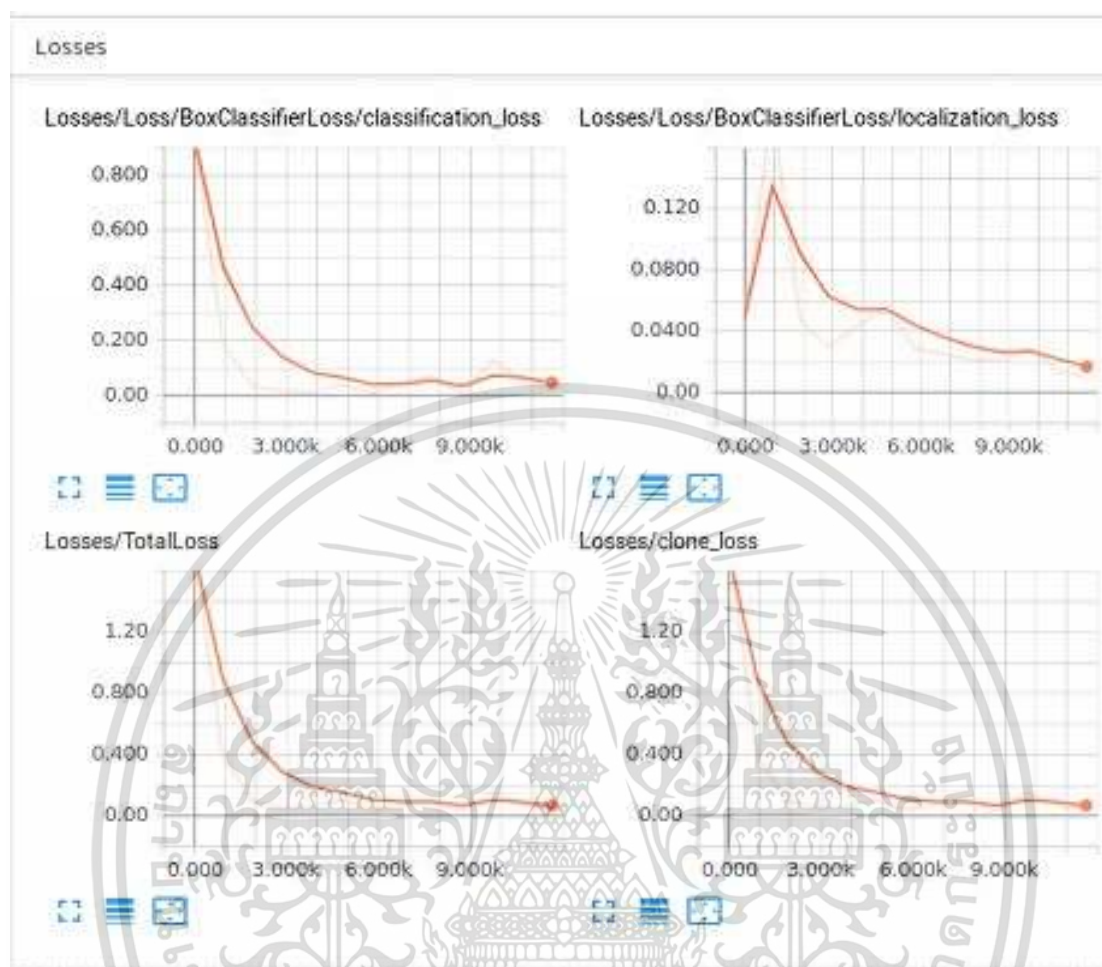
จากการเตรียมรูปภาพโดยกำหนดข้อมูลการเรียนรู้จำนวน 50 รูป และข้อมูลทดสอบจำนวน 31 รูปภาพ ในการจัดเตรียมข้อมูลในหนึ่งรูปภาพจะมีไฟล์ label กำกับตำแหน่งของวัตถุที่สนใจ โดยข้อมูลไฟล์จะประกอบไปด้วยเส้นทาง (path) ที่อยู่ของตำแหน่งไฟล์และ X, Y, width, height ซึ่งข้อมูลส่วนนี้จะถูกใช้เพื่อตรวจสอบความถูกต้องของผลลัพธ์ของปัญญาประดิษฐ์



ภาพที่ 4.1 กราฟ TensorBoard

จากภาพที่ 4.1 จะเห็นได้ว่าค่า loss จากการเรียนรู้ (Smoothed) เท่ากับ 0.02591 และค่า loss จากการทดสอบ (Smoothed) เท่ากับ 0.02851 จากการรันจำนวน 1,000 รอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



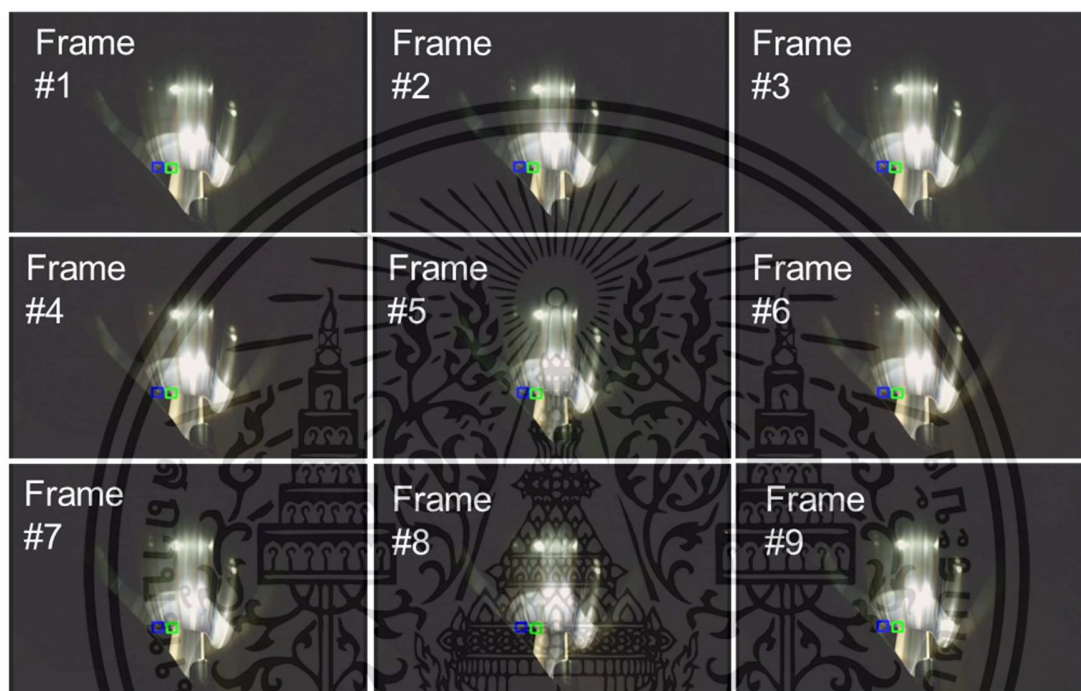
ภาพที่ 4.2 ผลลัพธ์ค่าสูญเสีย TensorBoard ที่ทำการรันจำนวน 10,000 รอบ

จากภาพที่ 4.2 กราฟ TensorBoard จะแสดงผลลัพธ์ที่ปัญญาประดิษฐ์สามารถฝึกสอนแบบจำลองได้โดยกราฟจะแสดงค่า loss ที่แสดงค่าความแม่นยำของแต่ละภาพในการเรียนรู้แบบจำลอง ซึ่งแกน x หมายถึง (รอบการทำงาน) แกน y หมายถึง ค่า loss โดยค่า loss จะแบ่งออกเป็น 2 ส่วนด้วยกันคือ loss ที่เกิดจากการหาตำแหน่งและ loss ที่เกิดจากการจำแนก ซึ่งในการทดสอบผลการรันจำนวน 10,000 รอบผลลัพธ์มีค่าเท่ากับ 0.02 และ 0.04 ตามลำดับ โดยปกติขั้นตอนการเรียนรู้จะมีป้อนข้อมูลการเรียนรู้ไปยังเทคนิคปัญญาประดิษฐ์และปัญญาประดิษฐ์จะวิเคราะห์หาผลลัพธ์เปรียบเทียบกับจากป้ายกำกับ ถ้ามีค่าผลลัพธ์แตกต่างกันมากก็จะทำการอัปเดตค่าน้ำหนักใหม่ไปเรื่อยๆจนกว่าค่าผลลัพธ์จะใกล้เคียงกับป้ายกำกับ สังเกตจากกราฟจะพบว่ากราฟจะค่อยๆลู่เข้าสู่ค่าศูนย์หรือเข้าสู่แกนในแนวนอน

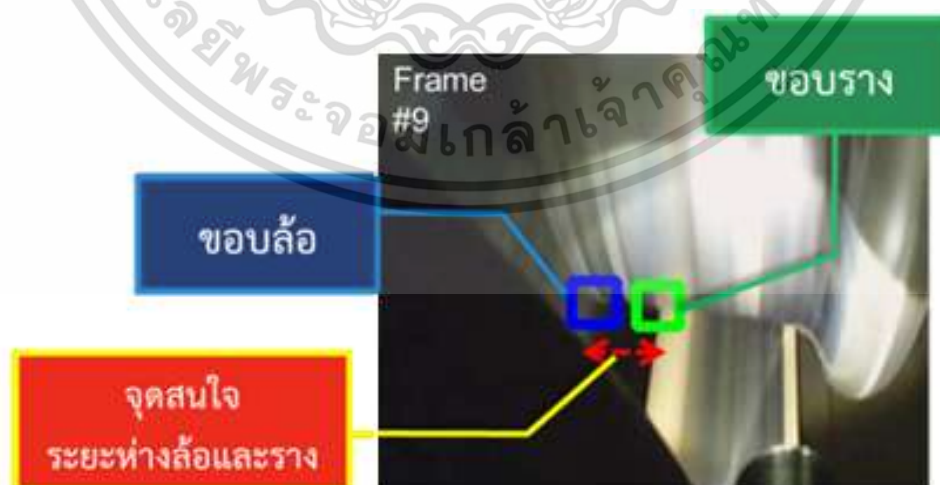
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.2 ผลจากการตรวจจับตำแหน่งและระยะห่างของล้อและรางด้วย Faster R-CNN

เมื่อแบบจำลองปัญญาประดิษฐ์ได้ผ่านกระบวนการเรียนรู้และปรับปรุ้ค่าน้ำหนักที่เหมาะสมแล้ว สามารถทดสอบโมเดลด้วยการนำภาพล้อและรางมาวิเคราะห์ตรวจจับตำแหน่งซึ่งสามารถแสดงได้ดังรูปภาพที่ 4.3



ภาพที่ 4.3 ผลลัพธ์การตรวจหาตำแหน่งของขอบล้อและราง



ภาพที่ 4.4 ผลลัพธ์การวิเคราะห์หาตำแหน่งล้อและรางจากเทคนิคปัญญาประดิษฐ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

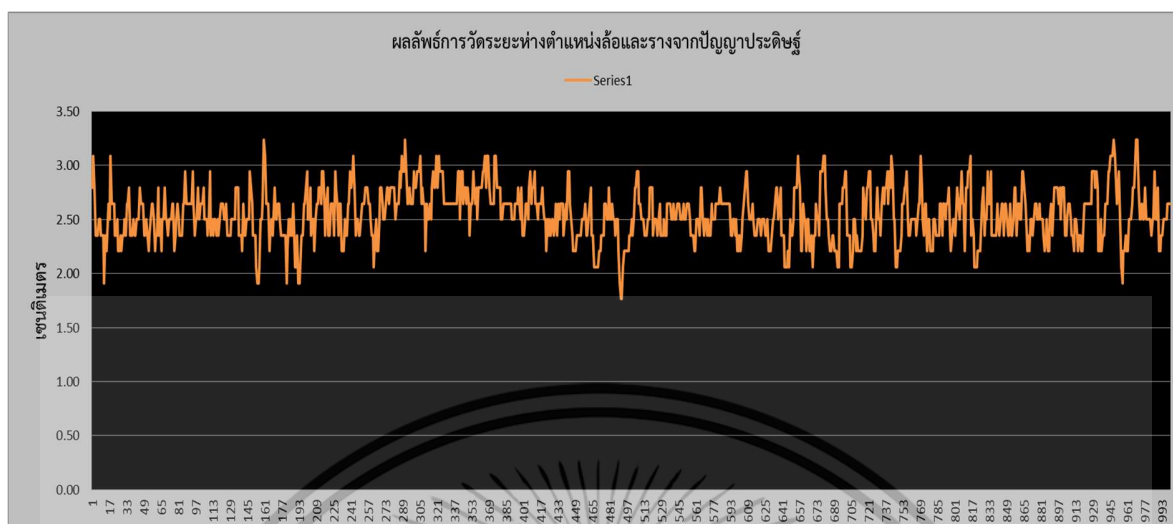
จากภาพที่ 4.4 กรอบสี่เหลี่ยมสีน้ำเงินจะแสดงบริเวณขอบล้อและกรอบสี่เหลี่ยมสีเขียวจะแสดงบริเวณขอบราง เมื่อนำทั้งสองตำแหน่งมาวิเคราะห์ระยะห่างเพื่อตรวจประเมินความปลอดภัยและสภาพทางวิ่งซึ่งระยะปลอดภัยของระยะห่างระหว่างล้อและรางคือ 2.50 เซนติเมตร ซึ่งเป็นระยะที่ล้ออยู่บนรางและบังใบของล้อไม่กระทบรางโดยทดสอบกับสัญญาณภาพวิดีโอจากรางรถไฟฟ้า Airport Rail link ความยาวประมาณ 2 นาที ซึ่งสามารถแบ่งเป็นสัญญาณภาพได้จำนวน 2,699 ภาพ และทำการแบ่งช่วงการวิเคราะห์ออกเป็น 3 ช่วงของสัญญาณรูปภาพ ผลการวัดระยะห่างสามารถแสดงได้ดังภาพที่ 4.5 ภาพที่ 4.6 และภาพที่ 4.7 ตามลำดับ



ภาพที่ 4.5 ผลลัพธ์การวัดระยะห่างของของตำแหน่งล้อและรางช่วง 1-1,000

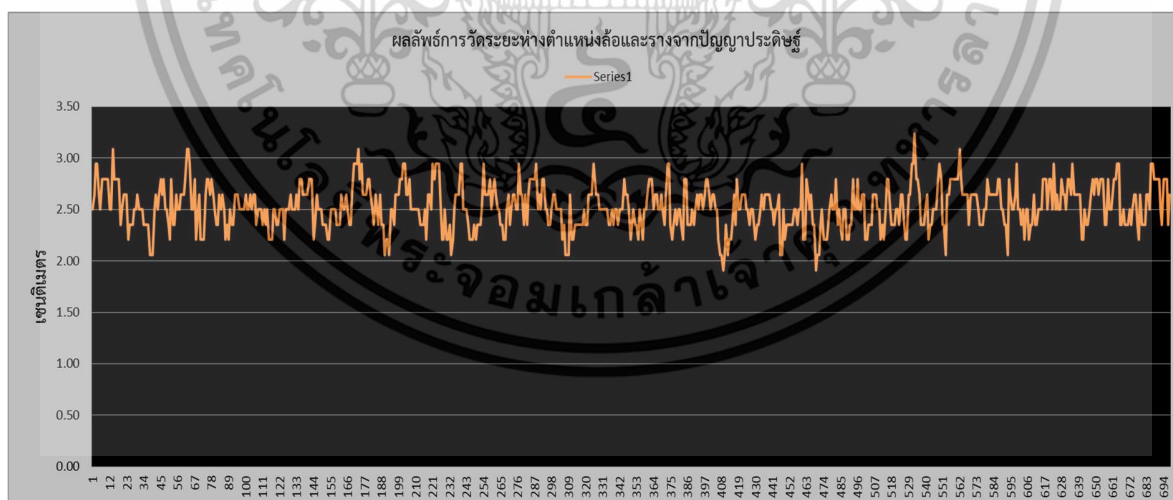
ช่วงแรกภาพที่ 4.5 วิเคราะห์รูปภาพในช่วงรูปภาพที่ 1 ถึงรูปภาพที่ 1,000 พบว่าค่าเฉลี่ยระยะห่างระหว่างล้อและรางอยู่ที่ 2.40 เซนติเมตร ค่าเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.2854 และตรวจพบการเคลื่อนที่ของล้อเข้าใกล้ขอบราง 2 ครั้ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 4.6 ผลลัพท์การวัดระยะห่างของตำแหน่งล้อและรางช่วง 1,001-2000

ช่วงสองภาพที่ 4.6 วิเคราะห์รูปภาพในช่วงรูปภาพที่ 1,001 ถึงรูปภาพที่ 2,000 พบว่าค่าเฉลี่ยระยะห่างระหว่างล้อและรางอยู่ที่ 2.54 เซนติเมตร ค่าเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.2408 และไม่พบการเคลื่อนที่ของล้อเข้าใกล้ขอบราง



ภาพที่ 4.7 ผลลัพท์การวัดระยะห่างของตำแหน่งล้อและรางช่วง 2,001-2,699

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 4.7 เป็นช่วงสุดท้ายของการวิเคราะห์ในช่วงรูปภาพที่ 2,001 ถึงรูปภาพที่ 2,699 พบว่าค่าเฉลี่ยระยะห่างระหว่างล้อและรางอยู่ที่ 2.53 เซนติเมตร ค่าเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.2145 และไม่พบการเคลื่อนที่สองล้อเข้าใกล้ขอบราง

สรุปจากข้อมูลทั้ง 3 ช่วง พบว่าค่าเฉลี่ยในรางช่วงระยะแรกมีค่าต่ำที่สุดโดยมีค่าเฉลี่ยอยู่ที่ 2.40 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.2854 และยังคงตรวจพบการเคลื่อนตัวของล้อเข้าใกล้ขอบรางจำนวน 2 ครั้ง ซึ่งสามารถสรุปรายละเอียดได้ดังตารางที่ 3.2

ตารางที่ 3.2 ตารางสรุปผลการวิเคราะห์

| ลำดับ | ช่วงของรูปภาพ | ค่าเฉลี่ย (cm.) | ค่าส่วนเบี่ยงเบนมาตรฐาน |
|-------|---------------|-----------------|-------------------------|
| 1 | 1-1,000 | 2.40 | 0.2854 |
| 2 | 1,001-2,000 | 2.54 | 0.2408 |
| 3 | 2,001-2,699 | 2.53 | 0.2145 |

จากตารางที่ 3.2 พบว่ารางในช่วงที่ 2 และช่วงที่ 3 มีค่าเฉลี่ยใกล้เคียงกันโดยมีค่าอยู่ที่ 2.54 และ 2.53 ตามลำดับ ค่าเฉลี่ยส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.2408 และ 0.2145 ตามลำดับ

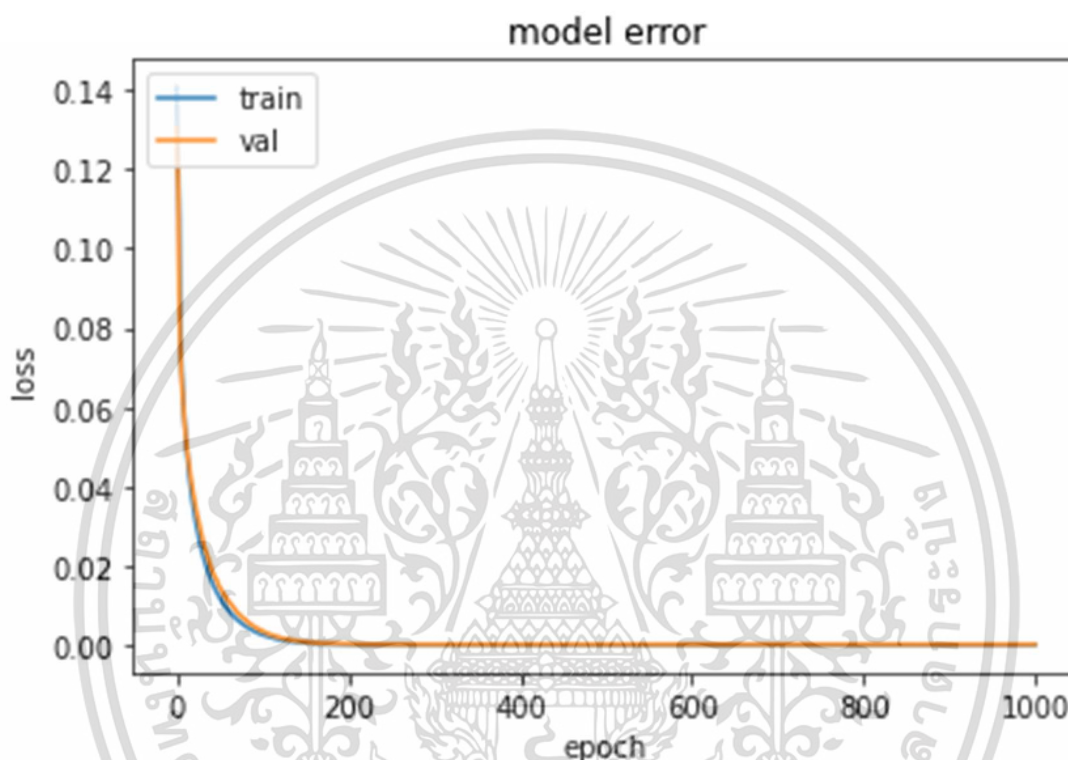
จากภาพที่ 4.3 และ 4.4 จะเห็นได้ว่าจุดที่สนใจคือจุดกึ่งกลางของบริเวณที่สนใจของล้อและราง จากการเก็บข้อมูลภาพด้วยกล้องวิดีโอซึ่งติดอยู่กับล้อรถไฟฟ้าเมื่อรถไฟฟ้ามีการเคลื่อนที่ที่สามารถกำหนดได้ว่าล้ออยู่กับที่ส่วนรางมีการขยับเปลี่ยนตำแหน่งไปมา ด้วยเหตุนี้จึงทำให้สามารถกำหนดจุดของล้อรถไฟฟ้าให้เป็นจุดอ้างอิงแล้วใช้ Faster R-CNN ด้วยโครงสร้างของ VGG-16 วิเคราะห์หาเพียงตำแหน่งของรางเพื่อลดข้อมูลในการวิเคราะห์ อีกทั้งยังทำการลดชั้นโครงสร้างของ VGG-16 ลงเพื่อลดเวลาในการประมวลผลเพื่อหาระยะห่างของล้อและราง

4.2 ผลการตรวจจับอันตรกิริยาระหว่างล้อและรางรถไฟด้วยเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 โดยกำหนดให้ขอบล้อเป็นจุดอ้างอิง

4.2.1 ผลจากการเรียนรู้และการทดสอบของปัญญาประดิษฐ์ด้วยเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 โดยกำหนดให้ขอบล้อเป็นจุดอ้างอิง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การตรวจจับตำแหน่งจากเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 เมื่อนำตำแหน่งที่ได้มาฝึกสอนแบบจำลองเพื่อให้เกิดการเรียนรู้และนำมาทดสอบจะได้กราฟค่าความผิดพลาดของน้ำหนัก ดังภาพที่ 3.16



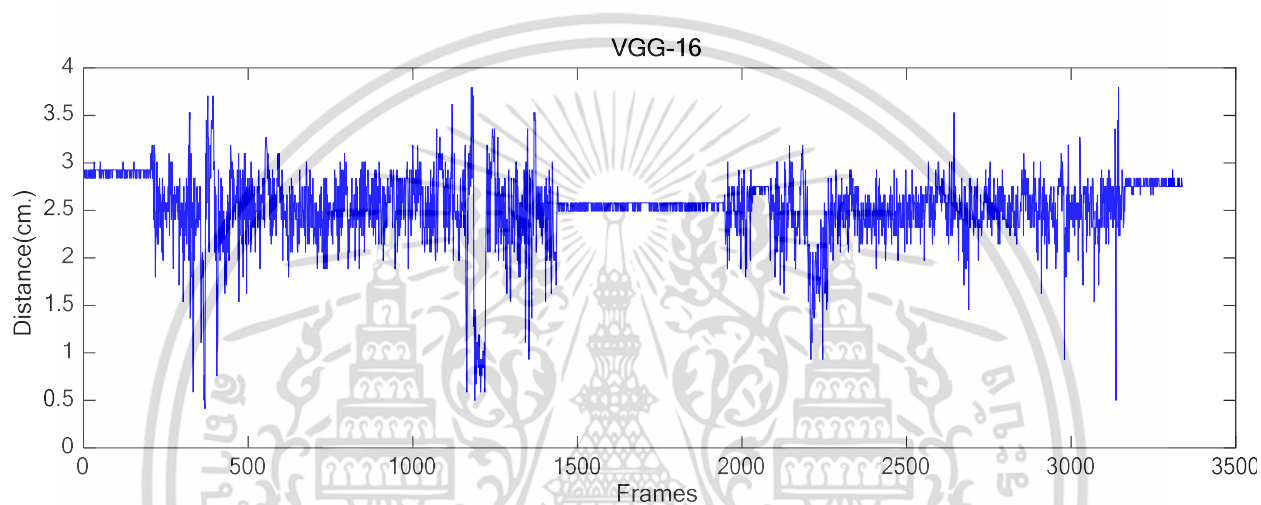
ภาพที่ 4.8 ค่าความผิดพลาดของน้ำหนักที่ทดสอบด้วยเทคนิคปัญญาประดิษฐ์โดยใช้โครงข่ายประสาทเทียมด้วยโครงสร้างของ VGG-16

จากภาพที่ 4.8 แสดงค่าความผิดพลาดที่เข้าใกล้ศูนย์ของน้ำหนักที่ทดสอบด้วยเทคนิคปัญญาประดิษฐ์ด้วยโครงข่ายประสาทเทียมบนโครงสร้างของ VGG-16 ที่ทดสอบจำนวน 1,000 รอบ ได้ค่า loss ของการเรียนรู้เท่ากับ 3.74×10^{-5} และได้ค่า loss ของการทดสอบเท่ากับ 2.78×10^{-5} ได้ค่าความคลาดเคลื่อนของแบบจำลองที่เรียนรู้กับที่ทดสอบจากวิธีการผลต่างกำลังสอง (Mean Square Error) เป็น 5×10^{-5}

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

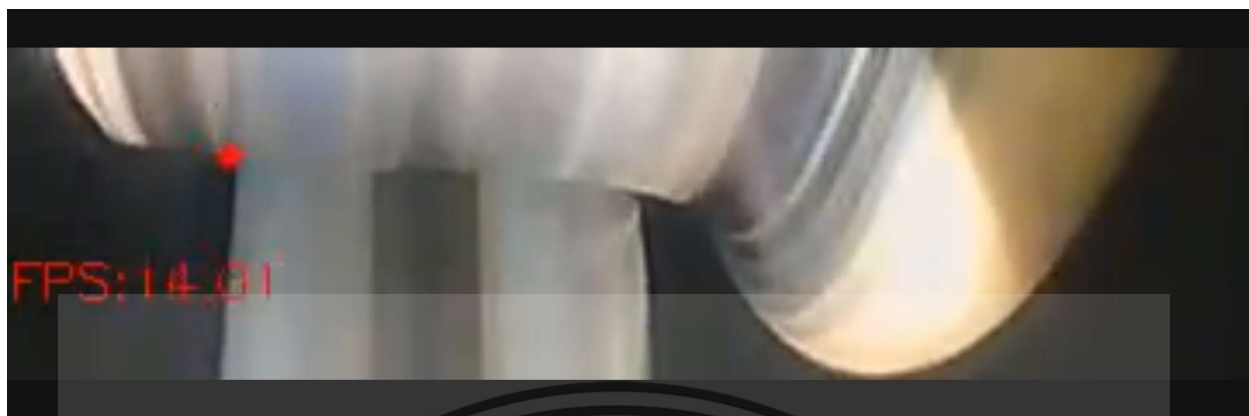
4.2.2 ผลจากการตรวจจับตำแหน่งและระยะห่างของล้อและรางด้วยเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 โดยกำหนดให้ขอบล้อเป็นจุดอ้างอิง

เมื่อนำเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 โดยกำหนดให้ขอบล้อเป็นจุดอ้างอิงมาตรวจจับระยะห่างระหว่างล้อและรางของรถไฟฟ้าจะได้กราฟ ดังภาพที่ 4.9



ภาพที่ 4.9 ผลจากการวัดระยะห่างของล้อกับรางด้วยการตรวจจับตำแหน่งของรางด้วย Faster R-CNN ด้วยโครงสร้างของ VGG-16

จากภาพที่ 4.9 จะเห็นได้ว่าผลจากการวัดระยะห่างของล้อกับรางด้วยการตรวจจับตำแหน่งของรางด้วย Faster R-CNN ด้วยโครงสร้างของ VGG-16 จากวิดีโอ โดยแปลงเป็นภาพจำนวน 3,339 ภาพ จะได้ค่าเฉลี่ยของระยะห่างของล้อกับรางตลอดการเคลื่อนที่เท่ากับ 2.50 เซนติเมตร ดังแสดงผลการตรวจจับตำแหน่งของรางในวิดีโอ โดยได้ตัวอย่างดังภาพที่ 4.10



ภาพที่ 4.10 การตรวจจับตำแหน่งของรางด้วย Faster R-CNN ด้วยโครงสร้างของ VGG-16

จากรูปที่ 4.10 แสดงให้เห็นถึงจุดที่ตรวจจับรางได้และได้ความเร็วในการคำนวณหรือค่ารูปต่อวินาที (Frame per Sec) เป็น 14.01

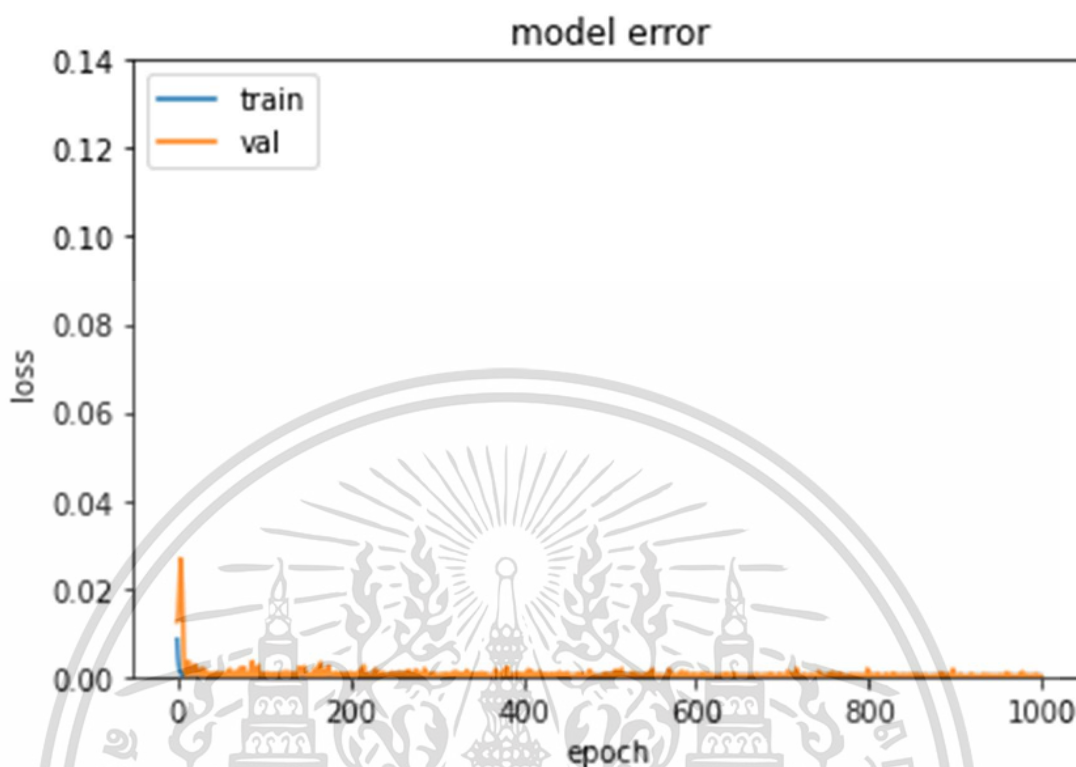
เมื่อนำไปใช้ตรวจจับภาพตำแหน่งของรางทั้งวิดีโอด้วย Faster R-CNN ด้วยโครงสร้างของ VGG-16 จะได้ค่าเฉลี่ยของความเร็วในการคำนวณเป็น 13.74 หรือ 0.073 วินาที

4.3 ผลการตรวจจับอันตรกิริยาระหว่างล้อและรางรถไฟด้วยเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 แบบลดชั้นโดยกำหนดให้ขอบล้อเป็นจุดอ้างอิง

4.3.1 ผลจากการเรียนรู้และการทดสอบของปัญญาประดิษฐ์ด้วยเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 แบบลดชั้นโดยกำหนดให้ขอบล้อเป็นจุดอ้างอิง

การตรวจจับตำแหน่งจากเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 แบบลดชั้นตามภาพที่ 3.10 โดยกำหนดให้ขอบล้อเป็นจุดอ้างอิง เมื่อนำตำแหน่งที่ได้มาฝึกสอนแบบจำลองเพื่อให้เกิดการเรียนรู้และนำมาทดสอบจะได้กราฟค่าความผิดพลาดของน้ำหนัก ดังภาพที่ 4.11

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



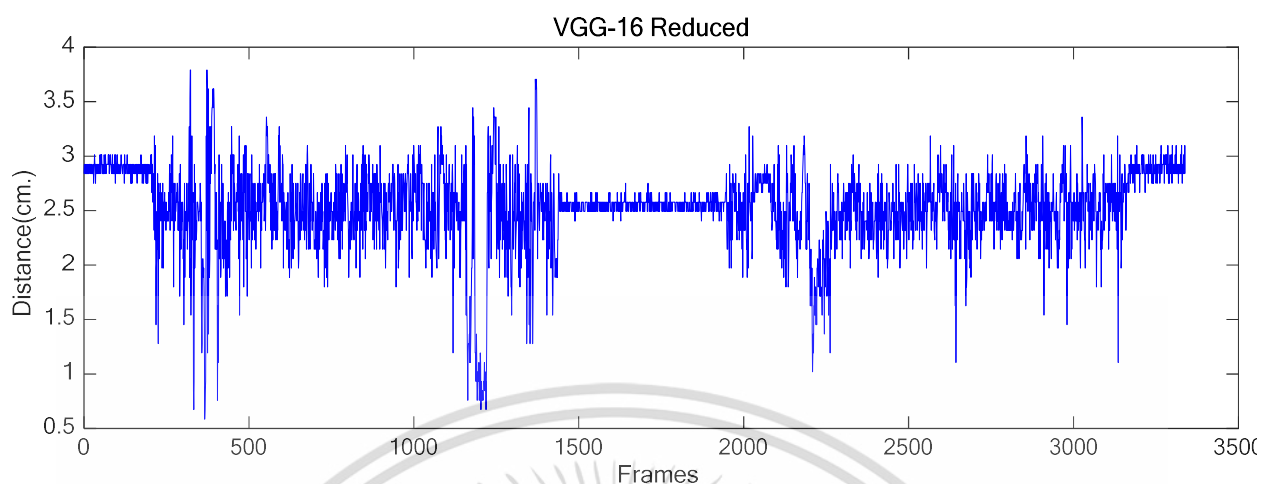
ภาพที่ 4.11 ค่าความผิดพลาดของน้ำหนักที่ทดสอบด้วยเทคนิคปัญญาประดิษฐ์โดยใช้โครงข่ายประสาทเทียมด้วยโครงสร้างของ VGG-16 แบบลดชั้น

จากภาพที่ 4.11 แสดงค่าความผิดพลาดที่เข้าใกล้ศูนย์ของน้ำหนักที่ทดสอบด้วยเทคนิคปัญญาประดิษฐ์โดยใช้โครงข่ายประสาทเทียมด้วยโครงสร้างของ VGG-16 ที่ทดสอบจำนวน 1,000 รอบ ได้ค่า loss ของการเรียนรู้เท่ากับ 1.48×10^{-4} และได้ค่า loss ของการทดสอบเท่ากับ 6.94×10^{-5} ได้ค่าความคลาดเคลื่อนของแบบจำลองที่เรียนรู้กับที่ทดสอบจากวิธีการผลต่างกำลังสอง (Mean Square Error) เป็น 6×10^{-5}

4.3.2 ผลจากการตรวจจับตำแหน่งและระยะห่างของล้อและรางด้วยเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 แบบลดชั้นโดยกำหนดให้ขอบล้อเป็นจุดอ้างอิง

เมื่อนำเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 โดยกำหนดให้ขอบล้อเป็นจุดอ้างอิงมาตรวจจับระยะห่างระหว่างล้อและรางของรถไฟฟ้าจะได้กราฟ ดังภาพที่ 4.12

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 4.12 ผลจากการวัดระยะห่างของล้อกับรางด้วยการตรวจจับตำแหน่งของรางด้วย Faster R-CNN ด้วยโครงสร้างของ VGG-16 แบบลดชั้น

จากภาพที่ 4.12 จะเห็นได้ว่าผลจากการวัดระยะห่างของล้อกับรางด้วยการตรวจจับตำแหน่งของรางด้วย Faster R-CNN ด้วยโครงสร้างของ VGG-16 แบบลดชั้นจากวิดีโอ โดยแปลงเป็นภาพจำนวน 3,339 ภาพ จะได้ค่าเฉลี่ยของระยะห่างของล้อกับรางตลอดการเคลื่อนที่เท่ากับ 2.52 เซนติเมตร ดังแสดงผลการตรวจจับตำแหน่งของรางในวิดีโอ โดยได้ตัวอย่างดังภาพที่ 4.13

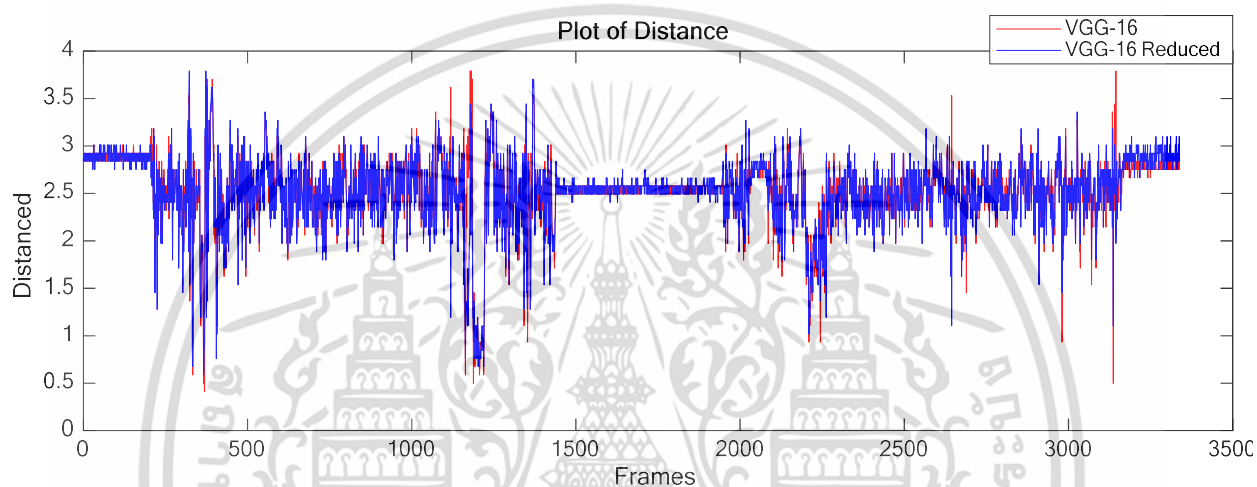


ภาพที่ 4.13 การตรวจจับตำแหน่งของรางด้วย Faster R-CNN โดยใช้โครงข่ายประสาทเทียมด้วยโครงสร้างของ VGG-16 แบบลดชั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4.13 แสดงให้เห็นถึงจุดที่ตรวจจับรางได้และได้ความเร็วในการคำนวณหรือค่ารูปต่อวินาที (Frame per Sec) เป็น 41.54

เมื่อนำไปใช้ตรวจจับภาพตำแหน่งของรางทั้งวิดีโอด้วย Faster R-CNN ด้วยโครงสร้างของ VGG-16 แบบลดชั้นจะได้ค่าเฉลี่ยของความเร็วในการคำนวณเป็น 47.69 หรือ 0.024 วินาที



ภาพที่ 4.14 ระยะห่างระหว่างล้อและรางรถไฟจากการตรวจจับด้วยวิธีการ Faster R-CNN ด้วยโครงสร้างของ VGG-16 กับ VGG-16 แบบลดชั้น

จากภาพที่ 4.14 เส้นสีน้ำเงินแสดงระยะห่างระหว่างล้อและรางรถไฟจากการตรวจจับด้วยวิธีการ Faster R-CNN ด้วยโครงสร้างของ VGG-16 ส่วนเส้นสีแดงแสดงระยะห่างระหว่างล้อและรางรถไฟจากการตรวจจับด้วยวิธีการ Faster R-CNN ขึ้นด้วยโครงสร้างของ VGG-16 แบบลดชั้น เมื่อนำผลของข้อมูลทั้งสองมาหาผลต่างจากวิธีการผลต่างกำลังสอง (Mean Square Error) จากระยะปลอดภัย (2.50 cm.) จะได้ผลต่างของวิธีการที่นำเสนอเท่ากับ 2.65%

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.1 ค่าต่างๆที่วัดได้จากเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 กับโครงสร้างของ VGG-16 แบบลดชั้นโดยกำหนดให้ขอบลือเป็นจุดอ้างอิง

| Faster R-CNN โดยกำหนดให้ ขอบลือเป็น จุดอ้างอิง | loss | | MSE ระหว่าง การเรียนรู้และ การทดสอบ | ความเร็วในการ วิเคราะห์ต่อภาพ (วินาที) | ค่าเฉลี่ยของ ระยะห่างระหว่าง ลือและราง (cm.) |
|---|-----------------------|-----------------------|---|--|---|
| | เรียนรู้ | ทดสอบ | | | |
| VGG-16 | 3.73×10^{-5} | 2.78×10^{-5} | 5×10^{-5} | 0.073 | 2.50 |
| VGG-16 แบบ ลดชั้น | 1.48×10^{-4} | 6.94×10^{-4} | 6×10^{-5} | 0.024 | 2.52 |

จากตารางที่ 4.1 จะเห็นได้ว่าค่า loss ในการเรียนรู้ของเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 มีค่า 3.73×10^{-5} ส่วนโครงสร้างของ VGG-16 แบบลดชั้น มีค่า loss เท่ากับ 1.48×10^{-4} ค่า loss ในการทดสอบของเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 มีค่า 2.78×10^{-5} ส่วนโครงสร้างของ VGG-16 แบบลดชั้น มีค่า loss เท่ากับ 6.94×10^{-4} ค่าความผิดพลาดผลต่างกำลังสองของเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 มีค่า 5×10^{-5} ซึ่งน้อยกว่าค่าความผิดพลาดผลต่างกำลังสองของโครงสร้างของ VGG-16 แบบลดชั้นซึ่งมีค่าเท่ากับ 6×10^{-5} ค่าความเร็วในการวิเคราะห์เทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 แบบลดชั้นใช้เวลา 0.024 วินาทีต่อภาพ ซึ่งรวดเร็วกว่าโครงสร้างของ VGG-16 ที่ใช้เวลา 0.073 วินาทีต่อภาพ ส่วนค่าระยะห่างระหว่างลือและรางที่ตรวจจับด้วยเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 มีความแม่นยำกว่าโครงสร้างของ VGG-16 แบบลดชั้น โดยโครงสร้างของ VGG-16 แบบลดชั้นคลาดเคลื่อนไปจากระยะปลอดภัย (2.5 เซนติเมตร) จากการวัดด้วยวิธีค่าความคลาดเคลื่อนผลต่างกำลังสอง 2.65%

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

ในวิทยานิพนธ์นี้ได้ใช้ระบบวิทัศน์แก้ปัญหาด้านความปลอดภัยของการเคลื่อนที่ของรถไฟฟ้าโดยอาศัยการประมวลผลและการจำแนกรูปภาพด้วยเทคนิคทางปัญญาประดิษฐ์ โดยนำเสนอวิธีการลดชั้นโครงสร้างของ VGG-16 เพื่อลดข้อมูลในการวิเคราะห์ทำให้เวลาในการวิเคราะห์เร็วขึ้นกว่าเดิม จากผลการทดลองที่ได้สามารถสรุปได้ดังนี้

5.1 สรุปผลการวิจัย

5.1.1 สรุปผลการตรวจจับอันตรายระหว่างล้อและรางรถไฟฟ้าด้วยเทคนิค Faster R-CNN

เทคนิคที่ได้นำเสนอเป็นการศึกษาพัฒนาระบบปัญญาประดิษฐ์วิทัศน์สำหรับตรวจวัดบนตัวรถไฟฟ้าให้สามารถตรวจประเมินความปลอดภัยและสภาพทางวิ่งโดยใช้ข้อมูลเส้นทางการวิ่งของรถไฟฟ้า แบบจำลองปัญญาประดิษฐ์สำหรับตรวจวัดสภาพรางและอันตรายระหว่างล้อและรางที่นำเสนอคือ Faster R-CNN ด้วยโครงสร้างของ VGG-16 ซึ่งผลการทดลองได้แสดงประสิทธิภาพในการตรวจประเมินสภาพรางและอันตรายการวัดระยะห่างล้อและราง นอกจากนี้การแสดงผลของปัญญาประดิษฐ์ยังสามารถระบุตำแหน่งของบริเวณขอบล้อและรางได้อย่างชัดเจนและใช้เวลาในการประมวลผลไม่มาก

ผลลัพธ์ในการวัดระยะห่างของตำแหน่งล้อและรางทั้ง 3 ช่วง แสดงให้เห็นถึงประสิทธิภาพการประเมินอันตรายระหว่างล้อและรางในแต่ละช่วง คือ ระยะห่างระหว่างล้อและรางช่วงแรกพบว่ามีค่าเฉลี่ยในช่วงระยะแรกมีค่าต่ำที่สุดโดยมีค่าเฉลี่ยอยู่ที่ 2.40 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.2854 และยังตรวจพบการเคลื่อนตัวของล้อเข้าใกล้ขอบรางจำนวน 2 ครั้ง และรางในช่วงที่ 2 และที่ 3 มีค่าเฉลี่ยใกล้เคียงกันโดยมีค่าอยู่ที่ 2.54 และ 2.53 ตามลำดับ ค่าเฉลี่ยส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.2408 และ 0.2145 ตามลำดับ ผลที่ได้สามารถนำไปพัฒนาและประยุกต์ใช้ตรวจประเมินความปลอดภัยและสภาพทางวิ่งให้เป็นแบบตามเวลาจริงในอนาคตได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.1.2 สรุปผลการตรวจจับอันตรายระหว่างล้อและรางรถไฟด้วยเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16

เทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 สำหรับตรวจวัดตำแหน่งของรางและวิเคราะห์ผลการวัดระยะห่างของล้อและราง พบว่าสามารถตรวจประเมินความปลอดภัยและสภาพทางวิ่งได้รวดเร็วโดยดูได้จากค่าเฉลี่ยของการประมวลผลของเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 ใช้เวลาการวิเคราะห์ผลต่อภาพ 0.073 วินาที โดยเทคนิคที่นำเสนอนำมาประมวลผลบนวิดีโอที่บันทึกภาพการเคลื่อนที่ของรถไฟฟ้าจะเห็นได้ว่าวิธีการ Faster R-CNN ด้วยโครงสร้างของ VGG-16 สามารถตรวจวัดตำแหน่งของรางและวิเคราะห์ผลการวัดระยะห่างของล้อและรางได้รวดเร็วและตามเวลาจริง

5.1.3 สรุปผลการตรวจจับอันตรายระหว่างล้อและรางรถไฟด้วยเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 แบบลดขั้น

เทคนิคที่ได้นำเสนอเป็นการศึกษาการลดขั้นของเทคนิค Faster R-CNN ขึ้นด้วยโครงสร้างของ VGG-16 เพื่อลดข้อมูลในการวิเคราะห์และลดเวลาในการคำนวณประมวลผลภาพเปรียบเทียบกับ VGG-16 แบบดั้งเดิม พบว่าการลดขั้นของโครงสร้างของ VGG-16 ทำให้สามารถประมวลผลการตรวจจับและวัดระยะห่างของล้อและรางได้รวดเร็ว ตามเวลาจริงและแม่นยำจึงสามารถตรวจประเมินความปลอดภัยและสภาพทางวิ่งได้รวดเร็วกว่า โดยดูได้จากค่าเฉลี่ยของการประมวลผลของโครงสร้างของ VGG-16 แบบดั้งเดิมซึ่งใช้เวลาการวิเคราะห์ผลต่อภาพ 0.073 วินาที ส่วนโครงสร้างของ VGG-16 แบบลดขั้นใช้เวลาการวิเคราะห์ผลต่อภาพเพียง 0.024 วินาที โดยนำเทคนิคที่นำเสนอมาประมวลผลบนวิดีโอที่บันทึกภาพการเคลื่อนที่ของรถไฟฟ้า จะเห็นได้ว่าวิธีการที่นำเสนอให้ผลใกล้เคียงกับเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 เนื่องจากมีค่าความคลาดเคลื่อนของผลการวัดระยะห่างระหว่างล้อและรางเพียง 2.65% แต่วิธีการที่นำเสนอใช้เวลาในการประมวลผลน้อยกว่า

5.2 ข้อเสนอแนะ

ระบบที่นำเสนอเป็นประมวลผลจากรูปภาพด้วยการคำนวณทางคณิตศาสตร์ด้วยคอมพิวเตอร์ซึ่งเทคนิคทางปัญญาประดิษฐ์เป็นวิธีการประมวลผลจากรูปภาพเพื่อวิเคราะห์หาจุดที่สนใจ จากผลการทดลองที่ได้สามารถสรุปได้ ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 5.2.1 กล้องที่ใช้ต้องมีความละเอียดเพียงพอในการรับภาพและรูปภาพที่นำมาใช้ในการวิเคราะห์ต้องมีความสว่างเพียงพอเพื่อให้เกิดความแตกต่างของวัตถุในภาพ
- 5.2.2 เนื่องจากวิธีการที่นำเสนอเป็นการประมวลผลด้วยการคำนวณทางคณิตศาสตร์ คอมพิวเตอร์ที่ใช้ในการประมวลผลจะต้องมีความสามารถเพียงพอในการรับข้อมูลเข้ามาประมวลผลในปริมาณมาก
- 5.2.3 การติดตั้งกล้องเพื่อบันทึกการเคลื่อนที่ของล้อและรางต้องแน่นหนาเนื่องจากบริเวณล้อรถไฟฟ้ามหานครที่มีความสั่นสะเทือน หากติดตั้งไม่ดีอาจทำให้กล้องเกิดความเสียหาย
- 5.2.4 ผลที่ได้จากการตรวจจับด้วยตรวจจับอันตรกิริยาระหว่างล้อและรางรถไฟฟ้ามหานครด้วยเทคนิค Faster R-CNN ด้วยโครงสร้างของ VGG-16 สามารถนำไปแปลงเวฟเล็ตเพื่อนำสัญญาณเวฟเล็ตที่ได้วิเคราะห์ประเมินสภาพการวิ่งของรถไฟฟ้ามหานครได้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

- [1] สำนักงานสภาพัฒนาการเศรษฐกิจและสังคมแห่งชาติ. “แผนพัฒนาเศรษฐกิจและสังคมแห่งชาติ ฉบับที่สิบสอง (พ.ศ. 2560-2564)” [Online] Available : https://www.nesdc.go.th/ewt_dl_link.php?nid=6422
- [2] Saraswat, S. and Yadava, G.S. "An overview on reliability, availability, maintainability and supportability (RAMS) engineering" **International Journal of Quality & Reliability Management**, vol. 25, no. 3, Jan. 2008.
- [3] FRA (Federal of Railway Association). “**Rail Safety Fact Sheet.**” [Online] Available : <https://railroads.dot.gov/>
- [4] Xavier G., Vishal M. and Rama C. “Deep Multitask Learning for Railway Track Inspection” **IEEE Trans. On Intelligent Transportation Systems**, vol. 18, no. 1, Jan. 2017.
- [5] Junbo L., Yaping H., Qi Z., Mei T. and Shengchun W. “Learning Visual Similarity for Inspecting Defective Railway Fasteners” **IEEE Sensors Journal**, vol. 19, no. 16, Apr. 2019.
- [6] Orhan Y., Mehmet K. and Erhan A. “A vision based diagnosis approach for multi rail surface faults using fuzzy classification in railways” **Proc. IEEE-ICCSE. Antalya, Turkey**, Oct. 2017. Pp. 713-718
- [7] Vonani R., Nico S. and Yskandar H. “3D Rail Modelling and Measurement for Rail Profile Condition Assessment” **Proc. IEEE-AFRICON. Cape Town, South Africa**, Nov. 2017. Pp. 1522-1527
- [8] Canan T., Orhan Y., Mehmet K. and Erhan A. “A real time interface for vision inspection of rail components and surface in railways” **Proc. IEEE-IDAP. Malatya, Turkey**, Nov. 2017.
- [9] Krishna M., Madhav P., Giridhar M. K. and Amit S. P. “Automated Traffic Monitoring System Using Computer Vision” **Proc. IEEE-ICTBIG. Indore, India**, Apr. 2017.
- [10] Bhaskar B., Clarence G., Shubham B. and Jignesh S. “A Self-Driving Car Implementation using Computer Vision for Detection and Navigation” **Proc. IEEE-ICCS. Madurai, India**, Apr. 2020.
- [11] Judith S. E, Deepa M. D. K, Ansy A., Anumary M V, Daine M., Fincy J. and Jinesh KJ, “Advanced License Plate Recognition System for Car Parking” **Proc. IEEE**

- ICES. Coimbarore, India, Nov. 2014.**
- [12] Harikrishnan J, Arya S., and Aravind A. “Vision-Face Recognition Attendance Monitoring System for Surveillance using Deep Learning Technology and Computer Vision” **Proc. IEEE-VITECoN. Vellore, Indis, Nov. 2019.**
- [13] S. M Amir K., Md. Nazmul I. S. and Moajjem H. C. “Smart Entrance System Using Computer Vision at Corporate Environment” **Proc. IEEE-ICECCT. Coimbatore, India, Oct. 2019.**
- [14] Ulagamuthalvi, J B Janet F and Abinaya D. “An Efficient Object Detection Model Using Convolution Neural Networks” **Proc. IEEE-ICOEI. Tirunelveli, India, Oct. 2019.**
- [15] Wei Z, Shihao W, Sophanyouly T., Jingzhou C. and Yuntao Q. “DECONV R-CNN for Small Object Detection on Remote Sensing Images” **Proc. IEEE-IGRSS. Valencia, Spain, Nov. 2018.**
- [16] Yan Z. and Ying H. “Video Smoke Detection Based on Convolution Neural Network” **Proc. IEEE-ICCTEC. Dalian, China, Aug. 2019.**
- [17] Sameeruddin Mohammed. “Analytics Vidhya: Image Convolution from Scratch” [Online] Available : <https://medium.com/analytics-vidhya/image-convolution-from-scratch-d99bf639c32a>. 2019.
- [18] Mailerdaimon. Stack overflow. “Why Gauss filter chooses such values” [Online] Available : <https://stackoverflow.com/questions/20041391/why-gauss-filter-chooses-such-values/20044618>. 2013.
- [19] Damir D., Emir S. and Amira S. “Performance of some image processing algorithms in TensorFlow” **Proc. IEEE-IWSSIP. Maribor, Slovenia, Jun. 2018.**
- [20] Melissa Y., Brain M., Natalia V. and Avid R. “Design of a Vehicle Driver Drowsiness Detection System through Image Processing using Matlab” **Proc. IEEE-CONCAPAN XXXVIII, San Salvador, El Salvador, Nov. 2018.**
- [21] Mogeab A. A. M., Abdul A. AL-Y. and Abdu G. “An Automatic Nuclei Cells Counting Approach Using Effective Image Processing Methods” **Proc. IEEE-ICSIP, Wuxi, Chaina, Jul. 2019.**
- [22] Wentao X. “Research on Fire Detection and Image Information Processing System Based on Image Processing” **Proc. IEEE-ICAACI, Ottawa, ON, Canada, Sep. 2020.**

[23] Tuples Edu. “What is an Artificial” [Online]. Available :

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
<https://becominghuman.ai/what-is-an-artificial-neuron-8b2e421ce42e>. 2017
 ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามเผยแพร่ต่อผู้อื่น และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [24] Krishnab. “**Neural network: Why can NNs have identical structure in the hidden units and still learn?**” [Online] Available : <https://stats.stackexchange.com/questions/413712/neural-network-why-can-nns-have-identical-structure-in-the-hidden-units-and-sti>. 2019
- [25] Edvin Beqari. “**A Very Basic Introduction to Feed-Forward Neural Networks**” [Online] Available : <https://dzone.com/articles/the-very-basic-introduction-to-feed-forward-neural>. 2018.
- [26] Super Data Science Team. “**Super Data Science: Convolutional Neural Networks**” [Online] Available : <https://www.superdatascience.com/blogs/convolutional-neural-networks-cnn-step-4-full-connection>. 2018
- [27] “**Course Website. CS231n Convolutional Neural Networks for Visual Recognition**” [Online] Available : http://cs231n.github.io/assets/nn1/neuron_model.jpeg
- [28] Savaram Ravindra. “**What is Image Recognition and why is it Used**” [Online] Available : <https://www.kdnuggets.com/2017/08/convolutional-neural-networks-image-recognition.html>
- [29] Jeremy Jordan. “**Data Science : Convolutional Neural Networks**” [Online] Available : <https://www.jeremyjordan.me/convolutional-neural-networks/>. 2017.
- [30] Austin G. Walters. “**Convolutional Neural Networks (CNN) to Classify Sentences**” [Online] Available : <https://austingwalters.com/convolutional-neural-networks-cnn-to-classify-sentences/image-filter-matrix-cnn/>
- [31] Arden Dertat. “**Towards data science : Applied Deep Learning-Part 4 : Convolutional Neural Networks**” [Online] Available : <https://towardsdatascience.com/applied-deep-learning-part-4-convolutional-neural-networks-584bc134c1e2>
- [32] Ujjwalkarn. “**The data science blog : An Intuitive Explanation of Convolutional Neural Networks**” [Online] Available : <https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/>
- [33] Ratthanapoom N. “**Convolutional Neural Network — Backbone of Computer Vision**” [Online] Available : <https://medium.com/botnoi-classroom/convolutional-neural-network-backbone-of-computer-vision-723ef3077219>

- [34] Afshine Amidi and Shervine Amidi. “**Convolutional Neural Networks cheatsheet**” [Online] Available : <https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-convolutional-neural-networks>
- [35] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” **Proc. IEEE-CCVPR. Columbus, OH, USA. Sep. 2014. Pp. 580-587.**
- [36] R. Girshick, “Fast R-CNN,” **Proc. IEEE-ICCV. Santiago, Chile, 2015. Pp. 1440-1448, 2015.**
- [37] Ren S., He K., Girshick R. and Sun J. 2017. “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks” **IEEE Transaction on pattern analysis and machine intelligence**, Vol. 39, No. 6. Jun. 2017
- [38] Bin liu, Wencang Zhao and Qiaoqiao, “Study of object detection based on Faster R-CNN”, **Proc. Chinese Automation Congress (CAC), Jinan, China. Oct. 2017. Pp. 6233-6236.**
- [39] Yu liu, “An Improve Faster R-CNN for object Detection”, **Proc. IEEE-ISCID Hangzhou, China. Dec. 2018. Pp. 119-123**
- [40] Li Y. T. and Guo J. I. A “VGG-16 Based Faster RCNN Model for PCB Error Inspection in Industrial AOI Applications.” **Proc. IEEE-ICCE-TW. Taichung, Taiwan. Aug. 2018.**
- [41] Kasthurirangan Gopalakrishnan, Siddhartha K. Khaitan, Alok Choudhary and Ankit Agrawal, “Deep Convolutional Neural Networks with transfer learning for computer vision-based data-driven pavement distress detection”, **ScienceDirect : Construction and Building Materials**, Vol. 157, 30 December 2017, pp. 322-330.
- [42] Fei-Fei Li, Justin Johnson amd Serena Yeung. “**Lecture 12: Detection and Segmentation**” [Online] Available : http://cs231n.stanford.edu/slides/2019/cs231n_2019_lecture12.pdf

ภาคผนวก
ผลงานวิจัยที่ได้รับการตีพิมพ์เผยแพร่

- [1] Anakkapon Saenthon, Somyot Kiatwanidvilai and Suparat Pumyoy,
“Development of Artificial-intelligence Vision System for Measurement On-
service Train for Train-track Inspection”, Sensor and Materials, Vol. 32, No. 2,
pp. 549-561, 2020



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Development of Artificial-intelligence Vision System for Measurement On-service Train for Train-track Inspection

Anakkapon Saenthon,¹ Somyot Kiatwanidvilai,² and Suparat Pumyoy^{2*}

¹Department of Manufacturing System Engineering, King Mongkut's Institute of Technology Ladkrabang, 1, Soi Chalongkrung 1, Chalongkrung Street, Ladkrabang, Bangkok, 10520

²Department of Electrical Engineering, King Mongkut's Institute of Technology Ladkrabang, 1, Soi Chalongkrung 1, Chalongkrung Street, Ladkrabang, Bangkok, 10520

(Received June 29, 2019; accepted January 24, 2020)

Keywords: rail, track, artificial intelligence, vision system inspection

We studied and inspected the evaluation of safety during the operation of a rail system. In addition, in order to find patterns in the interaction between wheels and rails, the interaction is investigated using an automatic vision system to analyze the patterns that occur. The model of a faster region with a convolutional neural network (R-CNN) can be used to design artificial intelligence (AI) models with vision systems for detecting abnormal operations that occur, causing rail accidents, and to monitor the measurement of tracks and interactions between wheels and tracks in real time.

1. Introduction

Thailand is developing a basic rail infrastructure with a length of more than 4500 km, including dual-rail trains and city buses, and mass transit trains. The length of the railway (Railway Track) is up to a distance of more than 7000 km. However, the system is required to be ready to provide service under the safety standards so the rail infrastructure must be reliable. According to the data of rail accidents collected by the Federal Railway Association (FRA), 45 % of rail accidents were caused by irregularities of railways and various devices on the trains. Risk prevention can be achieved by effective measurement and maintenance planning. The rails currently in use can be analyzed using images. Oh *et al.* presented a vision-based tracking system with a train signal system to stop trains when there is an emergency.⁽¹⁾ Rikhotso *et al.* presented 3D rail modelling and measurement for rail profile condition assessment.⁽²⁾ Yaman *et al.* presented a method for the detection of abnormalities of railway tracks from photos taken by cameras.⁽³⁾ Liu *et al.* used image processing to detect the damage of the rail surface.⁽⁴⁾ In addition, Girshick *et al.* presented the region with convolutional neural network (R-CNN) method together with photographs from which objects of interest are detected.⁽⁵⁾ Later, Girshick presented fast R-CNN, which takes less time than R-CNN.⁽⁶⁾ Therefore, we are interested in applying interdisciplinary knowledge, such as knowledge of electrical and electronic engineering, computer engineering, mechanical engineering, rail transport engineering, and

*Corresponding author: e-mail: suparat.pu@kmitl.ac.th

<https://doi.org/10.18494/SAM.2020.2529>

other related fields such as vision systems, artificial intelligence (AI), and Internet of Things (IoT), to monitor interactions between wheels and tracks in real time. Nassu and Ukai applied a basic vision system to evaluate the changes in train tracks.⁽⁷⁾ The main purpose of this research is to develop a tool for measuring the abnormalities of a rail system. Such a tool will be beneficial for the safe use of a rail system by both rail system providers and rail system users. It will also help to reduce the possibility of accidents involving rail systems, which would affect the credibility of the country as well.

2. Theory

In this paper, we present a vision system and AI based on image processing.

2.1 Vision system

The operation of the vision system involves image processing or computer calculation to obtain the required quality and quantity of information. There are various important steps to derating sharp images to eliminate interference in images. Objects in photographs are isolated to obtain an image of the object for analysis to acquire quantitative data such as size, shape, and the direction of movement of objects in the photograph. We can subsequently use these quantitative data for analysis and to create systems for use in various applications, such as the fingerprint recognition system, the quality inspection system of products in the manufacturing process of industrial plants, the grade separation or inspection system of the quality of agricultural crops, the automatic postal code reading system that sorts letters from images of the postal code on the envelope, the data collection system of vehicles that enter and exit a building using photographs of the vehicle license plate, the road traffic monitoring system in which the number of cars on the road is counted in a photograph taken by closed-circuit television (CCTV), the face recognition system for surveillance of terrorists in buildings, at landmarks, or immigration areas, the advanced license plate recognition system for car parking,⁽⁸⁾ the image processing technique in a real-time vision system for automatic weeding strategy,⁽⁹⁾ the vision system in the weaving loom industry,⁽¹⁰⁾ and the vision system for monitoring intermodal freight trains.⁽¹¹⁾ These systems require much image processing that must be repeated. It is impossible for human beings to analyze this data because it would be excessively time-consuming. Therefore, computers play an important role in these functions in place of humans.

2.2 AI

AI is a field of computer science concerning how to make computers have capabilities like those of humans or imitate human behavior, especially the ability to think by themselves or have intelligence. Such humanlike intelligence is created by the computer and therefore called AI. People differ depending on how they use intelligence considering behavior in the environment or the results of AI (the ability of AI to think). Hence, the definition of AI is based on the abilities of humans divided into the following four groups:

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- acting humanly,
- thinking humanly,
- thinking rationally, and
- acting rationally.

The neuron model structure used in this research is a faster R-CNN,⁽⁶⁾ which has the ability to detect the position of an object in an image and recognize the object type. For the Faster R-CNN, the main idea is to bring the selective search, which is the area of interest containing the target object, included in the neuron model in the same structure. However, Liu improved the Faster R-CNN for object detection.⁽¹²⁾ In addition, Cai *et al.* presented street object detection based on the Faster R-CNN.⁽¹³⁾

As shown in Fig. 1, this model has the following three main components.

- 1) the base that performs the extraction feature or considers the distinctive features of the image,
- 2) the region proposal network (RPN) area that separates “expecting or maybe” of an object from the feature map, and
- 3) the classification section that processes the feature map and region obtained from the RPN by region of interest (RoI) pooling to determine which areas of the image contain objects.

VGG-16 is a CNN model proposed by Simonyan and Zisserman.⁽¹⁶⁾ The VGG-16 structure shown in Fig. 2 can support colorized black and white images. The first layer can support pictures with a size of 224×224 pixels and RGB images. The work structure is divided into five convolution steps. Then, the final convolution is the fully connected classifier. Details are shown in Fig. 3.

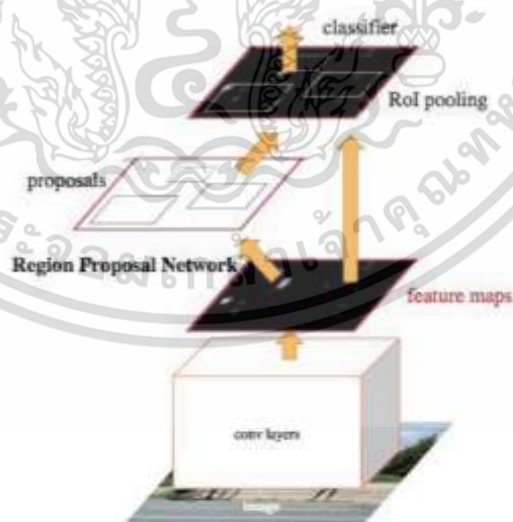


Fig. 1. (Color online) The faster R-CNN is a single, unified network for object detection. The RPN module serves as the ‘attention’ of this unified network.⁽¹⁴⁾

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

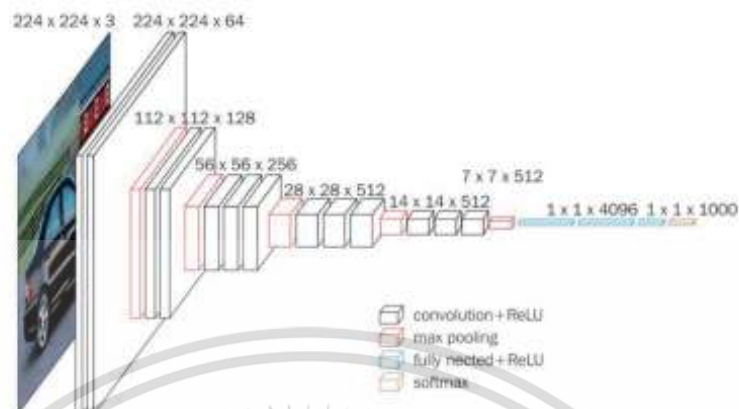


Fig. 2. (Color online) Structure of VGG-16 used as the base of the model.⁽¹⁵⁾

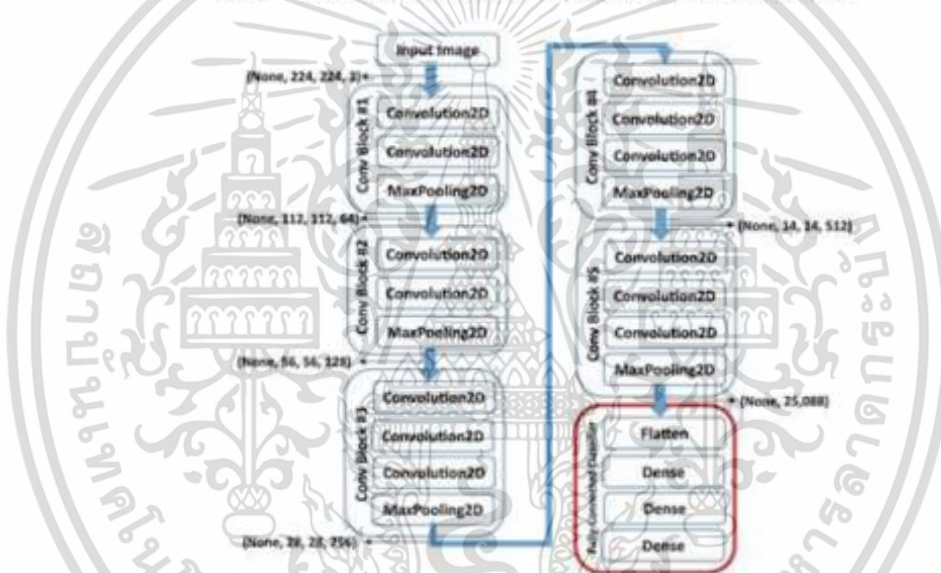


Fig. 3. (Color online) Details of each VGG-16 layer.

3. Experimental Results

3.1 Determining the scope of consideration for preparing images

Figure 4 shows an example of images from the video camera installed under the carriage. In data collection, a camera is used to record data as a video of the normal state of movement for the purpose of quality assessment analysis of safety and railway conditions. When using the image signal to consider the nature of the risk, the characteristics of wheels and rails are as shown in Fig. 5.

From Fig. 5, we found that the wheel rim was very close to the edge of the rail, causing the wheel rim to hit the track. This collision may result in damage to both the rail and the wheel. To determine the area of interest, the process converted the video to 32400 images suitable for studying and memorizing by AI (Fig. 6).

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

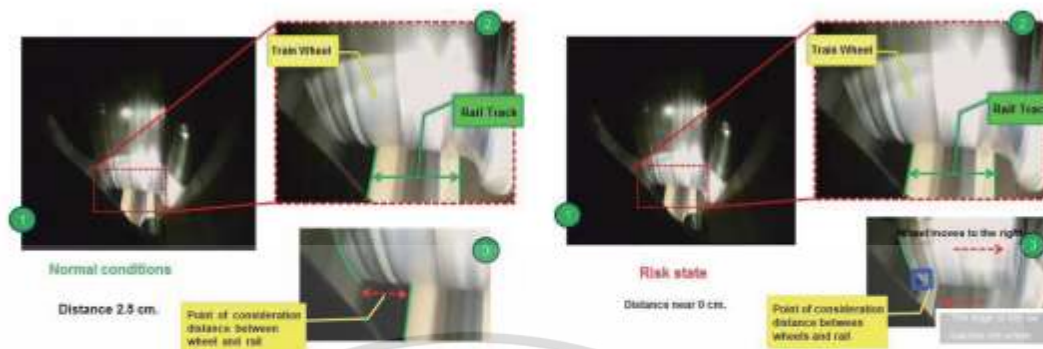


Fig. 4. (Color online) Images of wheel and rail area to determine the point of consideration.

Fig. 5. (Color online) Images showing wheels and the risk of damage to wheels.

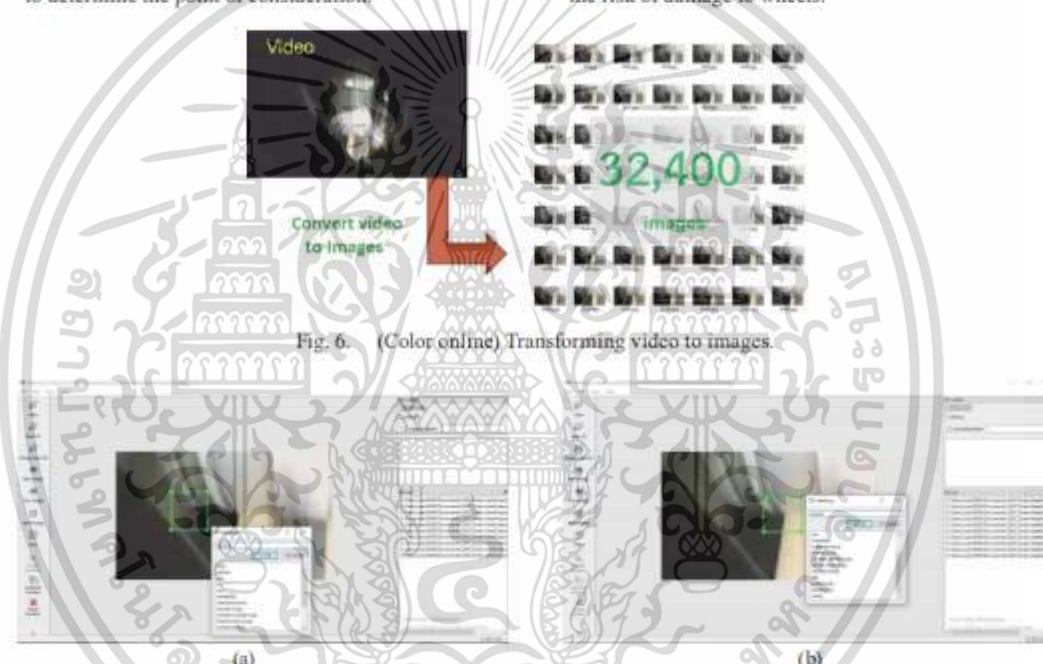


Fig. 6. (Color online) Transforming video to images.

Fig. 7. (Color online) Selection of the wheel rim and the rail rim area. (a) Rim of the wheel. (b) Rim of the rail.

On the basis of the total amount of data, the data were separated into the following two main parts: information used for learning by the AI and information used to test AI capabilities. The procedure specifies two areas, namely the wheel rim and the edge of the track, as shown in Figs. 7(a) and 7(b), respectively.

3.2 Detailed structure of AI model

Figure 8 shows the detailed structure of the Faster R-CNN. First, import the image into VGG-16 to find the hidden layer in a total of five repetitions. Then, perform convolution to find the strength of the object in the image and apply the rectified linear unit (ReLU) to eliminate

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

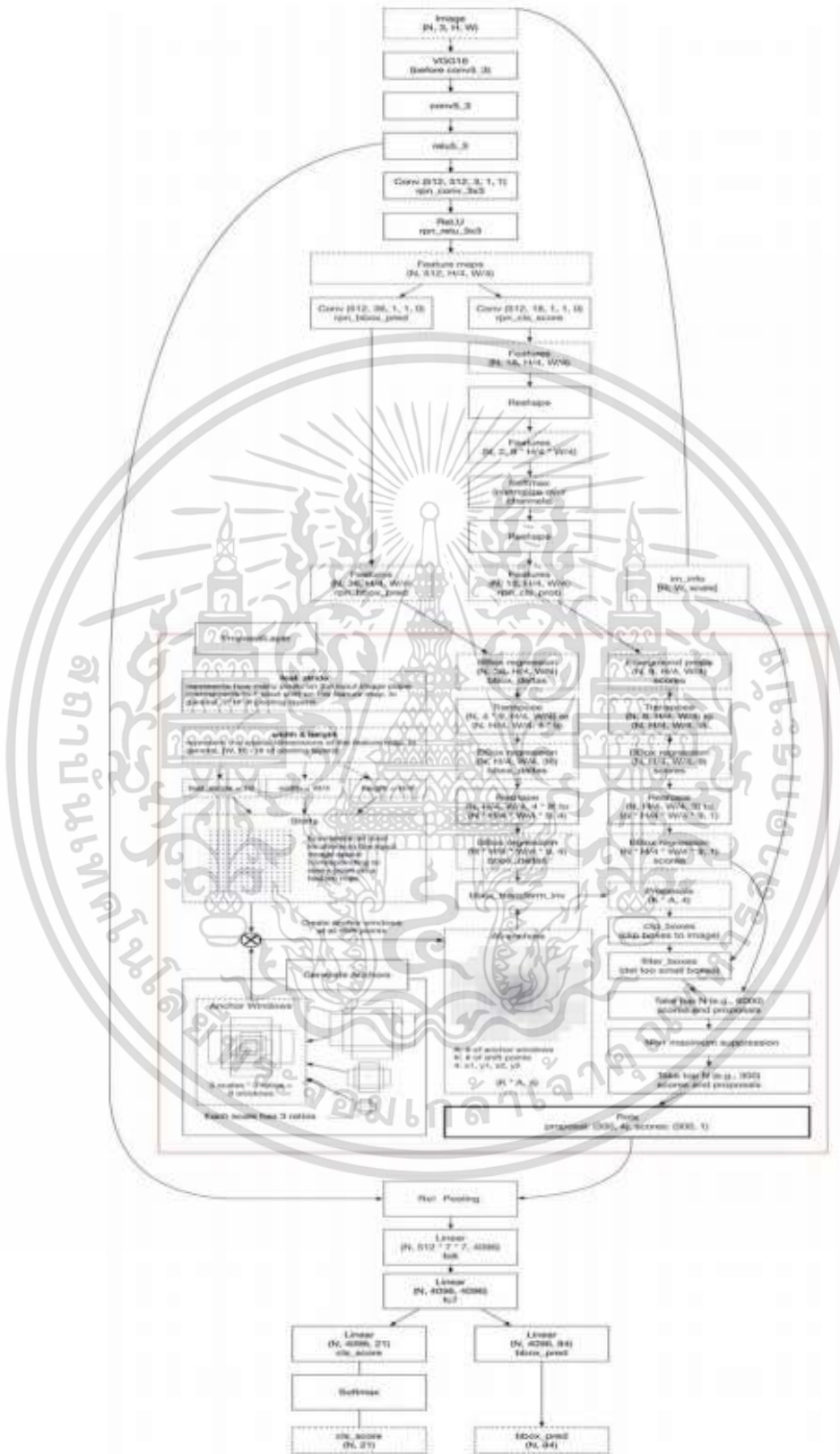


Fig. 8. (Color online) Structure of the Faster R-CNN.⁽⁵⁾

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

linearity. Next, import data into the RPN to separate the parts that are expected to be objects. Then, obtain images of many objects using only the feature map from the final convolution layer and determine the coordinates of the objects on the image. Use the object image to find the proposal layer to obtain the RoI. Then, carry out RoI pooling to obtain the fixed size vector. After that, send the information to the classification section.

3.3 Learning to remember to analyze the positions of wheels and rails

Figure 9 shows the determination of the area on the rim and the edge of the rail using the Labellmg program. A total of 81 images were used, 50 images for learning and 31 images to test the AI results.

The prepared image data for locating the object of interest is in the form of a Label file. The file information consists of the path, the location of the file, X position, Y position, width, and height. This information is used to verify the accuracy of the AI results.

Figure 10 shows the TensorBoard graph. The TensorBoard graph is a visualization tool that appropriates with the AI training process. The TensorBoard graph in Fig. 10 represents the

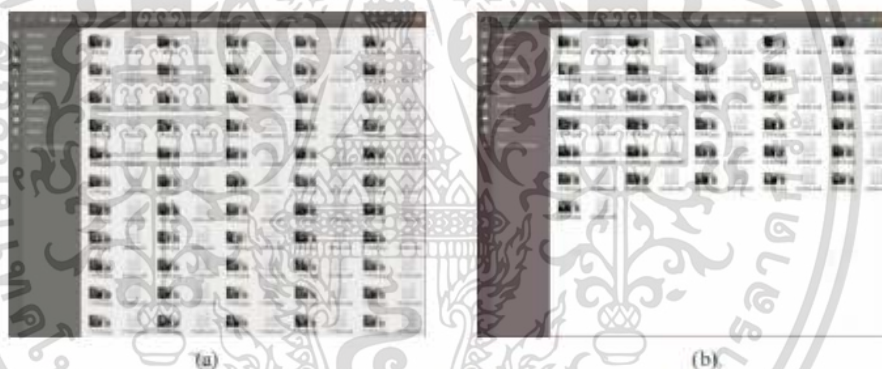


Fig. 9. (Color online) Images and labels for (a) training and (b) testing.



Fig. 10. (Color online) TensorBoard graph.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Cross Entropy graph, which is the result of training the model. The results are obtained from the differential value between the distributed probability regarding the actual data and the distributed probability regarding the utilized model as

$$H(p, q) = -\sum p \log q . \quad (1)$$

Then, when the differential value between the distributed probability regarding the actual data and the distributed probability regarding the utilized mode is improved to be as low as possible by utilizing the loss function, this particular loss function becomes the differential value between the output and the target, which is utilized to estimate and update the weight of the model.

Typically, in the training process, the learning data is fed to AI, then AI analyzes the results compared with the labels. If the results are largely different from the labels, the weight will be constantly updated until the results become close to the labels. The differential values of AI recognition results are called loss values. The loss values can be divided into two sections, which are the loss from the locating process and the loss from the classification process. After running the test for 10000 repetitions, the results show that the loss values are 0.02 and 0.04. Figure 11 delineates the graph of the results from AI, clearly showing that the graph is gradually converged into the horizontal axis.

3.4 Learning results for position by AI

The AI model is composed of the learning process and the adjustment of the appropriate weight. The test results for the analysis of the positions of wheels and rails are shown in Fig. 12.

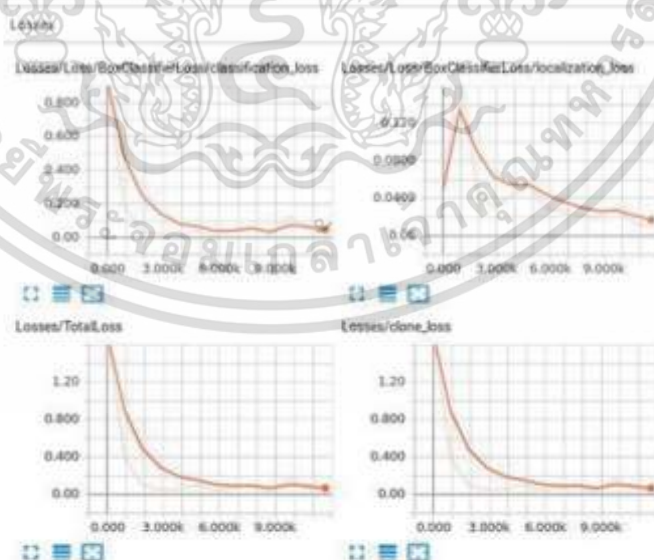


Fig. 11. (Color online) Resulting loss determined by TensorBoard running for 10000 repetitions.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

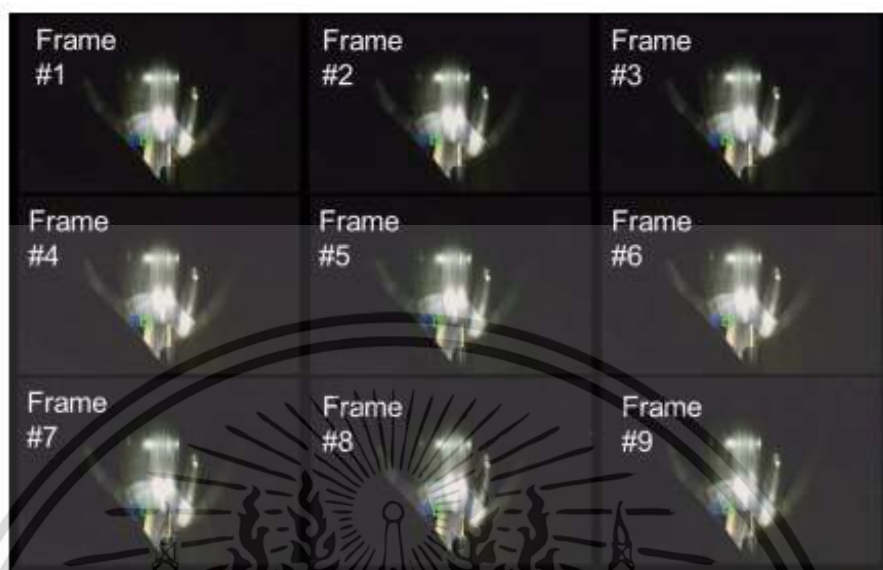


Fig. 12. (Color online) Results of wheel and rail edge detection.

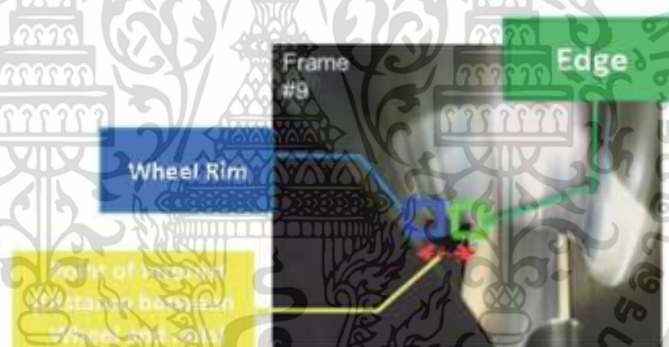


Fig. 13. (Color online) Results of analysis of wheel and rail positions by AI.

In Fig. 13, the blue square shows the wheel rim and the green square shows the trough. The distance between the two areas was analyzed to assess safety and railway conditions by testing using video signals of about 90 s length from the train tracks.

The data of 2699 image signals for rail position are divided into three ranges, as shown in Figs. 14–16.

The analysis results of the first range of signals from the 1st to the 1000th image are shown in Fig. 14. It was found that the average distance between wheels and rails was 2.40 cm. The standard deviation was 0.2854. Two wheels moving near the edge of the rail were detected two times.

The analysis results of the second range of signals from the 1001st to the 2000th image are shown in Fig. 15. It was found that the average distance between wheels and rails was 2.54 cm. The standard deviation was 0.2408. No wheel moving near the edge of the rail was detected.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

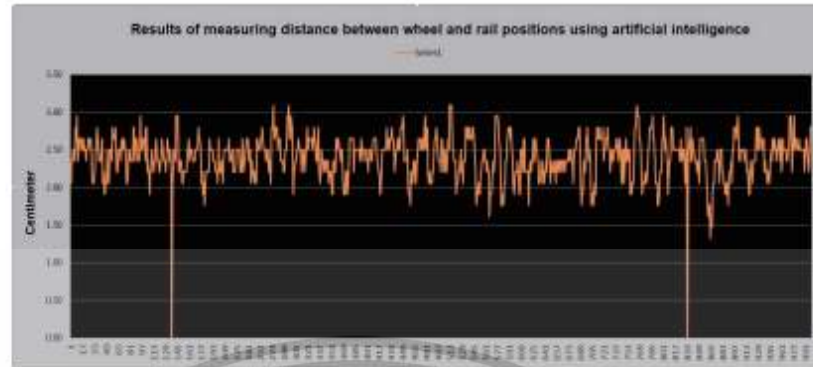


Fig. 14. (Color online) Results of measuring the distance between the wheel and rail positions for image signals from 1 to 1000.



Fig. 15. (Color online) Results of measuring the distance between the wheel and rail positions for image signals from 1001 to 2000.

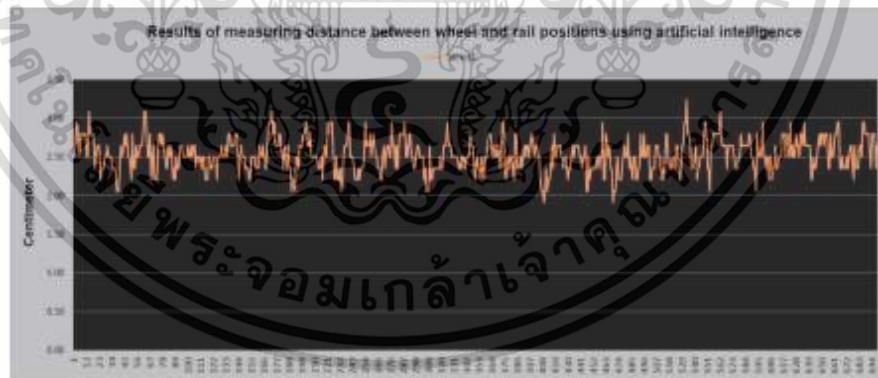


Fig. 16. (Color online) Results of measuring the distance between the wheel and rail for image signals from 2001 to 2699.

The analysis results of the last range of signals from the 2001st to the 2699th image are shown in Fig. 16. It was found that the average distance between wheels and rails was 2.53 cm. The standard deviation was 0.2145. Again, no wheel moving near the edge of the rail was detected.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Table 1
Mean and standard deviation of analyses results.

| Range | Image signals | Mean distance (cm) | Standard deviation |
|-------|---------------|--------------------|--------------------|
| 1 | 1–1000 | 2.40 | 0.2854 |
| 2 | 1001–2000 | 2.54 | 0.2408 |
| 3 | 2001–2699 | 2.53 | 0.2145 |

From the analysis of these three ranges of data, it was found that the average distance in the first trough was the lowest with an average of 2.40 cm and a standard deviation of 0.2854, and the movement of the wheel near the edge of the draft was detected two times. These results are shown in Table 1.

Table 1 shows that the distances in the second and third ranges have similar averages of 2.54 and 2.53 cm, with average standard deviations of 0.2408 and 0.2145, respectively.

4. Conclusions

The developed AI systems for the visual measurement of the train can be used to examine the safety and railway conditions used in the route information of the train. The AI model for measuring the images of the rail condition and the interaction between wheels and rails shown here is the Faster R-CNN (VGG-16). The results show the effectiveness of the model in determining the rail condition and the efficiency of the assessment of the interaction between wheels and rails by distance measurement. The result of measuring the distance between the wheel and rail positions from three consecutive sets of images shows the efficiency of the evaluation of the interaction between wheels and rails in the three periods. The average distance in the first range was the lowest at 2.40 cm with the standard deviation of 0.2854. The movement of the wheel near the edge of the rail was detected 2 times. In the second and third ranges of image analysis, the average values were similar, being 2.54 and 2.53 cm, with the average standard deviations of 0.2408 and 0.2145 cm, respectively. The results of this experiment can be further analyzed and used to evaluate the safety and conditions of rails in the future.

Acknowledgments

We express special gratitude to the Development Research Center, Test and Transfer Technology of High Speed Train of Thailand for providing research funding.

References

- 1 S. Oh, S. Park, and E. Joung: Proc. 2009 Int. Conf. Information and Multimedia Technology (IEEE, Jeju Island, 2019) 59–63. <https://doi.org/10.1109/ICIMT.2009.34>
- 2 V. Rikhotse, N. Steyn, and Y. Hamam: Proc. 2017 IEEE AFRICON (IEEE, Cape Town, 2017) 1522–1527. <https://doi.org/10.1109/AFRICON.2017.8095708>
- 3 O. Yaman, M. Karakose, and E. Akin: Proc. 2017 Int. Conf. Computer Science and Engineering (UBMK, Antalya, 2017) 713–718. <https://doi.org/10.1109/UBMK.2017.8093511>

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 4 Z. Liu, W. Wang, X. Zhang, and W. Jia: Proc. 2010 2nd Int. Asia Conf. Informatics in Control, Automation and Robotics (CAR, Wuhan, 2010) 472–475. <https://doi.org/10.1109/CAR.2010.5456793>
- 5 R. Girshick, J. Donahue, T. Darrell, and J. Malik: Proc. 2014 IEEE Conf. Computer Vision and Pattern Recognition (IEEE, 2014) 580–587. <https://doi.org/10.1109/CVPR.2014.81>
- 6 R. Girshick: Proc. 2015 IEEE Int. Conf. Computer Vision (ICCV, Santiago, 2015) 1440–1448. <https://doi.org/10.1109/ICCV.2015.169>
- 7 B. T. Nassu and M. Ukai: Proc. IEEE Trans. Intelligent Transportation Systems (IEEE, 2012) 1763–1771. <https://doi.org/10.1109/TITS.2012.2204052>
- 8 E. J. Sen, K. D. M. Dixon, A. Anto, M. V. Anumary, D. Meehal, F. Jose, and K. J. Jinesh: Proc. 2014 Int. Conf. Embedded Systems (ICES, Coimbatore, 2014) 162–165. <https://doi.org/10.1109/EmbeddedSys.2014.6953109>
- 9 M. M. Mustafa, A. Hussain, K. H. Ghazali, and S. Riyadi: Proc. 2007 IEEE Int. Symp. Signal Processing and Information Technology (IEEE, Giza, 2007) 632–635. <https://doi.org/10.1109/ISSPIT.2007.4458197>
- 10 G. Vladimir, B. Dmitriy, T. N. Win, and I. Evgen: 2017 IEEE Conf. Russian Young Researchers in Electrical and Electronic Engineering (EIConRus, St. Petersburg, 2017) 669–671. <https://doi.org/10.1109/EIConRus.2017.7910645>
- 11 A. Kumar, N. Ahuja, J. M. Hart, U. K. Visesh, P. J. Narayanan, and C. V. Jawahar: Proc. 2007 IEEE Workshop Applications of Computer Vision (WACV '07, Austin, TX, 2007). <https://doi.org/10.1109/WACV.2007.9>
- 12 Y. Liu: Proc. 2018 11th Int. Symp. Computational Intelligence and Design (ISCID, Hangzhou, 2018) 119–123. <https://doi.org/10.1109/ISCID.2018.10128>
- 13 W. Cai, J. Li, Z. Xie, T. Zhao, and K. Lu: Proc. 2018 37th Chinese Control Conference (CCC) 9500–9503. <https://doi.org/10.23919/ICCC.2018.8482613>
- 14 S. Ren, K. He, R. Girshick, and J. Sun: Proc. IEEE Trans. Pattern Analysis and Machine Intelligence (IEEE, 2017) 1137–1149. <https://doi.org/10.1109/TPAMI.2016.2577031>
- 15 Y. T. Li and J. L. Guo: Proc. 2018 IEEE Int. Conf. Consumer Electronics-Taiwan (ICCE-TW). <https://doi.org/10.1109/ICCE-China.2018.8448674>
- 16 K. Simonyan and A. Zisserman: Proc. 2015 6th Int. Conf. Learning Representations (ICLR 2018). <https://arxiv.org/pdf/1409.1556.pdf>

About the Authors



Anakkapon Saenthon received his B.S. and M.S. degrees from Narasuan University, Thailand, in 1996 and 2008, respectively, and his Ph.D. degree from King Mongkut's Institute of Technology Ladkrabang, Thailand, in 2011. Since 2015, he has been an assistant professor at King Mongkut's Institute of Technology Ladkrabang, Thailand. His research interests are in vision systems, image processing, and manufacturing systems. (anakkapon.sa@kmitl.ac.th)



Somyot Kiatwanidvilai received his B.S. and M.S. degrees from King Mongkut's Institute of Technology Ladkrabang, Thailand, in 1996 and 2000, respectively, and his Ph.D. degree from Asian Institute of Technology (AIT) in 2005. From 2006 to 2012, he was an assistant professor at King Mongkut's Institute of Technology Ladkrabang, Thailand. Since 2013, he has been an associate professor at King Mongkut's Institute of Technology Ladkrabang, Thailand. His research interests are in power electronics, machine drives and systems, electrical power and energy, control systems, and artificial intelligence. (somyot.ki@kmitl.ac.th)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



Suparat Pomyoy received her B.S. and M.S. degrees from King Mongkut's Institute of Technology Ladkrabang, Thailand, in 2009 and 2014, respectively. Her research interests are in power electronics, genetic algorithms, and vision systems. (suparat.pu@kmitl.ac.th)



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [2] Suparat Pumyoy, Somyot Kiatwanidvilai, Harn Sison, Poom Konghuayrob and Natchaya Yamnak, "Signal analysis by Discrete Wavelet Transform for Lateral Movement Inspection Data of Railway Rapid Transit from Machine Vision", The 5th National and The 1st International Conference on Informatics, Agriculture, Management Business administration, Engineering Sciences and Technology, pp. 99-105, 28-29 May 2020 (Online Conference).



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้




The 5th National Conference
The 1st International Conference

**on Informatics, Agriculture, Management
 Business Administration, Engineering
 Sciences and Technology**

28 – 29 May 2020
online Conference

King Mongkut's Institute of Technology Ladkrabang
 Prince of Chumphon Campus



<http://www.pcc.kmitl.ac.th/iambest>



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



Signal Analysis by Discrete Wavelet Transform for Lateral Movement Inspection Data of Railway Rapid Transit from Machine Vision

Suparat Pumyoy^{1*}, Somyot Kaitwanidvilai¹, Harn Sison¹, Poom Konghuayrob¹ and Natchaya Yamnak¹

¹ Department of Electrical Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand

*Corresponding author: Suparat Pumyoy, e-mail address: suparat.pu@kmitl.ac.th

Abstract

This paper presents the signal processing by discrete wavelet transform for hunting oscillation inspection of system which utilized the image processing data which is captured from high resolution speed camera. The obtained is the data which been through the image processing and data transformation system. The results are position of the flange and the railway. Since the analysis by the Fourier transform are unable to pinpoint the exact position where the noise signal is occurred, however, the discrete wavelet transform is able to classify and inform the change of the signal in real time.

Keywords: Inspection systems, Machine vision, Hunting frequency, Discrete wavelet transform, Railway

1. Introduction

Railway transportation is developed to become the energy efficiency and sustainable system which tends to be more utilized in the future. Railway transportation has begun in 1900s, in the South East Asia (SEA), Thailand is the first country that brought the public transportation innovation since the era of King Rama V (SomdetchPhraParamindrMaha Chulalongkorn). The railway transportation has been continuously developed until the light rail train and electric drive train has been utilized.

In order to inspect the electrical train regarding the research [1], which utilize the high resolution speed camera to consider the lateral movement of the tram from the image of the flange and the rail which is analyzed by the image processing technique to define the distance between the wheel and the rail. The obtained results are correct and accurate. Considering the results data, the data presents high quantity of the noise signal which requires the comparison of a lot of images to identify the source of the noise signal.

Fourier transform (FT) is widely use for evaluating the various pictures and images [2]. It had been utilized in more than 100 years until now, and was aware of the frequency-domain analysis is often superior to the time-domain analysis; frequency-domain analysis is not only easy to detect, but also simple analysis of multiple complicated signals. However, using more accurate digital approach, namely, Discrete Fourier Transform (DFT) spectral analysis, before the advent of the fast Fourier transform (FFT), this technique is impractical, due to the Discrete Fourier Transform (DFT) calculation is too computationally intensive. In 1965 Cooley Tukey developed the first fast Fourier transform (FFT) algorithm [3]. A. Sokolowski and T. Pardela proposed Fourier transform (FT) in melanoma image classification [8]. Lifang Dong, et al. proposed processing of images in gas discharge by spatial correlation function and fast Fourier transform (FFT) [4]. Muruganand Shanmugam and Balamurugan Rajamanickam proposed application of FFT in laser speckle image pattern recognition technique for the metrological measurement [5].

Wavelet applied the mathematics theory which can capture data into frequency components frames, and then analyze each component with a precise matched to figure scale. These technique role higher

performance over the normal Fourier methods in analysis physical situations where the signals contains discontinuities and sharp spikes [6]. K. Ramesh and Y. Mallikharjuna Rao present the high voltage transformer preventive by using wavelets [7]. Wavelet transform consists of two main types, first is Continuous Wavelet Transformation (CWT) and another is Discrete Wavelet Transformation (DWT) [6]. Moreover, in 2015, RasmirBajric, et al. [8] classified the abnormality regarding the fault within gear set connecting to the electric generating turbine utilized the discrete wavelet transform. The result revealed that it is able to classify the characteristics of the each wavelet that occurred correctly when compare to the abnormalities of each one. Furthermore, there is some research such [9] that utilized the discrete wavelet transform to classify the wavelet from the abnormalities of the pump [10] and also the research [10] that utilizes the discrete wavelet transform to detect the emotion via the Electroencephalography (EEG) technique which able to detect the sadness and enjoyment more accurately when compare to the conventional.

This research presents the use of Discrete wavelet transform to analyze and classify the abnormality in real time, this allow the problem consideration to proceed faster and less complex since each wavelet that occurred has specific characteristic and can be detected at the specific time. This technique has more advantages compared to the Fast Fourier Transform which cannot be able to specify the location of fault and the time of that fault occur in real time.



Figure 1 Images showing wheels and rail.

2. Theory

2.1 Fast Fourier Transform Theory

Fast Fourier transform (FFT) is the method that transform the data in the time domain to frequency domain which is developed from Discrete Fourier Transform (DFT) to become faster. Since the storage data is large, it is necessary to utilize the fast Fourier transform (FFT) and reduce the calculation of the N data set from $2N^2$ into $2N \log N$. The transform equations can be described by the following:

$$X(k) = \sum_{j=0}^{N-1} x(j) \omega_N^{j(k-i)} \quad (1)$$

$$x(j) = (1/N) \sum_{k=0}^{N-1} X(k) \omega_N^{-j(k-i)} \quad (2)$$

Where $\omega_N^{j(k-i)} = e^{(2\pi j/N)(k-i)}$ is the N^{th} root of unity.



By utilizing the Cooley–Tukey FFT algorithm with the bits reversion arrangement and transform N to $N/2$, the new Fast Fourier transform equation is presents as the following:

$$\sum_{n=0}^{N-1} a_n e^{-2\pi i n k / N} = \sum_{n=0}^{N/2-1} a_{2n} e^{-2\pi i (2n) k / N} + \sum_{n=0}^{N/2-1} a_{2n+1} e^{-2\pi i (2n+1) k / N} \tag{3}$$

$$= \sum_{n=0}^{N/2-1} a_n^{even} e^{-2\pi i n k (N/2)} + e^{-2\pi i k / N} \sum_{n=0}^{N/2-1} a_n^{odd} e^{-2\pi i n k (N/2)} \tag{4}$$

2.2 Discrete Wavelet Transform [3]

First Wavelet transform is the signals $f(t)$ is similar to the wavelet superposition of different position of signal, different scales, while the wavelet coefficients $W_f(a, b)$ shows the overall contribution of this location, the scales wavelet mother function by scaling and translation, is established, if a wavelet mother function $\Psi(t)$, scaling and translation factors, respectively a, b , then the continuous wavelet function of signal $f(t)$ as described below:

$$W_f(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \Psi\left(\frac{t-b}{a}\right) f(t) dt \tag{5}$$

$$\Psi_t = \frac{1}{\sqrt{a}} \Psi\left(\frac{t}{a}\right) \tag{6}$$

Wavelet mother function $\Psi(t)$, was expanded a times by utilizing the wavelet function. The purpose including \sqrt{a} guarantee $\Psi_a(t)$ is the square on the space $L^2(\mathbb{R})$ namely:

$$\|\Psi_a(t)\|^2 = \int \Psi_a(t)^2 dt < \infty \tag{7}$$

By equation (5):

$$W_f(a, b) = f(b) * \Psi_a(-b) \tag{8}$$

Respectively find the Fourier transform of the signal $f(b)$ and wavelet function $\Psi_a(-b)$ are $F(w), \Psi_a(w)$, then the Fourier transform of wavelet transform $W_f(a, w)$, by the convolution theorem available:

$$W_f(a, w) = F(w) * \Psi_a^*(w) \tag{9}$$

Secondly, in order to ensure the creative of inverse wavelet transform, the wavelet mother function, the Fourier transform $\Psi_a(w)$ should be met as:

$$C_\Psi = \int_0^\infty \frac{|\Psi(w)|}{w} dw < \infty \tag{10}$$

To apply the equation (10) Fourier transform of wavelet function, this $\Psi_a(w)$ at frequencies w equal to band-pass filter component, you can deduce, for different scales $\Psi_a(w), (\Psi_a(0) = 0)$. Therefore, the wavelet transform can be seen as $ess0$ must be 0 or $\Psi_a(0) = 0$, the status of the signal can indicate the wavelet transform is essentially the signal through verity bandwidth of band-pass filter data signals. To facilitate computer data processing, usually scaling and translation factors a, b are convert to be discrete,



taking $a=2n$, $b=m$. Then the wavelet transform equation (5) transform into a discrete binary wavelet transform:

$$W_f(n, m) = 2^{\frac{n}{2}} \int_{-\infty}^{\infty} \Psi\left(\frac{\tau-m}{2^n}\right) f(\tau) d\tau \tag{11}$$

If $\Psi(x)$ is taken to the form (5) non-compactly supported orthogonal wavelet, $f(x) \in L^2(\mathbb{R})$, then:

$$W_f(\lambda, x) = \left[\frac{\sqrt{\lambda}}{(2\pi)} \right] \left[\int_{-\frac{\pi}{\lambda}}^{\frac{\pi}{\lambda}} F(w) \exp(-jxw) dw + \int_{\frac{\pi}{\lambda}}^{2\pi/\lambda} F(w) \exp(-jxw) dw \right] \tag{12}$$

Proof of the theorem is clearly available, it is a little bundle.

For the discrete signal $f(n)$, $n=0,1,\dots,N-1$; $N=2^i$, $1 \in \mathbb{Z}$, set the discrete Fourier transform of $f(n)$ as follows:

$$F(k) = \left(\frac{1}{N}\right) \sum_{n=0}^{N-1} f(n) \exp\left(-\frac{j2\pi kn}{N}\right), k = 0, 1, \dots, N-1 \tag{13}$$

The discrete FT of $F(k)$ as follows:

$$F(k) = (1/N) \sum_{n=0}^{N-1} f(n) \exp\left(-\frac{j2\pi kn}{N}\right), m = 0, 1, \dots, N-1 \tag{14}$$

$$F_i(k) = 2^i f(k + N_i), N_i = 2^{i-1} N, i = 1, 2, \dots \tag{15}$$

From equation (15) available:

$$\begin{aligned} W_f(2^{-i}, 2^{-i}m) &= \left[\frac{2}{2^{-i}\sqrt{2^{-i}}} \right] FF_i(m), FF_i(m) \\ &= \left(\frac{1}{N_i}\right) \sum_{k=0}^{N_i-1} F_i(k) \exp\left(-\frac{j2\pi km}{N_i}\right) \end{aligned} \tag{16}$$

Discrete wavelet transform is the signal processing method which is developed from Fourier transform of discrete wavelet transform which return the result in term of the pattern and the time following the Mother wavelet. In this research, two filter banks are utilized for the transformation. The frequency will be classified into two components, first is the component that pass through the low pass filter called "Approximation" the second one is the component pass through the high pass filter called "Details" via the Multiple-level decomposition and Down sampling, this procedure will reduce the size of the data to be the half and also able to reverses transform in case of the data that less than the starting data. The next transformation is to classify the obtained low frequency signal again in the tree mapping as delineate in the Fig. 2.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

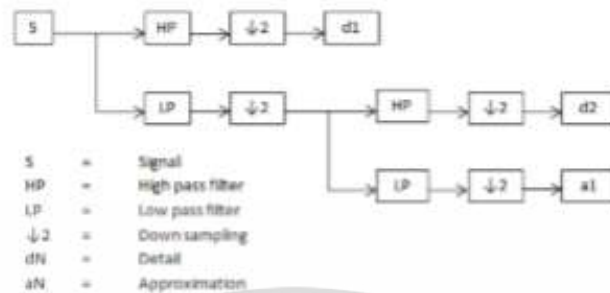


Figure 2 The tree chart shows the conversion into multiple layers.

3. Experimental Results

The researcher brought the signal from the research [1] to analyze as shown in Fig. 3 which are the distance between the wheel and the train rail compare to the number of frame per second (fps) captured. When process the Fourier transform, the results are in the form of frequency which cannot be able to determine the time of these frames was captured.

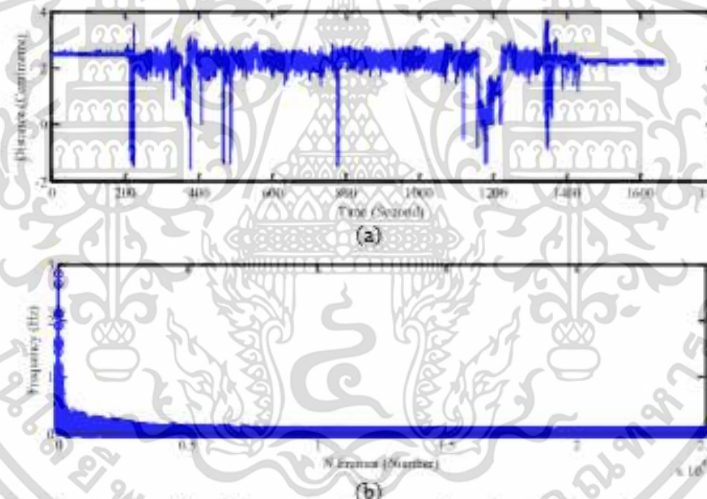


Figure 3 (a) Signal image from Signal Processing (b) Image from Fourier Transform

Then, the previous data is utilized in the Discrete wavelet transform by using the mother wavelet with Daubechies4 (db4) type to decompose the high frequency and low frequency via Decomposition level 8. All wavelet frequency is separated to be low frequency wavelet and high frequency wavelet. The low frequency wavelet is utilized to be primary signal of the system that is not included the high frequency wavelet into it.

Coefficient graph a8 shows the primary signal which clearly seen in Fig.4 that the high frequency signal is filtered form the original. Form the result, the primary signal a8 can be identified as the normal

lateral movement condition. High wavelet frequency coefficient d1 identifies the rail shuffle of the train which occurs in many positions in the graph result, while coefficient d4 consists of many small signals which are much more difference from the rail shuffle condition when compare to the signal of d1, d2 and d3 causing by the lateral sway movement and the sway signal disappear at the end of graph before new wavelet occur on the coefficient of the low frequency graph d6, d7 and d8. This compared result can be used for describing that the cause of this signal is from speed reducing of train before the train stops at the platform.

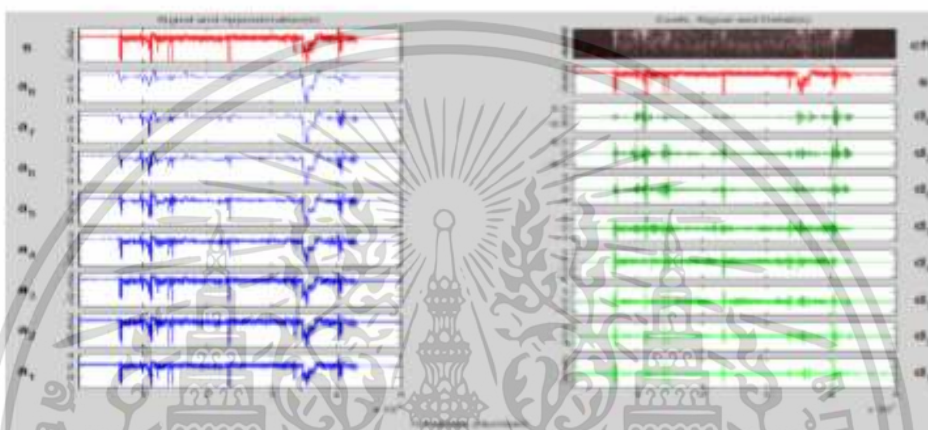


Figure 4 The signal after Discrete wavelet transform with Daubechies4 (db4) and Decomposition level 8.

4. Conclusion

Discrete wavelet transform can utilizes to analyze the change of wavelet in real time which can totally classify the railway turnout since the scale I which can quickly detect the problem occur in the railway. Additionally, it can also cut off the noise signal and only have the primary signal left. Comparison of the signal in each scale shows that the wavelets is different in both scale and time that it occur. This procedure is conducted in order to indicate the characteristics of the occurred wavelet. Therefore, the wavelet transform can be utilized to classify the abnormality by referencing the experimental wavelet results.

5. Acknowledgement

This research work is supported by faculty of engineering, King Mongkut's Institute of Technology Ladkrabang and the KMITL Research Fund under the research grant no. ERF186202.

6. References

- [1] Kaitwanidvilai S., Parichatpreecha R. 2015. Development for lateral movement inspection data of railway rapid transit from machine vision, In: Proceeding of The 2nd Thailand Rail Symp. (TRAS-2), Phitsanulok, Thailand.

- [2] Sokoloski and Pardela T. 2013. Fourier Transforms in Melanoma Image Classification. In: Proceeding of The 6th International Conference on Human System Interactions (HSI), Sopot, Poland, pp. 386-389.
- [3] Li L. 2010. A New Method of Wavelet Transform Based on FFT for Signal Processing. In: Proceeding of The 2010 Second WRI Global Congress on Intelligent Systems, Wuhan, China, pp. 203-206.
- [4] Dong L., Chen J., Xiao H., Yue H., Li Y. 2010. Processing of The Image on Gas Discharge by spatial Correlation Function and Fast Fourier Transform. In: Proceeding of 2010 International Conference on Electrical and Control Engineering, Wuhan, China, pp.1507-1510.
- [5] Rajamanickam B. and Shanmugam M. 2010. Application of Fast Fourier Transform (FFT) in Laser Speckle Image Pattern Correlation Technique for the Metrological Measurement. In: Proceeding of The 2012 Fourth International Conference on Advanced Computing (ICoAC), Chennai, India.
- [6] Shukla D. and Sahu J. 2013. Wavelets: Basic Concepts. International Journal of Electrical and Electronic Engineering & Telecommunications (IJEETC). Vol. 2. No. 4. 30-36.
- [7] Ramesh K. and Mallikharjuna Y. Power Transformer Protection Using Wavelets. International Journal of Electrical and Electronic Engineering & Telecommunications (IJEETC). Vol. 3. No. 4. 14-21.
- [8] Bajric R., Zuber N., Skrimpas G. A. and Mijatovic N. Feature Extraction Using Discrete Wavelet Transform for Gear Fault Diagnosis of Wind Turbine Gearbox. Shock and Vibration. Vol. 2016. 1-10.
- [9] Ramadevi R. 2014. Cavitation signal analysis using discrete wavelet transform. In: Proceeding of The 2014 International Conference on Control Instrumentation, Communication and Computational Technologies (ICCICCT). Kanyakumari, India, pp. 558-560.
- [10] Yohanes R. E. J., Ser W. and Huang G. B. 2012. Discrete Wavelet Transform coefficients for emotion recognition from EEG signals. In: Proceeding of The 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. San Diego, USA. pp. 2251-2254.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [3] Suparat Pumyoy, Anakkapon Seanthon, Somyot Kiatwanidvilai and Poom Konghuayrob, "Detection of the Interaction Between Wheel and Railway of the Electrical Train Utilizing Faster R-CNN Technique", The 5th National and The 1st International Conference on Informatics, Agriculture, Management Business administration, Engineering Sciences and Technology, pp. 122-131, 28-29 May 2020 (Online Conference).



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



Detection of the Interaction Between Wheel and Railway of the Electrical Train Utilizing Faster R-CNN Technique

Suparat Pumyoy^{1*}, Anakkapon Saenthon², Somyot Kiatwanidvilai¹ and Poom Konghuayrob²

¹ Department of Electrical Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand

² College of Advanced Manufacturing Innovation, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand

*Corresponding author: Suparat Pumyoy, e-mail address: suparat.pu@kmitl.ac.th

Abstract

This paper presents the Faster R-CNN detection technique which is capable of detecting the location of the object in the image rapidly and efficiently. This particular technique has been used to display the performance and also evaluate the safety of the railway system by detecting the interaction between the wheel and the railway of the electric train, the first techniques is Faster R-CNN and the second one is Harris Corner Detection techniques which is the conventional technique. The result indicates that the Faster R-CNN technique and the Border Technique give a good agreement, moreover, the Faster R-CNN techniques require less calculating time.

Keywords: Artificial Intelligence, Visual Inspection system, Detection, Rail, Track

1. Introduction

Currently, the transportation in Thailand has been growing excessively, particularly a rail transport along with the reduction of fossil fuel trend and traffic issue. The government is highly interested to extend the railway route in the capital city to support the expanding in the future. The safety is the most significant aspect regarding the rail system, particularly the wheel and railway system must be well prepared and constantly inspected. Bogdan Tomoyuki Nassu and Masato Ukai applied the basic vision to evaluate the change of train tracks [1]. The railway safety analysis can be performed by using the picture, Sehchan Oh, Sunghyuk Park and Euijin Joung presented a vision-based tracking system with a train signal system to stop trains when there is an emergency [2]. Vonani Rikhotso, Nico Steyn and Yskandar Hamam presented the 3D rail modeling and measurement for rail profile condition assessment [3]. Orhan Yaman, Mehmet Karakose and Erhan Akun presented the detection of abnormalities of railway tracks from photos taken by cameras [4]., Ze Liu, Wei Wang, Xiaofei Zhang and Wei Jia used image processing to detect the damage of the rail surface [5]. From the aforementioned reasons, an estimation technique regarding the train driving efficiency is presented, this technique recording the motion pictures and calculates the distance between the wheel and the railway from the picture with R-CNN technique which is the fast technique. Faster R-CNN technique is presented by Ross Girshick [6] which takes less time than R-CNN technique to detecting the railway condition and the interaction between wheel and railway in real-time.

2. Theory

2.1 Vision system

Vision system basically is used to process the image or calculate using the computer to obtain the required result in both quantity and quality, the significant processes are to make the higher resolution of the image, eliminate the noise, detect the interesting object from the image and to analyze the quantitative data such as size, shape, and also direction regarding the motion of an object in the image, this performance will be analyzed and utilized to shape the system to help in many tasks.

2.2 The Harris Corner Detection [7]

The angle analysis using the Harris method is popular for finding points of interest, due to its outstanding accuracy in rotation, scale, variability, image distortion and image noise. The principle of the Harris Corner Detector is utilized to find the corner of the image by looking through the small piece of image that is established which inside will be the lines of the image then move the small window that has been set up in various directions along the line. When encountering a lot of path changes or extreme line refraction will consider that point as an angle point.

That can be written as the refraction equation of the line (Change of intensity) when the shift $[u, v]$ is set as follows.

$$E(u, v) = \sum_{x, y} w(x, y) [I(x+u, y+v) - I(x, y)]^2 \quad (1)$$

When $w(x, y)$ is Windows function
 $(x+u, y+v)$ is Shifted intensity
 $[I(x+u, y+v) - I(x, y)]^2$ is Intensity
 and Window function $w(x, y) =$



For scrolling $[u, v]$ can use bilinear estimation as follows

$$E(u, v) \approx [u \quad v] M \begin{bmatrix} u \\ v \end{bmatrix} \quad (2)$$

Which M is Matrix 2x2 that has been calculated from the received image:

$$M = \sum_{x, y} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (3)$$

By value of λ_1, λ_2 is specific Eigenvalue of M

All of above equations can be written as an equation for calculating angles by Harris Corner Detector as follows :

$$R = \det M - k(\text{trac} M)^2 \quad (4)$$

Where $\det M$ can be found from $\det M = \lambda_1 \cdot \lambda_2$. The trace $M = \lambda_1 + \lambda_2$ and k is a constant with a value between 0.04-0.06. The R value will vary according to the specific value of M where R is worth a corner and will have a very negative value when the line is edge. When the absolute value R is given, the small value indicates that it is a flat.

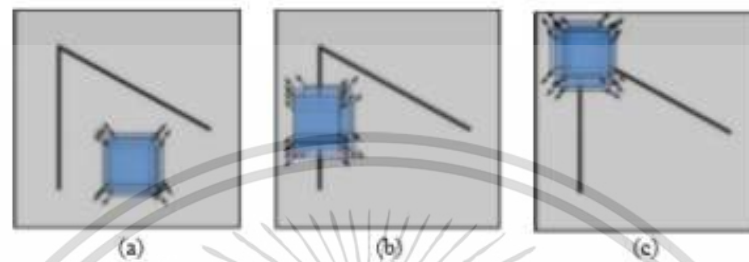


Figure 1 Harris Corner Detection: (a) Flat Region: No change in all direction, (b) Edge: No change along the edge direction, and (c) Corner: Significant change in all direction [8]

Harris algorithm has a good performance on its stability and robustness, but it is also the direct reason of the limitation of its computing speed.

2.3 Faster R-CNN

- Artificial Intelligence

Artificial Intelligence (AI) is one of the fields where computer science is related to make the computer capable of doing human tasks or imitate the human behaviour, particularly an ability to think or having the artefact. This ability was created by human, thus, it has been called the "Artificial Intelligence". The aspect regarding the AI is various depend on what kind of intelligence we required. This mainly depends on the behaviour that affects the environment or depends on how the AI thinks. Therefore, the AI has been defined by its abilities which are divided into 4 main groups:

- Acting Humanly: where the AI can act like the human being
- Thinking Humanly: where the AI can think like human being
- Thinking rationally: where the AI is able to think logically or correctly regarding the logic to finding the reasonable answer.
- Acting rationally: Where AI is able to act logically or rationally.

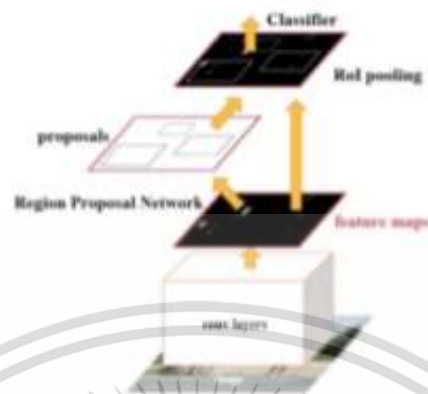


Figure 2 The structure of the VGG-16 used as the base of the model [9]

- Faster R-CNN [6]

For this research, the Faster R-CNN neuron model structure is utilized, this structure is capable of detecting the location of the object and extracting the required object on demand. The principle of the Faster R-CNN is to apply the Selective Search and include in the same neuron structure. This neuron model structure consists of three main sections:

- The base section which will extract the feature or consider the remarkable feature in the image.
- The Region Proposal Network (RPN) section. This section is functioning to distract the "expecting or possibility" from the object location.

Categorizing section which utilize the feature map and region from the RPM and processing by utilized the ROI pooling in order to find out the location which contains the object.



Figure 3 The VGG-16 Structure Utilized as A Base Model. [10]

The VGG-16 structure can support back-to-colour black and white images. The first layer can support pictures with the size 224 x 224 and RGB images. The work structure is divided into 5 convolution steps. Then the final convolution is fully-connected Classifier. Details can be shown as in Figure 4.

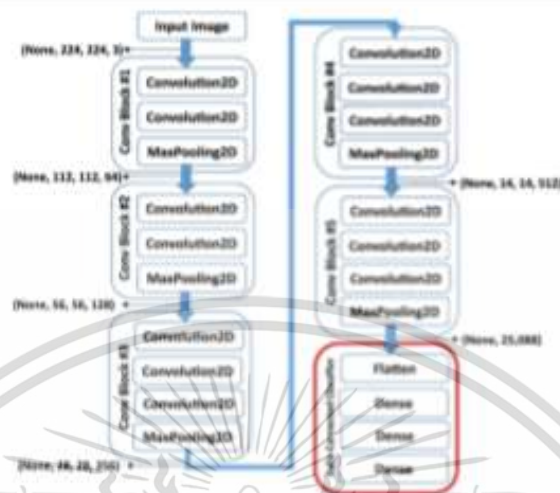


Figure 4 The process of each layer of VGG-16.

3. The Experimental Result

The experiment was conducted by convert the railway and its wheel video into the image, the number of the image used in the process is 200 pictures.

A. Measuring the distance using The Harris Corner Detection

This method uses for measuring the distance of rims and rail from the Harris Corner Detector as shown in Figure 5. This distance is utilized to verify the behavior of both rims and wheel during the electric train moving.



Figure 5 Measurement by The Harris Corner Detector

B. Measuring the distance using Faster R-CNN technique

For this technique, the data has been categorized into two main sections, the first section is for AI training purpose and another section will be utilized for the AI performance testing purpose. Two main

focuses of the image are the rim of the wheel and the edge of the railway this process is performed by using the Labeling software (to determine the training the system or learning, for 50 images used to training and 31 images used for testing)

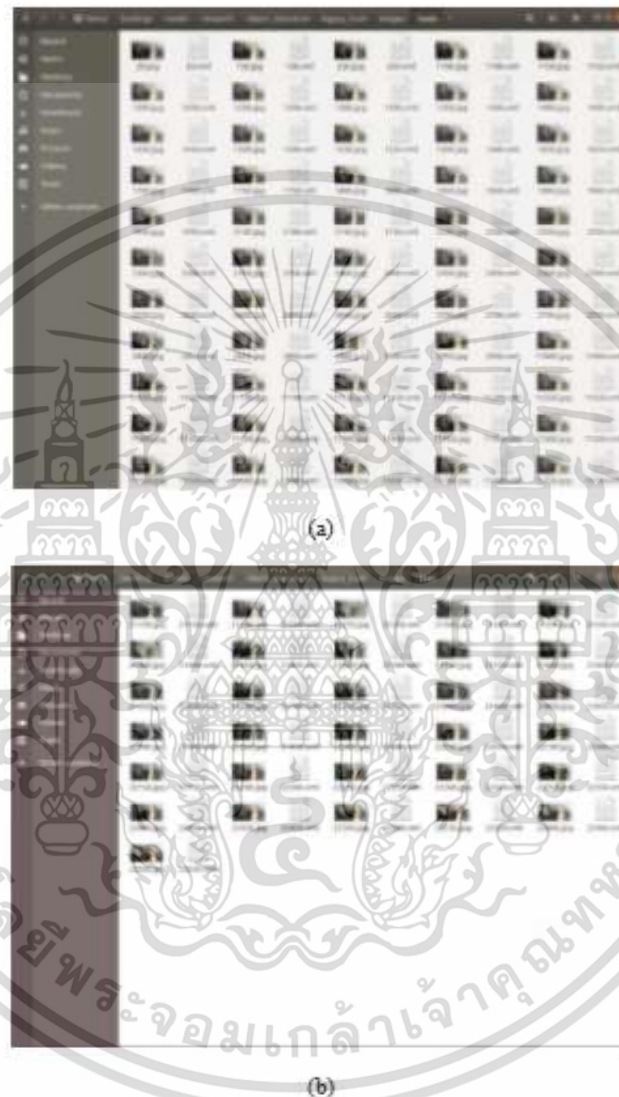


Figure 6 Image and label in experimental test: (a) training, and (b) testing.

To prepare the data set of the image, each image will be contained the label which will regulate the location of the interesting object. The file consists of the path of the file location, X value, Y value, Width, and Height, this data will be called to verify the correctness of the result from AI.

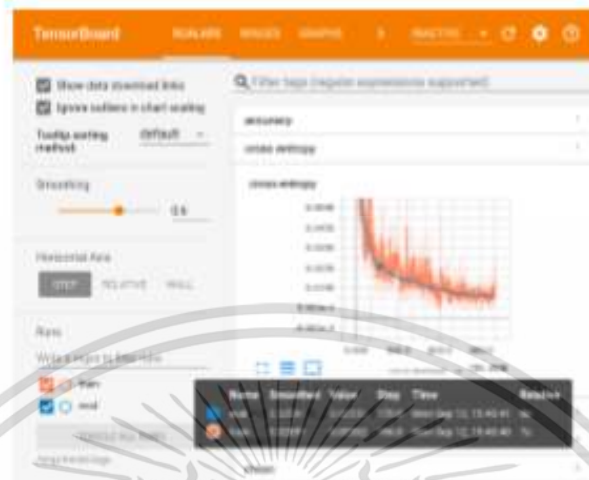


Figure 7 TensorBoard.

TensorBoard graph in figure 7 displays the result which the AI is able to recognize called a loss value. Loss value can divide into two parts which are the locating loss and categorizing loss. After running test for 10,000 times, the loss result is 0.002 and 0.004 respectively. Typically, the training process will feed the data to the AI, and the AI will analyses and find the solution from the label, if the result is too difference, the weight will be update until the value is closed to the label. The observation from the graph shows that the graph is continuously converging to the horizontal axis zero.

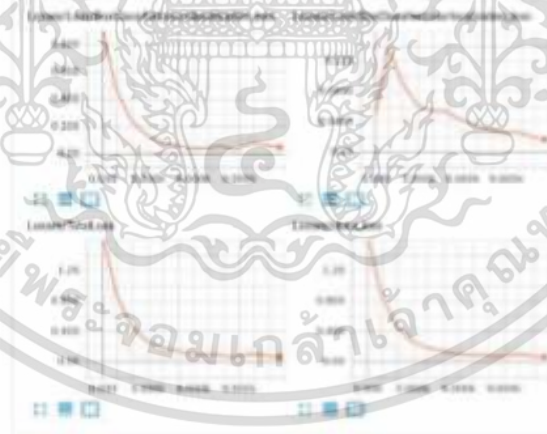


Figure 8 The loss values regarding the TensorBoard graph after running for 10,000 times.

When the model of AI passed through the training and the proper weight value improving, then the model after train can be conducted. The process was conducted by using the wheel and railway image to analyses and detects distance as shown in Fig.9.

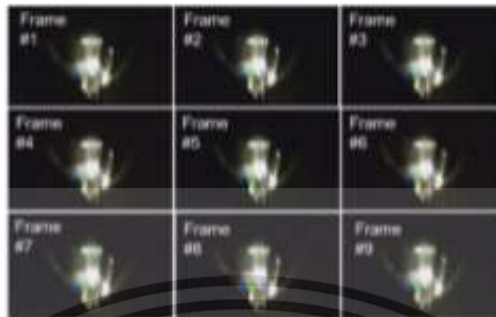


Figure 9 The test results of detecting the location regarding the rim of the wheel and the edge of the railway.



Figure 10 The test results of detecting the location regarding the rim of the wheel and the edge of the railway from the AI.

Fig. 10 shows the result after train the proposed algorithm Faster R-CNN technique. As seen in figure 10 brown box represents the rim of the wheel and the green box represents the edge of the railway. In order to compare 2 techniques Faster R-CNN technique and conventional Harris Corner Detection (HCD), the results between two detection techniques is shown in Fig.11 below.

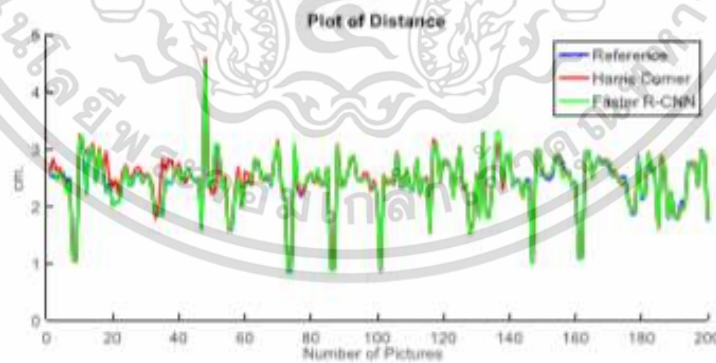


Figure 11 Measurement results by The Harris Corner Detector and Faster R-CNN technique.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

As seen in Fig. 11, the result of the proposed Faster R-CNN quite similar to the reference data, while the conventional (redline) some results is difference from the reference picture. In order to show the performance of the proposed algorithm, Least Square Error (LSE) is utilized for comparing the values obtained from both proposed techniques with reference obtained by manual measurement. The Harris Corner Detection gives an error equal 2.63% while the proposed Faster R-CNN give an error equal 0.33% respectively.

4. Conclusions

Regarding the image measuring results of the wheel and the rail via Harris Corner Detection which utilizes the angle finding technique compared to the Faster R-CNN which the interested area is presented and utilize AI to classify the attribute of object. The results indicate that the Faster R-CNN can provide more precise measuring regarding the distance between the wheel and the rail compare to the conventional with Harris Corner Detection technique. The experimental results also confirm the benefit of the precise measuring results of the proposed Faster R-CNN technique that utilizes less image processing time when compare to the Harris Corner Detection technique which has a time constrain when process the large amount of image. Therefore, the proposed Faster R-CNN can consider to be a real-time detection applied to real electrical railway or electric train safety control system that can monitor the fault between wheel and rail condition while the train is moving in the real environment.

5. Acknowledgements

Our special gratitude is for the Development Research Center, Test and Transfer Technology High speed Train of Thailand for its research funding

6. References

- [1] Nassu B. T. and Ukai M. 2012. A Vision-Basde Approach for Rail Extraction and its Application in a Camera Pan-Tilt Control System. *IEEE Transactions on Intelligent Transportation Systems*. Vol. 13, Issue 4. 1763-1771.
- [2] Oh S., Park S., and Joung E. 2009. Approaches for Connection of Vision Basad Monitoring System with Railway Signal System for Train Emergency Stop. In: *Proceeding of The 2009 International Conference on Information and Multimedia Technology*. Jeju Island, South Korea. pp. 59-63.
- [3] Rikhotso V., Steyn N., and Hamam Y. 2017. 3D Rail Modelling and Measurement for Rail Profile Condition Assessment. In: *Proceeding of the 2017 IEEE AFRICON (IEEE, Cape Town, 2017)*. Cape Town, South Africa. pp. 1522-1527.
- [4] Yaman O., Karakose M., and Akin E. 2017. A Vision Based Diagnosis Approach for Multi Rail Surface Faults Using Fuzzy Classification on Railways. In: *Proceeding of The 2017 International Conference on Computer Science and Engineering (UBMK, Antalya, 2017)*. Antalya, Turkey. pp. 713-718.
- [5] Liu Z., Wang W., Zhang X., and Jia W. 2010. Inspection of Rail Surface Defects Based on Image Processing. In: *Proceeding of The 2010 2nd International Asia Conference on Informatics in Control, Automation and Robotics (CAR, Wuhan, 2010)*. Wuhan, China. pp. 472-475.

- [6] Girshick R. 2015. Fast R-CNN. In: Proceeding of The 2015 IEEE International Conference on Computer Vision (ICCV, Santiago, 2015). Santiago, Chile. pp. 1440-1448.
- [7] Chenguang G, Xianglong L., Linfeng Z. and Xiang L. 2009. A Fast and Accurate Corner Detector Based on Harris Algorithm. In: Proceeding of The 2009 Third International Symposium on Intelligent Information Technology Application (Shanghai, 2009). Shanghai, China. pp. 49-52.
- [8] Deshmukh A. C., Gaikwad P. R. and Patil M. P. 2015. A Comparative Study of Feature Detection Methods. International Journal of Electrical and Electronic Engineering & Telecommunications (IJEETC 2015) Vol. 4, No. 2. 1-7
- [9] Ren S., He K., Girshick R. and Sun J. 2017. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Transaction on pattern analysis and machine intelligence, Vol. 39, No. 6.
- [10] Li Y. T. and Guo J. L. A VGG-16 Based Faster RCNN Model for PCB Error Inspection in Industrial AOI Applications. In: Proceeding of The 2018 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW). Taichung, Taiwan.



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

| | |
|------------------|---|
| ชื่อ-นามสกุล | นางสาวสุภารัตน์ พุ่มย้อย |
| วัน เดือน ปีเกิด | 9 ธันวาคม 2530 |
| ที่อยู่ | 8 ซอยฉลองกรุง 8 แขวงทับยาว เขตลาดกระบัง จังหวัดกรุงเทพมหานคร 10520 |
| ประวัติการศึกษา | พ.ศ. 2552 วิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหาร ลาดกระบัง พ.ศ. 2557 วิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหาร ลาดกระบัง |
| ประวัติการทำงาน | พ.ศ. 2554-ปัจจุบัน ตำแหน่งวิศวกร ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง |

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้