

การพัฒนาโครงข่าย CNN-RNN สำหรับตรวจคัดกรองโรคโควิด-19
จากภาพเอกซเรย์ และภาพเอกซเรย์คอมพิวเตอร์ทรวงอก

CNN-RNN NETWORK INTEGRATION FOR THE DIAGNOSIS OF COVID-19 FROM
CHEST X-RAY AND CT IMAGES



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรดุษฎีบัณฑิต
สาขาวิชาวิศวกรรมไฟฟ้า
คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ. 2566

KMITL-2024-EN-D-018-260

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

CNN-RNN NETWORK INTEGRATION FOR THE DIAGNOSIS OF COVID-19 FROM
CHEST X-RAY AND CT IMAGES



A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF
DOCTOR OF ENGINEERING IN ELECTRICAL ENGINEERING
SCHOOL OF ENGINEERING
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
2023
KMUTL-2024-EN-D-018-260

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



COPYRIGHT 2023

SCHOOL OF ENGINEERING

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์	การพัฒนาโครงข่าย CNN-RNN สำหรับตรวจคัดกรองโรคโควิด-19 จากภาพเอกซเรย์ และภาพเอกซเรย์คอมพิวเตอร์ทรวงอก
นักศึกษา	นาย กษิ เตียงหงษากุล
รหัสประจำตัว	64601017
ปริญญา	วิศวกรรมศาสตรดุษฎีบัณฑิต
สาขาวิชา	วิศวกรรมไฟฟ้า
พ.ศ.	2566
อาจารย์ที่ปรึกษาวิทยานิพนธ์	รศ.ดร. อรรถสิทธิ์ หล้าสกุล

บทคัดย่อ

วิทยานิพนธ์ฉบับนี้นำเสนอการพัฒนาการรวมโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution neural network, CNN) เข้ากับโครงข่ายประสาทเทียมแบบวนกลับ (Recurrent neural network, RNN) สำหรับตรวจคัดกรองโรคโควิด-19 (COVID-19) จากภาพเอกซเรย์ทรวงอก (X-ray image) และภาพเอกซเรย์คอมพิวเตอร์ทรวงอก (Computerized tomography image, CT image) เพื่อเป็นอีกหนึ่งแนวทางในการระบุผู้ติดเชื้อ ซึ่งช่วยควบคุมอัตราการแพร่ระบาดของเชื้อ และสามารถวางแผนการรักษาผู้ป่วยให้ได้ทันเวลาที่ โดยวิธีตรวจหาแบบอัตโนมัตินั้นช่วยให้การตรวจคัดกรองมีความสะดวก และรวดเร็วมากยิ่งขึ้น ในการพัฒนาการรวมโครงข่าย Convolution neural network ถูกใช้สำหรับการสกัดคุณลักษณะเด่นภายในภาพ (Feature extraction) และ Recurrent neural network ใช้สำหรับการหาความสัมพันธ์ของคุณลักษณะภายในภาพ โดยโครงข่ายประสาทเทียมแบบคอนโวลูชันที่ผ่านการเรียนรู้ข้อมูลภาพแล้ว (Pre-trained model) ได้แก่ VGG-19 ResNet152V2 และ DenseNet121 ถูกนำมารวมเข้ากับ Long short-term memory (LSTM) และ Gated recurrent unit (GRU) เพื่อใช้จำแนกภาพถ่ายรังสีทรวงอกของผู้ป่วยโรคโควิด-19 โรคปอดอักเสบ และภาวะปกติ ซึ่งมีจำนวนทั้งหมด 16,210 ภาพ และเปรียบเทียบหาโครงข่ายที่มีประสิทธิภาพสูงสุดในการจำแนกภาพถ่ายรังสีทรวงอกในภาวะต่าง ๆ ผลลัพธ์ที่ได้จากการทดสอบพบว่าแบบจำลอง ResNet152V2 ร่วมกับ GRU ที่ใช้ภาพต้นฉบับในการประมวลผลมีประสิทธิภาพความแม่นยำในการจำแนกภาพถ่ายรังสีทรวงอกของผู้ป่วยโรคโควิด-19 สูงที่สุด โดยมีค่าความแม่นยำ อยู่ที่ 94.14% ค่า F1-score 92.57% ค่า precision 90.58% และค่า recall 94.64% อีกทั้งยังมีประสิทธิภาพความแม่นยำโดยรวมในการจำแนกทั้งสามภาวะโรคสูงที่สุดอีกเช่นกัน โดยมีค่าความแม่นยำ 93.37% และใช้เวลาเฉลี่ยในการทำนายผล 0.02 วินาทีต่อภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Thesis	CNN-RNN NETWORK INTEGRATION FOR THE DIAGNOSIS OF COVID-19 FROM CHEST X-RAY AND CT IMAGES
Student	Mr. Kasi Tenghongsakul
Student ID.	64601017
Degree	Doctor of Engineering
Program	Electrical Engineering
Year	2023
Thesis Advisor	Assoc. Prof. Dr. Attasit Lasakul

ABSTRACT

This thesis introduces the development of a combined Convolutional neural network (CNN) and Recurrent neural network (RNN) architecture for the screening and detection of COVID-19 from X-ray and computed tomography images. This serves as another approach to identify infected individuals, contributing to the control of the spread of the virus and enabling timely treatment planning. The proposed automated detection method enhances the convenience and efficiency of screening procedures, making them more convenient and faster. In the development of the integrated network, CNN models, specifically pre-trained models such as VGG-19, ResNet152V2, and DenseNet121, were utilized for extracting various features within the images. Simultaneously, Recurrent neural network models, including Long short-term memory and gated recurrent unit, were employed to capture the relationships among these extracted features. The classification task centered on chest X-ray and CT images of patients with COVID-19, pneumonia, and normal conditions, with a comprehensive dataset comprising a total of 16,210 images. The results obtained from the testing indicate that the ResNet152V2 model combined with GRU demonstrates the highest accuracy in classifying X-ray images of patients with COVID-19. The performance metrics include an accuracy of 94.14%, an F1-score of 92.57%, precision of 90.58%, and recall of 94.64%. Additionally, it exhibits overall accuracy in classifying all three conditions, with an accuracy of 93.37%. The average prediction time per image is 0.02 seconds.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วย ความกรุณาและความอนุเคราะห์จากท่านอาจารย์ที่
ปรึกษา รศ.ดร. อรรถสิทธิ์ หล้าสกุล ที่ให้คำปรึกษา และคำแนะนำต่าง ๆ ในการทำวิทยานิพนธ์ฉบับ
นี้ รวมถึงมอบความรู้ให้แก่ข้าพเจ้า ๆ ขอขอบพระคุณเป็นอย่างยิ่ง

ขอขอบคุณท่าน รศ.ดร. บุญยชนะ ภูระหงษ์ และ รศ.ดร.ชวลิต เบญจางคประเสริฐ ที่มอบ
ความช่วยเหลือทั้งในเรื่องโอกาสทางการศึกษา การวิจัย และในด้านอื่น ๆ แก่ข้าพเจ้า

ขอขอบคุณท่าน ดร. ไศศุรย์ กาญจนสุรัตน์ ที่ให้คำปรึกษาในทุกด้าน ตั้งแต่แนวทางการศึกษา
การวิจัย มอบความรู้ พร้อมให้ความช่วยเหลือ และกำลังใจแก่ข้าพเจ้าเสมอมา ตลอดจนวิทยานิพนธ์
ฉบับนี้เสร็จสมบูรณ์

ขอขอบคุณท่าน ดร.ธนวิชัย อนุวงศ์พินิจ ที่ให้คำปรึกษาเรื่องระเบียบ และข้อบังคับ รวมถึง
มอบโอกาสในสายงานให้แก่ข้าพเจ้า

ขอกราบขอบพระคุณ บิดา มารดา และครอบครัวของข้าพเจ้าที่ให้การอุปการะอบรมเลี้ยงดู
ตลอดจนส่งเสริมการศึกษา และให้กำลังใจแก่ข้าพเจ้าเป็นอย่างดี ประโยชน์อันใดที่พึงได้จาก
วิทยานิพนธ์ฉบับนี้ ข้าพเจ้าขอมอบเป็นกตัญญูแก่แต่บิดามารดา และบูรพาจารย์ที่เคยอบรมสั่ง
สอน รวมทั้งผู้มีพระคุณทุกท่าน

กษิ เตียงหงษากุล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	I
บทคัดย่อภาษาอังกฤษ	II
กิตติกรรมประกาศ	III
สารบัญ.....	IV
สารบัญตาราง.....	VII
สารบัญรูป	VIII
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 จุดมุ่งหมายและวัตถุประสงค์.....	2
1.3 แนวคิดที่ใช้ในการวิจัย.....	2
1.4 ขอบเขตของงานวิจัย	2
1.5 ขั้นตอนในการศึกษา.....	3
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง	4
2.1 รังสีวินิจฉัย (Radiology).....	4
2.1.1 ภาพทางรังสี (Radiography)	4
2.1.1.1 ภาพเอกซเรย์ (X-ray)	4
2.1.1.2 ภาพเอกซเรย์คอมพิวเตอร์ (Computed tomography, CT)	5
2.1.2 ภาพวินิจฉัยทางรังสีวิทยาในโควิด-19	7
2.2 การประมวลผลภาพดิจิทัล.....	9
2.2.1 ภาพดิจิทัล.....	10
2.2.1.1 ภาพขาวดำ (Binary image).....	11
2.2.1.2 ภาพระดับสีเทา (Gray scale image).....	11
2.2.1.3 ภาพสีแบบ RGB (RGB image).....	12
2.2.2 ฮิสโตแกรมและคอนทราสต์ของภาพ	12
2.2.2.1 ฮิสโตแกรมของภาพ (Image Histogram).....	12
2.2.2.2 คอนทราสต์ของภาพ (Contrast).....	13
2.2.3 ตัวดำเนินการสำหรับการประมวลผลภาพ (Image processing operators).....	14
2.2.3.1 การดำเนินการทางเลขคณิตของภาพ (Image arithmetic)	14

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
2.2.4 การปรับปรุงภาพเชิงรังสี (Radiometric image enhancement).....	17
2.2.4.1 การปรับปรุงความสว่างของภาพแบบเชิงเส้น.....	17
2.2.4.2 การปรับปรุงความสว่างของภาพแบบไม่เป็นเชิงเส้น	18
2.2.4.3 การปรับปรุงความสว่างของภาพแบบช่วง	22
2.3 โครงข่ายประสาทเทียม (Neural network)	22
2.3.1 องค์ประกอบของโครงข่ายประสาทเทียม.....	22
2.3.2 การเรียนรู้ของโครงข่ายประสาทเทียม.....	25
2.3.3 ฟังก์ชันกระตุ้น (Activation Functions).....	26
2.3.3.1 ฟังก์ชันกระตุ้นแบบเชิงเส้น (Linear activation function)	26
2.3.3.2 ฟังก์ชันกระตุ้นแบบไม่เป็นเชิงเส้น (Non-linear activation function)..	27
2.4 การเรียนรู้เชิงลึก (Deep learning)	30
2.4.1 องค์ประกอบของการเรียนรู้เชิงลึก.....	30
2.4.1.1 ไฮเปอร์พารามิเตอร์ (Hyperparameters).....	31
2.4.1.2 ฟังก์ชันการสูญเสีย (Loss function).....	31
2.4.1.3 อัลกอริทึมปรับปรุงประสิทธิภาพ (Optimization algorithm).....	32
2.4.2 โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional neural network)....	33
2.4.3 แบบจำลองการถ่ายโอนการเรียนรู้เชิงลึก (Deep transfer learning models) .	36
2.4.4 โครงข่ายประสาทเทียมแบบวนกลับ (Recurrent neural network).....	37
2.5 ทบทวนวรรณกรรม	41
บทที่ 3 วิธีดำเนินการทดลอง.....	43
3.1 การรวบรวมชุดข้อมูลภาพรังสีทรวงอก	44
3.2 การปรับปรุงคุณภาพสีของชุดข้อมูลภาพรังสีทรวงอก	47
3.3 การพัฒนาโครงข่าย CNN และ CNN-RNN สำหรับจำแนกโรคทรวงอกจากภาพรังสี.....	47
3.3.1 การพัฒนาโครงข่าย CNN สำหรับจำแนกโรคทรวงอกจากภาพรังสี.....	48
3.3.2 การพัฒนาโครงข่าย CNN-RNN สำหรับจำแนกโรคทรวงอกจากภาพรังสี	49
3.4 การเตรียมข้อมูลเพื่อใช้ในกระบวนการเรียนรู้และทดสอบโครงข่าย	52
3.5 กระบวนการเรียนรู้ข้อมูลในการจำแนกโรคทรวงอกจากภาพถ่ายรังสีของแบบจำลอง	53
3.6 การจำแนกหมวดหมู่โรคทรวงอก และการประเมินค่าประสิทธิภาพของแบบจำลอง.....	55

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
บทที่ 4 ผลการทดลอง	57
4.1 ผลลัพธ์การปรับปรุงคุณภาพสีของชุดข้อมูลภาพรังสีทรวงอก.....	57
4.2 ผลลัพธ์การพัฒนาโครงข่าย CNN และ CNN-RNN สำหรับจำแนกโรคทรวงอกจากภาพถ่ายรังสี.....	59
4.3 การเรียนรู้ข้อมูลในการจำแนกโรคทรวงอกจากภาพถ่ายรังสีของแบบจำลอง	60
4.3.1 ข้อมูลสำหรับการเรียนรู้.....	60
4.3.2 ผลการเรียนรู้การจำแนกโรคทรวงอกจากภาพถ่ายรังสีของแบบจำลอง.....	62
4.4 ผลการจำแนกหมวดหมู่โรคทรวงอก และการประเมินค่าประสิทธิภาพของแบบจำลอง...	66
4.5 เวลาในการประมวลผลของกระบวนการเรียนรู้ และทำนายผลการจำแนกภาวะโรคทรวงอกจากภาพถ่ายรังสี	75
บทที่ 5 สรุปผลการวิจัยและการต่อยอดในอนาคต	77
5.1 สรุปผลการวิจัย.....	77
5.2 การพัฒนาต่อยอดในอนาคต.....	78
บรรณานุกรม	79

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่	หน้า
3.1 จำนวนข้อมูลที่ใช้ในงานวิจัย	46
3.2 จำนวนภาพของชุดข้อมูลที่ใช้ในกระบวนการต่าง ๆ ของแบบจำลอง	52
4.1 ข้อมูลโครงสร้างแบบจำลองที่พัฒนาขึ้นสำหรับจำแนกโรคทรวงอกจากภาพถ่ายรังสี	59
4.2 ค่าผลลัพธ์การเรียนรู้และการตรวจสอบของโครงข่ายแบบจำลอง CNN	64
4.3 ค่าผลลัพธ์การเรียนรู้และการตรวจสอบของโครงข่ายแบบจำลอง CNN-RNN	65
4.4 ค่าประสิทธิภาพการจำแนกหมวดหมู่โรคทรวงอกของโครงข่ายแบบจำลอง CNN	67
4.5 ค่าประสิทธิภาพการทำนายผลการจำแนกหมวดหมู่โรคทรวงอกของโครงข่าย CNN-RNN ที่ใช้ แบบจำลอง CNN เป็น ResNet152V2	68
4.6 ค่าประสิทธิภาพการทำนายผลการจำแนกหมวดหมู่โรคทรวงอกของโครงข่าย CNN-RNN ที่ใช้ แบบจำลอง CNN เป็น DenseNet121	70
4.7 ค่าประสิทธิภาพการทำนายผลการจำแนกหมวดหมู่โรคทรวงอกของโครงข่าย CNN-RNN ที่ใช้ แบบจำลอง CNN เป็น VGG19	72
4.8 ผลลัพธ์ประสิทธิภาพที่ดีที่สุดของแต่ละโครงข่ายตามแบบจำลอง CNN ที่ใช้	73
4.9 การเปรียบเทียบประสิทธิภาพของวิธีการที่นำเสนอกับงานวิจัยอื่น	74
4.10 เวลาที่ใช้ในการเรียนรู้ และทำนายผลการจำแนกภาวะโรคทรวงอกจากภาพถ่ายรังสี	75

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่	หน้า
2.1 ลักษณะของลำรังสีเอกซ์.....	5
2.2 ตัวอย่างภาพเอกซเรย์.....	5
2.3 ขั้นตอนการเก็บชุดข้อมูลภาพฉายรังสี และกระบวนการประมวลผลเพื่อสร้างภาพใหม่ของการ ถ่ายภาพเอกซเรย์คอมพิวเตอร์.....	6
2.4 ตัวอย่างภาพ CT	6
2.5 ภาพเอกซเรย์ทรวงอกแสดงรอยโรคในถุงลมเป็นฝ้าจาง (ground glass opacity) โดยกระจายอยู่ ที่ปอดกลีบล่างทั้งสองข้าง	8
2.6 ภาพเอกซเรย์คอมพิวเตอร์ทรวงอกแสดงลักษณะเป็นฝ้าจาง (ground glass opacity) โดยพบ รอยโรคที่ขอบชายปอด.....	9
2.7 เมทริกซ์ภาพดิจิทัล.....	10
2.8 ภาพขาวดำ และค่าความสว่างของแต่ละจุดภาพ	11
2.9 ภาพระดับสีเทา และค่าความเข้มแสงของแต่ละจุดภาพ	11
2.10 ภาพสีแบบ RGB และค่าความเข้มของแต่ละช่องสัญญาณสีในหนึ่งจุดภาพ	12
2.11 กราฟฮิสโตแกรมของภาพ	13
2.12 ตัวอย่างภาพคอนทราสต์ต่ำ	14
2.13 ตัวอย่างภาพคอนทราสต์สูง.....	14
2.14 ตัวอย่างการดำเนินการทางภาพแบบบวกระหว่างภาพกับค่าคงที่	15
2.15 ตัวอย่างการดำเนินการทางภาพแบบหักล้างระหว่างภาพกับค่าคงที่	15
2.16 ตัวอย่างการดำเนินการทางภาพแบบคูณระหว่างภาพกับค่าคงที่	16
2.17 ตัวอย่างการดำเนินการทางภาพแบบหารระหว่างภาพกับค่าคงที่	16
2.18 กราฟความสัมพันธ์ของการขยายพิสัยค่าความสว่างแบบเชิงเส้น.....	17
2.19 การปรับปรุงความสว่างของภาพแบบเชิงเส้น	18
2.20 กราฟความสัมพันธ์ของการขยายค่าความสว่างของภาพแบบไม่เป็นเชิงเส้น	18
2.21 การปรับปรุงความสว่างแบบ Histogram Equalization	19
2.22 การกระจายความถี่ของพิกเซลที่เกินขีดจำกัดของ Clip limit	20
2.23 ตัวอย่างภาพและฮิสโตแกรมของภาพหลังจากกระบวนการปรับปรุงแบบ Contrast Limited Adaptive Histogram Equalization	21
2.24 กราฟความสัมพันธ์ของการปรับปรุงความสว่างของภาพแบบช่วง	22
2.25 โครงสร้างหน่วยการทำงานของโครงข่ายประสาทเทียม	23
2.26 โครงข่ายประสาทเทียมหลายชั้นแบบป้อนไปข้างหน้า	24

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อการศึกษาเท่านั้น เมื่ออนุญาตให้ท่านไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่	หน้า
2.27 แผนภาพรูปแบบการเรียนรู้ของโครงข่ายประสาทเทียม	26
2.28 กราฟฟังก์ชันกระตุ้นแบบเชิงเส้น	27
2.29 กราฟฟังก์ชันกระตุ้น Rectified linear unit	27
2.30 กราฟฟังก์ชันกระตุ้น Sigmoid	28
2.31 กราฟฟังก์ชันกระตุ้น Hyperbolic tangent	29
2.32 ความแตกต่างของการจำแนกภาพระหว่าง Deep learning และ Machine learning แบบ ดั้งเดิม	30
2.33 ชั้นการทำงานของโครงข่ายประสาทเทียมแบบคอนโวลูชัน	33
2.34 การทำงานของชั้นคอนโวลูชัน	33
2.35 ผลลัพธ์การคอนโวลูชัน	34
2.36 ตัวอย่างการพูลลิงในแต่ละเทคนิคฟังก์ชัน	35
2.37 ชั้นเชื่อมโยงสมบูรณ์ในโครงข่าย CNN	35
2.38 โครงสร้างแบบจำลอง VGG19	36
2.39 โครงสร้างแบบจำลอง ResNet152V2	37
2.40 โครงสร้างแบบจำลอง DenseNet121	37
2.41 โครงสร้างของโครงข่ายประสาทเทียมแบบวนกลับ	38
2.42 โครงสร้างของโครงข่าย Long short-term memory	39
2.43 โครงสร้างของโครงข่าย Gated recurrent unit	40
3.1 แผนภาพลำดับขั้นตอนการทำงานของวิธีการที่นำเสนอ	44
3.2 ตัวอย่างข้อมูลภาพรังสีทรวงอกที่ใช้ในการทดลอง	46
3.3 ขั้นตอนการปรับปรุงคุณภาพสีของภาพรังสีทรวงอก	47
3.4 แผนผังสถาปัตยกรรมแบบจำลอง CNN ที่นำเสนอ	49
3.5 โครงสร้างแบบจำลอง CNN-RNN สำหรับจำแนกโรคทรวงอกจากภาพรังสีที่นำเสนอ	50
3.6 แผนผังสถาปัตยกรรมแบบจำลอง CNN-RNN ที่นำเสนอ	51
3.7 ขั้นตอนการจัดเตรียมข้อมูลเพื่อใช้ในกระบวนการเรียนรู้ และทดสอบโครงข่าย	53
3.8 ขั้นตอนการเรียนรู้ของโครงข่ายแบบจำลอง	54
3.9 ขั้นตอนการการทำนายผลของโครงข่ายแบบจำลอง	55
4.1 ตัวอย่างภาพรังสีทรวงอกต้นฉบับ	57
4.2 ตัวอย่างภาพรังสีทรวงอกหลังผ่านกระบวนการปรับปรุงคุณภาพสี	58

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่	หน้า
4.3 ตัวอย่างข้อมูลภาพเอกซเรย์ และป้ายกำกับที่ใช้ในการเรียนรู้	60
4.4 ตัวอย่างข้อมูลภาพ CT และป้ายกำกับที่ใช้ในการเรียนรู้	61
4.5 ตัวอย่างการเพิ่มข้อมูลเพิ่มความหลากหลายของข้อมูลภาพเอกซเรย์ และภาพ CT.....	61
4.6 การปรับช่วงค่าความเข้มสีในแต่ละพิกเซลของภาพให้อยู่ในช่วงระหว่าง 0 ถึง 1.....	62
4.7 กราฟประวัติค่าความแม่นยำ (Accuracy) และค่าความคลาดเคลื่อน (Loss) ในกระบวนการเรียนรู้ของโครงข่ายที่ใช้แบบจำลอง CNN เป็น ResNet152V2 ด้วยชุดข้อมูลภาพต้นฉบับ	63
4.8 กราฟประวัติค่าความแม่นยำ (Accuracy) และค่าความคลาดเคลื่อน (Loss) ในกระบวนการเรียนรู้ของโครงข่ายที่ใช้แบบจำลอง CNN เป็น DenseNet121 ด้วยชุดข้อมูลภาพต้นฉบับ.....	63
4.9 กราฟประวัติค่าความแม่นยำ (Accuracy) และค่าความคลาดเคลื่อน (Loss) ในกระบวนการเรียนรู้ของโครงข่ายที่ใช้แบบจำลอง CNN เป็น VGG19 ด้วยชุดข้อมูลภาพต้นฉบับ	64
4.10 เมตริกชี้วัดประสิทธิภาพการทำนายผลของแบบจำลอง CNN	66
4.11 เมตริกชี้วัดประสิทธิภาพการทำนายผลของโครงข่าย CNN-RNN ที่ใช้แบบจำลอง CNN เป็น ResNet152V2	67
4.12 เมตริกชี้วัดประสิทธิภาพการทำนายผลของโครงข่าย CNN-RNN ที่ใช้แบบจำลอง CNN เป็น DenseNet121	69
4.13 เมตริกชี้วัดประสิทธิภาพการทำนายผลของโครงข่าย CNN-RNN ที่ใช้แบบจำลอง CNN เป็น VGG19	71

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

เมื่อปลายปี ค.ศ. 2019 เริ่มเกิดการแพร่ระบาดของโรคโควิด-19 (Coronavirus disease 2019, COVID-19) ทั่วโลก ซึ่งถัดมาในเดือนมีนาคม ปี ค.ศ. 2020 องค์การอนามัยโลก (World health organization, WHO) ได้ประกาศภาวะฉุกเฉินด้านสาธารณสุขระหว่างประเทศ และยกระดับการแพร่ระบาดของโรคโควิด-19 เป็นภาวะการระบาดใหญ่ทั่วโลก ด้วยจำนวนผู้ป่วยมากกว่า 118,000 ราย และผู้เสียชีวิต 4,291 ราย ใน 114 ประเทศ [1] อีกทั้งยังมีแนวโน้มของอัตราการแพร่ระบาดสูงขึ้นเรื่อย ๆ ซึ่งส่งผลกระทบต่อทั้งในด้านสุขภาพ ด้านเศรษฐกิจ และสังคมอย่างมหาศาล [2]

โรคโควิด-19 เกิดจากการติดเชื้อไวรัสโคโรนาสายพันธุ์ใหม่ ที่จัดอยู่ในตระกูลเดียวกับเชื้อไวรัสโคโรนาสายพันธุ์ที่ถูกรับก่อนหน้าคือ โรคซาร์ส (SARS) และโรคเมอร์ส (MERS) ซึ่งก่อให้เกิดการอักเสบในระบบทางเดินหายใจต่อผู้ติดเชื้อ เชื้อไวรัสชนิดนี้มีการแพร่กระจายจากมนุษย์สู่มนุษย์ได้โดยส่งผ่านสารคัดหลั่งของผู้ติดเชื้อที่ฟุ้งกระจายออกมาด้วยการไอ การจาม หรืออื่น ๆ จึงทำให้เกิดการติดต่อของโรค และแพร่กระจายได้อย่างรวดเร็ว [3] ในช่วงปี ค.ศ. 2019 ถึง ค.ศ. 2021 วิธีที่นิยมใช้มากที่สุด และเป็นที่ยอมรับในการตรวจคัดกรองโรคโควิด-19 คือ Real-time reverse transcription-polymerase chain reaction (RT-PCR) ซึ่งในบางสถานพยาบาลมีทรัพยากรไม่เพียงพอต่อจำนวนผู้ป่วยที่เข้ารับการตรวจคัดกรอง อีกทั้งวิธีการนี้ยังใช้เวลานานในการให้ผลลัพธ์ โดยใช้เวลาสูงที่สุดถึง 48 ชั่วโมง ด้วยเหตุผลเหล่านี้การพัฒนาวิธีการอื่นในการระบุผู้ติดเชื้อโรคโควิด-19 จึงมีความสำคัญต่อการช่วยควบคุมอัตราการแพร่ระบาด และวางแผนการรักษาผู้ป่วยให้ได้ทันเวลาที่

จากผลการศึกษาผู้ป่วยที่ติดเชื้อโรคโควิด-19 นั้นสรุปได้ว่าสามารถใช้ภาพเอกซเรย์ทรวงอก (X-ray image) และภาพเอกซเรย์คอมพิวเตอร์ทรวงอก (Computerized tomography image) ในการคัดกรองผู้ป่วยโรคโควิด-19 ได้อย่างแม่นยำจากการวินิจฉัยรอยโรคภายในปอดโดยแพทย์ผู้เชี่ยวชาญ [4,5] อีกทั้งในปัจจุบันแบบจำลองการเรียนรู้เชิงลึก (Deep learning) เป็นที่นิยมนำมาประยุกต์ใช้ในทางการแพทย์สำหรับวินิจฉัยโรคต่าง ๆ เนื่องจากมีประสิทธิภาพสูง และช่วยอำนวยความสะดวกต่อบุคลากรทางการแพทย์ เนื่องจากแบบจำลองการเรียนรู้เชิงลึกนั้นสามารถกำหนดข้อมูลที่ต้องการให้เรียนรู้ และประมวลผลผลลัพธ์ตามข้อมูลตัวอย่างที่ถูกป้อนให้เรียนรู้ได้เป็นอย่างดี ซึ่งการเลือกนำภาพเอกซเรย์ และภาพ CT มาใช้ร่วมกับโครงข่ายแบบจำลองการเรียนรู้เชิงลึกสำหรับตรวจคัดกรองนั้นเป็นอีกหนึ่งทางเลือก

ที่ช่วยลดปัญหาการขาดแคลนทรัพยากรที่ใช้ในการตรวจคัดกรองผู้ป่วยโรคโควิด-19 อีกทั้งยังสามารถทราบผลได้ในทันที

ในงานวิจัยนี้จึงเสนอการพัฒนาแบบจำลองการเรียนรู้เชิงลึก สำหรับตรวจคัดกรองผู้ป่วยโรคโควิด-19 จากทั้งภาพเอกซเรย์ และภาพ CT ทรวงอก ที่มีความแตกต่างในด้านคุณลักษณะของภาพได้โดยใช้โครงข่ายเดียวกัน นอกจากนี้ยังสามารถต่อยอดพัฒนาเพื่อใช้ในการคัดกรองโรคอื่นได้อีกด้วย

1.2 จุดมุ่งหมายและวัตถุประสงค์

งานวิจัยนี้มีจุดมุ่งหมายและวัตถุประสงค์ เพื่อศึกษาและพัฒนาโครงข่ายประสาทเทียมสำหรับจำแนกโรคของผู้ป่วยจากภาพเอกซเรย์ และภาพ CT ทรวงอก เพื่อคัดกรองผู้ป่วยที่ติดเชื้อโรคโควิด-19 รวมถึงประเมินเปรียบเทียบประสิทธิภาพของโครงข่ายด้วยชุดข้อมูลสาธารณะจากหลายแหล่งที่มา เพื่อหาโครงข่ายที่เหมาะสมในการคัดกรองโรคโควิด-19 และสามารถต่อยอดในการใช้จำแนกภาวะโรคอื่น ๆ ได้

1.3 แนวคิดที่ใช้ในการวิจัย

หลักการที่ใช้ในงานวิจัยนี้คือ การสร้างโครงข่ายการเรียนรู้แบบมีผู้สอน (Supervised learning) ที่เกิดจากการนำแบบจำลอง Convolution neural network (CNN) ที่มีคุณสมบัติเด่นในการสกัดคุณลักษณะ (Feature extraction) ของรูปภาพ มารวมเข้ากับแบบจำลอง Recurrent neural network (RNN) ในการหาความสัมพันธ์ของคุณลักษณะที่ได้จากการสกัดโดย CNN ในการสร้างความเข้าใจเชิงลึกเกี่ยวกับความสัมพันธ์ระหว่างข้อมูลภาพ และหมวดหมู่ของภาพได้ดีขึ้น เพื่อใช้ในการจำแนกภาวะโรคทรวงอกจากภาพเอกซเรย์ และภาพ CT ทรวงอก โดยแบบจำลอง CNN ที่ผ่านการเรียนรู้แล้ว (Pre-trained model) ซึ่งมีประสิทธิภาพสูงในการจำแนกหมวดหมู่ภาพ ได้แก่ ResNet152V2 DenseNet121 และ VGG19 ถูกนำมาประยุกต์ใช้ด้วยเทคนิคการถ่ายโอนข้อมูลการเรียนรู้ (Transfer learning) ก่อนนำมาเชื่อมเข้ากับ LSTM และ GRU ซึ่งเป็นแบบจำลองประเภท RNN โดยผลลัพธ์การจำแนกหมวดหมู่โรคทรวงอกที่ได้จากแต่ละโครงข่ายจะถูกนำมาประเมิน เพื่อหาโครงข่ายที่ให้ประสิทธิภาพสูงสุด

1.4 ขอบเขตของงานวิจัย

1.) ชุดข้อมูลที่ใช้ในการเรียนรู้และประเมินผลของโครงข่ายเป็นภาพเอกซเรย์ และภาพ CT ทรวงอกที่ถูกรวบรวมมาจากแหล่งที่มาสาธารณะ ซึ่งประกอบไปด้วย ภาพทรวงอกของผู้ป่วยโรคโควิด-19 ภาพทรวงอกในภาวะปกติ และภาพทรวงอกของผู้ป่วยโรคปอดอักเสบ

2.) ประยุกต์ใช้เทคนิคการประมวลผลภาพดิจิทัลต่าง ๆ ในการปรับปรุงคุณภาพสีของชุดข้อมูลภาพ

3.) ประยุกต์ใช้แบบจำลองการเรียนรู้เชิงลึกแบบ CNN และ RNN ต่าง ๆ ที่นิยมใช้กันอย่างแพร่หลาย ในการสร้างโครงข่ายประสาทเทียมสำหรับจำแนกสามภาวะโรคปอด ได้แก่ ภาวะโรคโควิด-19 ภาวะปกติ และภาวะโรคปอดอักเสบ

4.) ประเมินประสิทธิภาพของโครงข่ายในการจำแนกภาวะโรคปอดจากตัวบ่งชี้สี่ค่า ได้แก่ ค่าความแม่นยำ (Accuracy) ค่าความเที่ยงตรง (Precision) ค่าความระลึก (Recall) และค่าคะแนน F1 (F1-score) และเปรียบเทียบผลลัพธ์ประสิทธิภาพของแต่ละโครงข่าย

1.5 ขั้นตอนของการศึกษา

ขั้นตอนของการศึกษาการตรวจคัดกรองโรคโควิด-19 จากภาพเอกซเรย์ และภาพ CT ทรวงอก มี 7 ขั้นตอนดังต่อไปนี้

- 1.) ทบทวนวรรณกรรม และงานวิจัยที่เกี่ยวข้องที่เกี่ยวกับการตรวจคัดกรองโรคโควิด-19 จากภาพเอกซเรย์ และภาพ CT ทรวงอก
- 2.) กำหนดปัญหา และตั้งสมมุติฐานของวิธีการตรวจคัดกรองโรคโควิด-19 จากภาพเอกซเรย์ และภาพ CT ทรวงอก
- 3.) รวบรวมชุดข้อมูลภาพเอกซเรย์ และภาพ CT ทรวงอก จากแหล่งที่มาสาธารณะต่าง ๆ
- 4.) ศึกษา และเลือกใช้เทคนิคการประมวลผลภาพดิจิทัลต่าง ๆ เพื่อนำมาประยุกต์ใช้ในการปรับปรุงคุณภาพสีของชุดข้อมูลภาพ
- 5.) ศึกษา และพัฒนาโครงข่ายประสาทเทียมแบบ CNN และ RNN สำหรับจำแนกภาวะโรคปอดจากภาพเอกซเรย์ และภาพ CT ทรวงอก
- 6.) ประเมิน และเปรียบเทียบผลลัพธ์ประสิทธิภาพของแต่ละโครงข่ายในการจำแนกภาวะโรคปอดจากภาพเอกซเรย์ และภาพ CT ทรวงอก
- 7.) สรุปผลการทดลองการจำแนกภาวะโรคปอดจากภาพเอกซเรย์ และภาพ CT ทรวงอก

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

บทนี้จะกล่าวถึงทฤษฎี และวรรณกรรมที่เกี่ยวข้องกับวิทยานิพนธ์ฉบับนี้ อันประกอบไปด้วย รังสีวินิจฉัย การประมวลผลภาพดิจิทัล โครงข่ายประสาทเทียม การเรียนรู้เชิงลึก และทบทวนวรรณกรรม

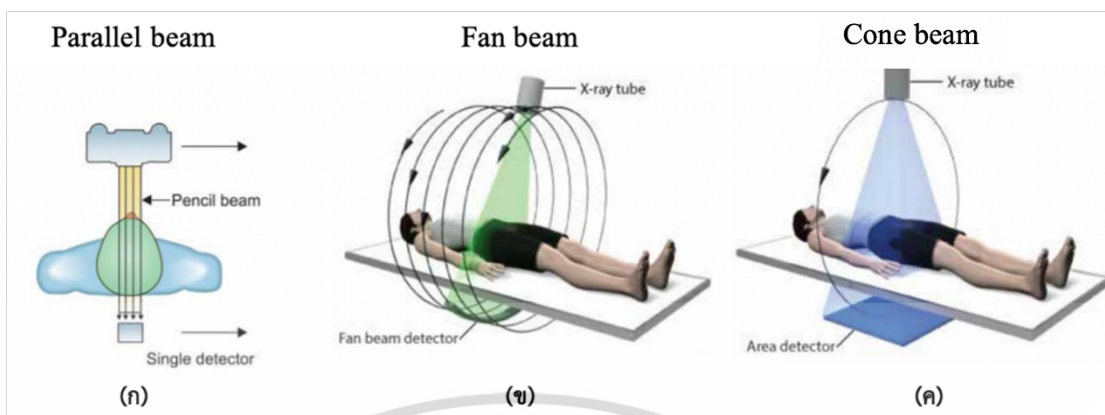
2.1 รังสีวินิจฉัย (Radiology)

รังสีวินิจฉัยในปัจจุบัน คือ วิชาการวินิจฉัยโรคโดยอาศัยภาพถ่ายอวัยวะภายในร่างกายมนุษย์ ด้วยพลังงานคลื่นชนิดต่าง ๆ กัน เพื่อถ่ายภาพความผิดปกติ หรือพยาธิสภาพของโรคภายในร่างกายมนุษย์ อันนำไปสู่การวินิจฉัยโรคได้อย่างถูกต้อง มีประสิทธิภาพและคุ้มค่า โดยสามารถยกตัวอย่างวิธีการทางเทคนิค ได้แก่ วิธีการตรวจโดยใช้คลื่นรังสีทำให้เกิดกระบวนการแตกตัวของไอออน (Ionizing radiation) เช่น การตรวจโดยภาพรังสีเอกซ์ (X-ray) การตรวจโดยเอกซเรย์คอมพิวเตอร์ (Computed tomography, CT) เป็นต้น [4]

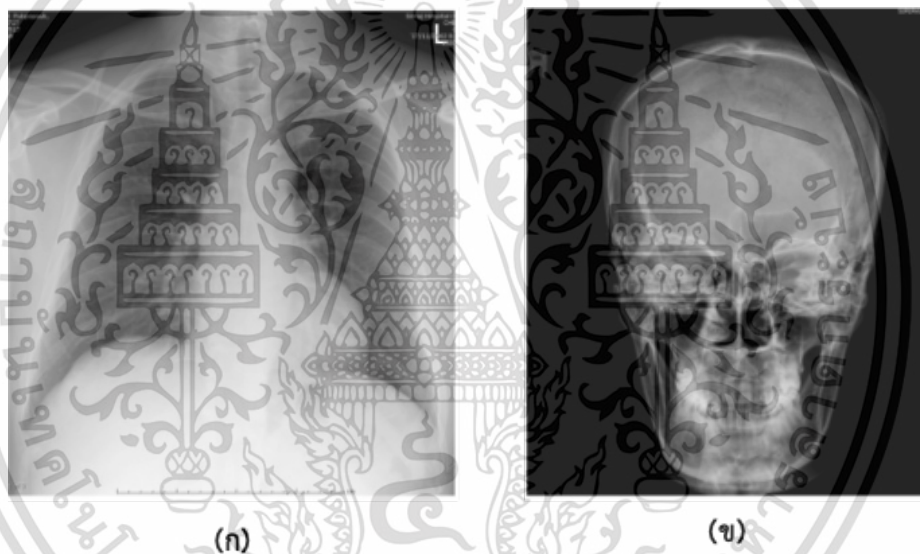
2.1.1 ภาพทางรังสี (Radiography)

2.1.1.1 ภาพเอกซเรย์ (X-ray) [5]

ภาพเอกซเรย์เกิดจากการใช้รังสีเอกซ์ซึ่งคือ คลื่นแม่เหล็กไฟฟ้าที่มีพลังงานอยู่ในช่วงระดับประมาณ 10 ถึง 100 keV มีความยาวคลื่นในช่วงประมาณ 0.01 ถึง 0.1 nm แล่งกำเนิดรังสีเอกซ์สำหรับใช้งานกับเครื่องสแกนของการถ่ายภาพเอกซเรย์คอมพิวเตอร์ ส่วนใหญ่เป็นหลอดกำเนิดรังสีเอกซ์ และในเครื่องถ่ายรุ่นใหม่จะเป็นเครื่องเร่งอนุภาคอิเล็กตรอน โดยทั่วไปรังสีเอกซ์ทำอันตรกิริยากับร่างกายส่วนที่เคลื่อนที่ผ่านได้สองรูปแบบ คือ การชนแบบโฟโตอิเล็กทริก และการชนแบบคอมพ์ตัน โดยสามารถที่จะเคลื่อนทะลุผ่านร่างกายไปยังระบบตรวจวัดได้ ระดับพลังงานและความเข้มของรังสีเอกซ์ควรเลือกให้เหมาะสมกับการใช้งาน โดยการปรับความต่างศักย์และกระแสไฟฟ้าของหลอดกำเนิดรังสีเอกซ์ เพื่อให้ได้ภาพถ่ายที่คมชัดเห็นถึงบริเวณส่วนที่ต้องการศึกษา แต่จะ คำนึงถึงความปลอดภัยทางรังสีและขีดจำกัดปริมาณรังสีที่ผู้ป่วยได้รับเป็นสิ่งสำคัญ นอกจากนี้รูปแบบลักษณะของลำรังสีเอกซ์ที่ออกมาจากแหล่งกำเนิดก็มีบทบาทสำคัญต่อคุณภาพ และกระบวนการในการสร้างภาพสามมิติ ซึ่งแบ่งเป็น 3 แบบ คือ ลำรังสีแบบขนาน (Parallel beam) ลำรังสีบานปลายแบบรูปพัด (Fan beam) และลำรังสีแบบโคนบีบ (Cone beam) ดังแสดงในรูปที่ 2.1 โดยมีตัวอย่างภาพเอกซเรย์ที่ได้ดังรูปที่ 2.2



รูปที่ 2.1 ลักษณะของลำรังสีเอกซ์ (ก) แบบขนาน (ข) แบบพัด และ (ค) แบบโคนปี่ม



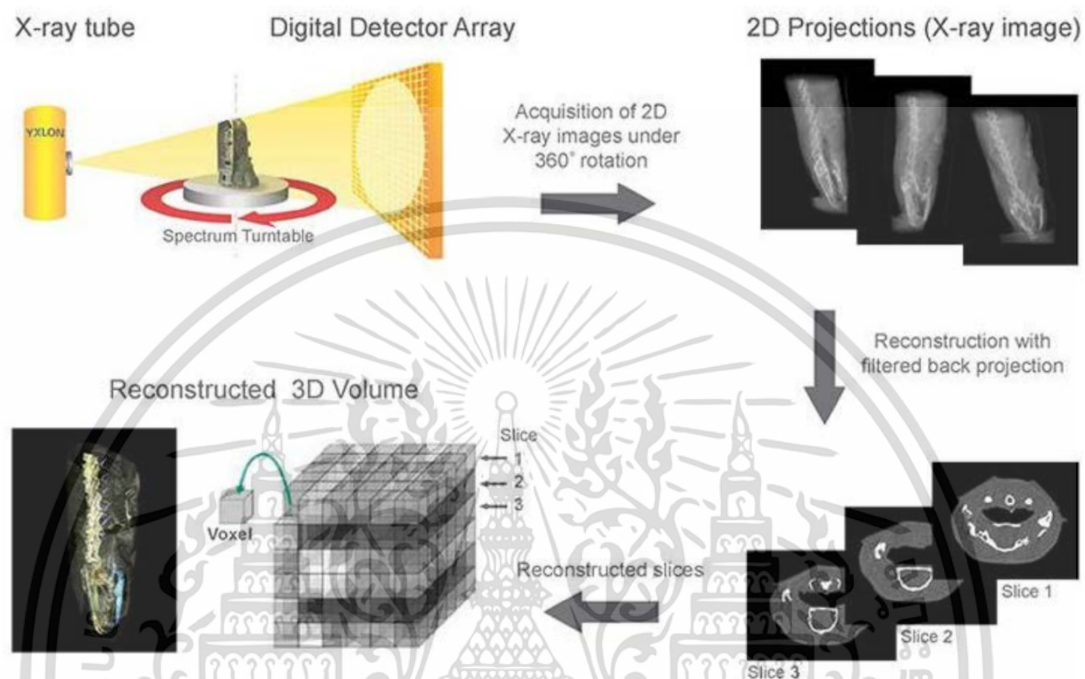
รูปที่ 2.2 ตัวอย่างภาพเอกซเรย์ (ก) ภาพเอกซเรย์ทรวงอก (ข) ภาพเอกซเรย์กะโหลกศีรษะ

2.1.1.2 ภาพเอกซเรย์คอมพิวเตอร์ (Computed tomography, CT) [5]

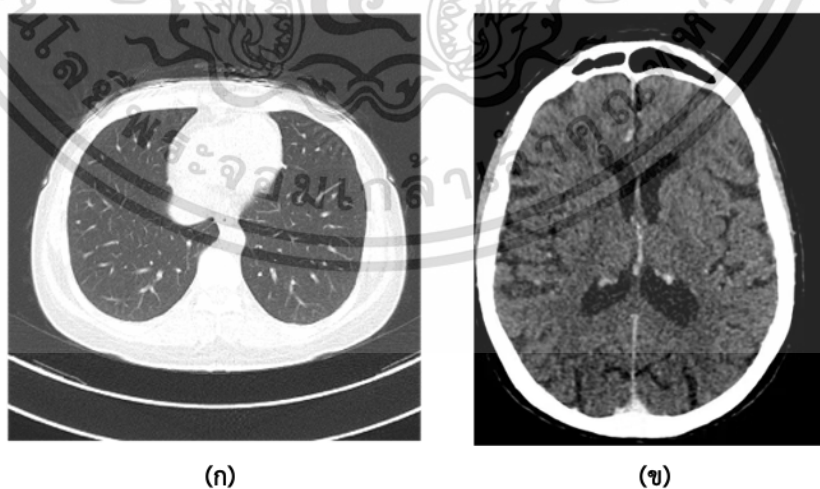
ภาพเอกซเรย์คอมพิวเตอร์ สร้างจากใช้ภาพฉายของรังสีเอกซ์หลายภาพจากมุมฉายที่แตกต่างกัน มารวบรวมประมวลผลใหม่เพื่อสร้างภาพภาคตัดขวาง โดยโปรแกรมคอมพิวเตอร์ ทำให้ได้เป็นรูปร่างในแนวมิติที่สามขึ้นมา คล้ายกับการนำก้อนชิ้นงานมาหันออกเป็นชิ้นบาง ๆ ให้มองเห็นภาพภาคตัดขวางภายใน ชุดภาพฉายแต่ละชุดของแต่ละแผ่นชิ้นงานของการถ่ายภาพเอกซเรย์คอมพิวเตอร์ได้มาจากการหมุนถ่ายภาพรอบชิ้นงาน หรือตัว ผู้ป่วย 360° หรือเพียงแค่ 180° ดังแสดงในรูปที่ 2.3 เมื่อฉายภาพที่มุมฉายต่างกันทำให้ได้ภาพฉายที่ได้จากมุมฉายแต่ละมุมก็จะต่างกัน จึงได้รับข้อมูลของวัตถุที่แตกต่างกัน คือถ้ายังมีจำนวนมุมถ่ายเยาะจะยิ่งได้ข้อมูลเกี่ยวกับวัตถุมากขึ้นเท่านั้น ดังนั้นการสแกนการถ่ายภาพสาม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

มิติแบบซีทีสแกน หรือการสร้างภาพแบบตัดขวางด้วยคอมพิวเตอร์ จึงมีบทบาทสำคัญในกระบวนการประมวลผลเพื่อการสร้างภาพจำลองโดยมีตัวอย่างภาพ CT ดังแสดงในรูปที่ 2.4



รูปที่ 2.3 ขั้นตอนการเก็บชุดข้อมูลภาพฉายรังสี และกระบวนการประมวลผลเพื่อสร้างภาพใหม่ของการถ่ายภาพเอกซเรย์คอมพิวเตอร์



รูปที่ 2.4 ตัวอย่างภาพ CT (ก) ภาพ CT ทรวงอก (ข) ภาพ CT กระโหลกศีรษะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยในปัจจุบันการตรวจด้วยภาพ CT สามารถใช้ตรวจระบบของร่างกายได้ 4 ระบบดังต่อไปนี้

- 1.) ระบบสมอง ได้แก่ การตรวจเอกซเรย์คอมพิวเตอร์ของสมอง ต่อมใต้สมอง ตา ต่อมไทรอยด์ และคอ เป็นต้น ในการตรวจนี้ จะต้องฉีดสารทึบรังสีเข้าทางหลอดเลือดดำ เพื่อช่วยให้เห็นพยาธิสภาพของโรคชัดเจนขึ้น
- 2.) ระบบช่องท้องและทรวงอก ได้แก่ การตรวจเอกซเรย์คอมพิวเตอร์ภายในช่องท้อง และอุ้งเชิงกราน ในการตรวจระบบช่องท้องบางกรณีรังสีแพทย์อาจพิจารณาให้ผู้ป่วยดื่มสารทึบ หรือสวนสารทึบเข้าทางทวารหนักเพื่อแยกลำไส้ออกจากเนื้อเยื่ออื่น ๆ ของช่องท้อง และในการตรวจระบบทรวงอกบางกรณีอาจจำเป็นต้องฉีดสารทึบรังสีเข้าทางหลอดเลือดดำเพื่อช่วยให้เห็นพยาธิสภาพของโรคชัดเจนขึ้น ทั้งนี้ขึ้นอยู่กับดุลยพินิจของรังสีแพทย์
- 3.) ระบบกระดูก กล้ามเนื้อ ข้อต่อและกระดูกสันหลัง ซึ่งมักใช้ในการวินิจฉัยโรคเนื้องอกของกล้ามเนื้อ กระดูก หรือการอักเสบของข้อต่อต่างๆ และลักษณะทางกายวิภาคของกระดูกสันหลัง โดยสามารถให้การวินิจฉัยโรคกระดูกได้ดีกว่าการตรวจเอกซเรย์ทั่วไป
- 4.) ระบบหลอดเลือด ได้แก่ การตรวจเอกซเรย์คอมพิวเตอร์ของหลอดเลือดสมอง หลอดเลือดหัวใจ หลอดเลือดแดงใหญ่ หลอดเลือดแดงไต และหลอดเลือดแดงที่ขา เป็นต้นในการตรวจนี้จำเป็นต้องฉีดสารทึบรังสีเข้าทางหลอดเลือดดำ

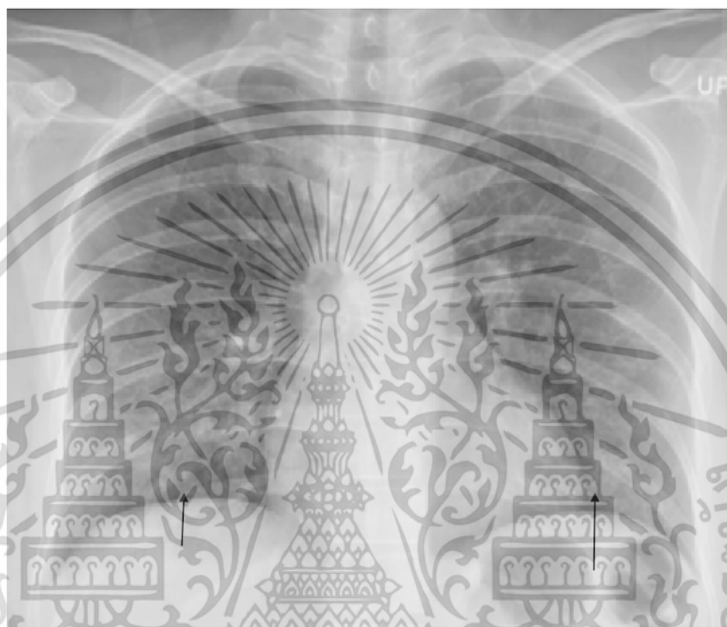
2.1.2 ภาพวินิจฉัยทางรังสีวิทยาในโควิด-19 [6,7]

โรคติดเชื้อไวรัสโคโรนาชนิดใหม่ หรือที่ องค์การอนามัยโลก (WHO) บัญญัติชื่อว่า โควิด-19 เป็นโรคอุบัติใหม่ โดยเริ่มกำเนิดจากเมืองอู่ฮั่น ประเทศสาธารณรัฐประชาชนจีน เมื่อเดือนธันวาคม 2019 และแพร่ขยายเป็นวงกว้างทั่วโลกอย่างรวดเร็ว จนองค์การอนามัยโลกให้คำจำกัดความของโรคนี้อีกครั้งเมื่อเดือนมีนาคม 2020 ว่าเป็นโรคระบาดใหญ่ทั่วโลก (pandemic) ณ ปัจจุบัน มีผู้ติดเชื้อ และเสียชีวิตทั่วโรคเป็นจำนวนมาก ผู้ติดเชื้อไวรัสโคโรนาชนิดใหม่นี้ส่วนใหญ่มีอาการทางระบบทางเดินหายใจ ตั้งแต่ ไอเล็กน้อย ปวดอักเสบ หรืออาจรุนแรงถึงภาวะการหายใจล้มเหลวและเสียชีวิตได้ ภาพวินิจฉัยทางรังสีวิทยาจึงมีบทบาทสำคัญทั้งในการช่วยวินิจฉัย ติดตามการรักษา และช่วยพยากรณ์โรค เนื่องจากเป็นโรคอุบัติใหม่ เครื่องมือและวิธีการต่างๆ ในการตรวจวินิจฉัยโรคจึงต้องมีการพัฒนาและปรับปรุงให้เหมาะสมกับการตรวจพบและการรักษาโรคนี้ให้มีประสิทธิภาพ นอกจากนี้ ความเข้าใจเกี่ยวกับการแพร่ระบาดและการป้องกันการติดเชื้อยังมีความสำคัญอย่างยิ่งในการควบคุมโรคและลดการแพร่ระบาดของไวรัสโคโรนาชนิดใหม่ในระดับสากล

ความผิดปกติในภาพเอกซเรย์ทรวงอกที่พบได้บ่อย ซึ่งใช้ยึดถือเป็นแบบฉบับ (Typical pattern) ในระยะแรก ๆ ที่มีการรายงานจากประเทศจีนคือ รอยโรคในถุงลม (Alveolar opacity) โดยรอยโรคนี้สามารถปรากฏเป็นฝ้าจาง (Ground glass opacity, GGO) โดยมีการกระจายอยู่ที่ปอดทั้งสองข้าง

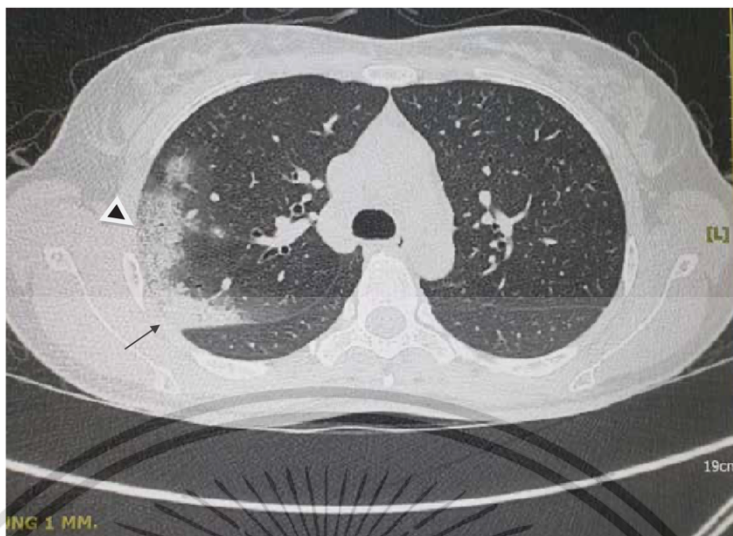
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังแสดงในรูปที่ 2.5 และพบบ่อยที่ปอดกลีบล่างมากกว่ากลีบบน ต่อมา เมื่อการระบาดแพร่กระจายไปทั่วโลก รายงานจากประเทศฝั่งตะวันตกพบว่ารอยโรคชนิด Interstitial opacity ที่มีลักษณะเป็นเส้น ๆ พบบ่อยกว่ารอยโรคชนิด Alveolar opacity การเปลี่ยนแปลงนี้อาจสะท้อนถึงความหลากหลายทางภูมิศาสตร์และประชากรที่ติดเชื้อ ซึ่งอาจมีปัจจัยที่ทำให้ภาพรังสีวิทยามีความแตกต่างกันในแต่ละพื้นที่



รูปที่ 2.5 ภาพเอกซเรย์ทรวงอกแสดงรอยโรคในถุงลมเป็นฝ้าจาง (ground glass opacity) โดยกระจายอยู่ที่ปอดกลีบล่างทั้งสองข้าง

ความผิดปกติในภาพเอกซเรย์คอมพิวเตอร์ หรือ ภาพ CT ทรวงอก เป็นเช่นเดียวกับภาพเอกซเรย์ทรวงอก แต่มีความไวมากกว่า โดยเอกซเรย์คอมพิวเตอร์ให้ความไวถึงร้อยละ 84-99 ขึ้นอยู่กับระยะเวลาการเจ็บป่วย ในบางรายงานจึงแนะนำให้ใช้การเอกซเรย์คอมพิวเตอร์ปอดเป็นเครื่องมือชนิดหนึ่งในการคัดกรองวินิจฉัยโรค COVID-19 เนื่องจากการส่งตรวจทางห้องปฏิบัติการเสียเวลามากกว่า นอกจากนี้หลายสถาบันใช้การเอกซเรย์คอมพิวเตอร์ปอดในการประเมินความรุนแรงของโรคและติดตามการรักษา ลักษณะของความผิดปกติในเนื้อปอดที่พบได้จากภาพเอกซเรย์คอมพิวเตอร์ที่พบได้บ่อยคือ GGO พบได้ร้อยละ 50 mixed pattern ร้อยละ 49-71 ดังแสดงในรูปที่ 2.6 ตำแหน่งในปอดที่พบความผิดปกติได้บ่อยคือ ที่ขอบชายปอด ร้อยละ 67-89 เป็นมากกว่า 1 กลีบของปอด ร้อยละ 81-93 เป็นที่ปอดทั้งสองข้าง ร้อยละ 59-91 พบที่ปอดส่วนหลังมากกว่าส่วนหน้า ร้อยละ 93 และพบที่ปอดส่วนล่างมากกว่าส่วนบน โดยกลีบปอดล่างขวาพบมากที่สุด ร้อยละ 93



รูปที่ 2.6 ภาพเอกซเรย์คอมพิวเตอร์ทรวงอกแสดงลักษณะเป็นฝ้าจาง (ground glass opacity) โดยพบรอยโรคที่ขอบชายปอด

การใช้ภาพถ่ายรังสี เช่น ภาพเอกซเรย์ และภาพ CT ทรวงอก ในการตรวจโรคโควิด-19 นั้นจึงเป็นอีกหนึ่งทางเลือกในการตรวจคัดกรอง และวินิจฉัย เนื่องจากมีลักษณะที่สามารถใช้ในการตรวจจับ และวินิจฉัยโรคได้อย่างรวดเร็ว และแม่นยำ อีกทั้งยังช่วยบอกความรุนแรงของโรค ติดตามการรักษา และพยากรณ์โรคได้ อย่างไรก็ตามความผิดปกติที่พบสามารถพบได้ในโรคปอดอักเสบ อื่น ๆ การแปลผลจึงต้องประกอบกับอาการทางคลินิก และผลตรวจทางห้องปฏิบัติการการเลือกใช้เครื่องมือทางรังสีวิทยาขึ้นอยู่กับปัจจัยทางด้านทรัพยากร และบริบทต่าง ๆ ของสถานพยาบาล

2.2 การประมวลผลภาพดิจิทัล (Digital image processing) [8-9]

การประมวลผลภาพดิจิทัลเป็นการใช้กระบวนการที่มากกระทำต่อข้อมูลภาพเพื่อให้ได้ซึ่งภาพที่มีคุณลักษณะตามต้องการ ทั้งในเชิงปริมาณและในเชิงคุณภาพ อาทิเช่น การกำจัดสัญญาณรบกวนออกจากภาพ การแบ่งส่วนของวัตถุที่สนใจออกมาจากภาพ โดยคอมพิวเตอร์นั้นมีความสามารถในการคำนวณ และประมวลผลข้อมูลได้อย่างรวดเร็ว จึงมีประโยชน์ในการเพิ่มประสิทธิภาพ และวิเคราะห์ข้อมูลที่ได้จากภาพในระบบ

การประมวลผลภาพมีจุดเริ่มต้นจาก ความต้องการในการส่งข้อมูลภาพ และการพิมพ์ภาพในอุตสาหกรรมหนังสือพิมพ์ โดยในช่วงหลังสงครามโลกครั้งที่หนึ่ง ช่วงปี ค.ศ. 1920 ได้มีความพยายามส่งภาพผ่านสายเคเบิลใต้น้ำในการติดตามข่าวจากพื้นที่ห่างไกล ซึ่งการส่งสัญญาณภาพในช่วงแรกมีปัญหาในเรื่องของระดับความเข้ม (Intensity levels) ของภาพ ส่งผลให้คุณภาพของภาพที่ได้ไม่ดีนัก ต่อมาจึงได้มี

การปรับปรุงโดยใช้เครื่องพิมพ์เทเลกราฟที่สื่อสารโดยใช้โทรสัญญาณไฟฟ้าในรูปแบบมอร์สโค้ด (Morse code) โดยอาศัยสัญลักษณ์ และพื้นที่ว่างระหว่างสัญลักษณ์เพื่อแสดงผล ซึ่งช่วยให้ได้ภาพที่มีคุณภาพสูง ขึ้น จากเหตุการณ์ข้างต้นเห็นได้ว่า นัยสำคัญของการประมวลผลภาพ คือ การสร้างภาพ หรือสร้างภาพ ขึ้นมาใหม่จากภาพเดิม เพื่อให้ภาพนั้นเหมาะสมต่อการนำไปประยุกต์ใช้งานในแต่ละประเภท โดยเฉพาะ ในยุคปัจจุบันที่มีการจัดเก็บข้อมูลภาพในรูปแบบดิจิทัล ไม่ว่าจะ เป็นกล้องดิจิทัล ที่เปลี่ยนมาจากกล้อง ฟิล์ม กล้องวงจรปิดที่จัดเก็บลงในจานบันทึกแบบแข็ง (Hard Disk) แทนการจัดเก็บลงในเทป (Tape) เป็นต้น ซึ่งการจัดเก็บภาพด้วยอุปกรณ์ต่าง ๆ เหล่านี้ล้วนต้องใช่วิธีการทางการประมวลผลภาพเข้าไปช่วย เพื่อให้ภาพนั้นสามารถจัดเก็บได้อย่างมีประสิทธิภาพ

2.2.1 ภาพดิจิทัล

ภาพดิจิทัล (Digital image) เกิดจากการเปลี่ยนภาพที่ได้จากการสะท้อนแสง (Light reflection) ลงบนเซ็นเซอร์รับภาพ (Image sensor) ซึ่งเป็นอุปกรณ์ที่ไวต่อการรับแสง (Photo sensitive devices) เช่น CCD/CMOS ให้เป็นข้อมูลภาพทางด้านดิจิทัล ซึ่งภาพที่เรามองเห็นบนจอภาพ หรือบนกระดาษเกิด จากการผสมสีเข้าด้วยกัน เช่น จอภาพใช้การรวมตัวของสีจากโมเดลสี RGB ซึ่งเป็นการรวมสี แดง เขียว และน้ำเงิน เข้าด้วยกัน เพื่อให้ได้สีตามที่ต้องการ เครื่องพิมพ์แบบฉีดหมึก (Ink jet) ใช้การรวมสีจาก โมเดลสี CMYK โดยอาศัยหลักการดูดซับแสง (Light absorption) ซึ่งจุดภาพแต่ละจุด ถูกเรียกว่า พิกเซล (Pixel) โดยเมื่อนำพิกเซลมาเรียงกันเป็นรูปทรงใด ๆ นั้นทำให้เกิดเป็นภาพในลักษณะต่าง ๆ ขึ้นมา กล่าวคือภาพเกิดจากเรียงตัวของพิกเซล ดังนั้นคุณภาพของภาพที่สร้างขึ้นมา จึงขึ้นอยู่กับคุณภาพของแต่ละ พิกเซล การประมวลผลภาพเป็นการนำตัวเลขในแต่ละพิกเซลมาวิเคราะห์ และประมวลผล ทั้งการ วิเคราะห์เพียงจุดภาพเดียว การวิเคราะห์เป็นกลุ่มพิกเซล หรือการวิเคราะห์ทั้งภาพ ซึ่งผลลัพธ์ที่ได้ในแง่ การคำนวณผลลัพธ์การประมวลผล อาจเป็นเมทริกซ์ (Matrix) เวกเตอร์ (Vector) หรือสเกลาร์ (Scalar) และในแง่ของการแสดงผลอาจเป็นภาพอีกหนึ่งภาพ กราฟ หรือค่าสี ขึ้นอยู่กับวิธีการที่ใช้ในการคำนวณ และการแสดงผล

ภาพดิจิทัลสามารถนิยามเป็นฟังก์ชันสองมิติ โดยที่ และเป็นพิกัดของจุดภาพและค่าของ แทน ระดับความสว่างที่พิกัด หากกำหนดให้ภาพ มีขนาดแถว และคอลัมน์ โดยมีจุดกำเนิดของภาพอยู่ที่ ตำแหน่ง (0,0) สามารถเขียนในรูปเมทริกซ์ได้ดังรูปที่ 2.7

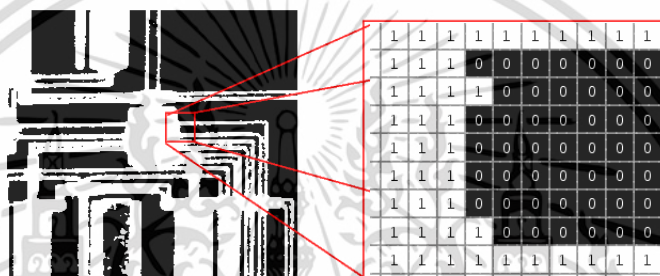
$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0, N-1) \\ f(1,0) & f(1,1) & \dots & f(1, N-1) \\ f(M-1,0) & f(M-1,1) & \dots & f(M-1, N-1) \end{bmatrix}$$

รูปที่ 2.7 เมทริกซ์ภาพดิจิทัล

โดยทั่วไปรูปแบบภาพที่มักนำมาใช้ในการประมวลผลภาพมีสามรูปแบบตามหลักการแสดงสีของคอมพิวเตอร์ ได้แก่ ภาพขาวดำ ภาพระดับเทา และภาพสีในปริภูมิสีแบบ RGB

2.2.1.1 ภาพขาวดำ (Binary image)

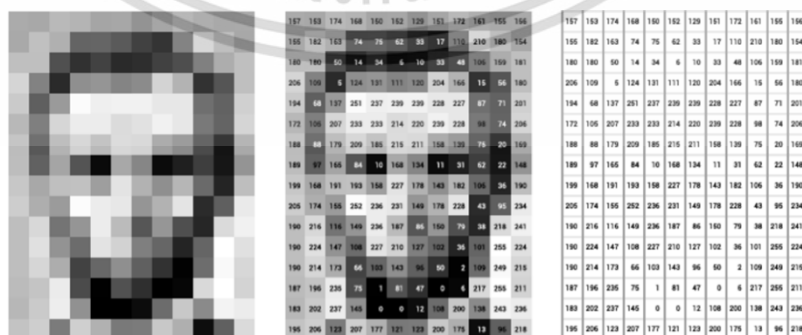
ภาพขาวดำ เป็นภาพที่แสดงข้อมูลเพียงสองสี ได้แก่ สีขาวและสีดำ กล่าวคือ ในหนึ่งจุดภาพมีโอกาสแสดงข้อมูลเป็นสีขาวหรือสีดำเท่านั้น ซึ่งภาพขาวดำถูกแทนค่าความสว่างด้วยเลขฐานสอง จำนวน 1 บิต คือ 0 หมายถึงสีดำ และ 1 หมายถึงสีขาว โดยแสดงตัวอย่างได้ดังรูปที่ 2.8



รูปที่ 2.8 ภาพขาวดำ และค่าความสว่างของแต่ละจุดภาพ [10]

2.2.1.2 ภาพระดับสีเทา (Gray scale image)

ภาพระดับสีเทา คือ ภาพที่มีค่าระดับความเข้มแสงในแต่ละจุดภาพแตกต่างกันขึ้นอยู่กับจำนวนบิต โดยแสดงความเข้มแสงตั้งแต่โทนสีขาว ไปจนถึงโทนสีดำ ภาพระดับสีเทามักถูกกำหนดให้มีค่าระดับความเข้มแสงเท่ากับ 8 บิต กล่าวคือ มีระดับความเข้มแสงทั้งหมด 256 ระดับ หรืออยู่ในช่วง 0 ถึง 255 โดยที่ค่าระดับความเข้มแสงเท่ากับ 0 คือ สีดำ และค่าระดับความเข้มแสงเท่ากับ 255 คือ สีขาว ตัวอย่างภาพระดับสีเทาแสดงดังรูปที่ 2.9

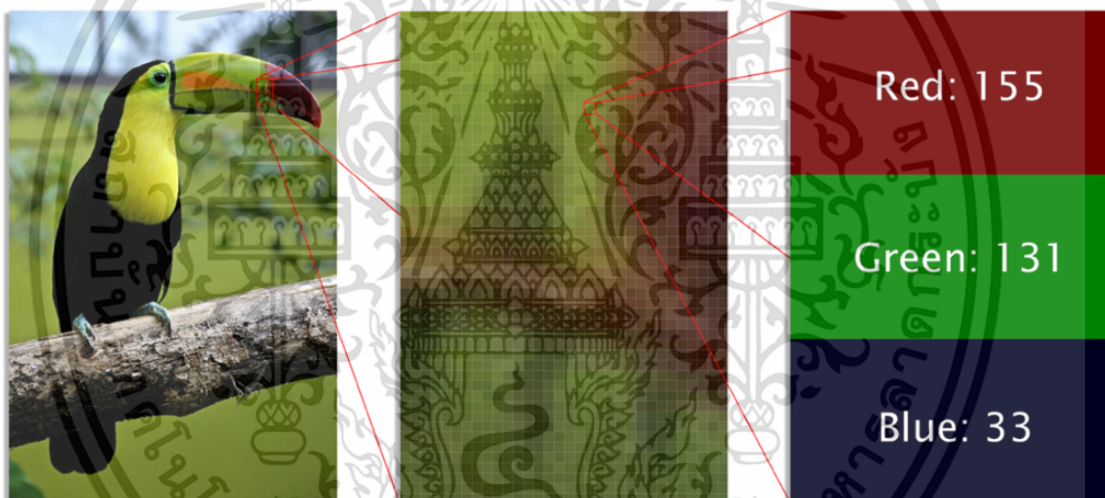


รูปที่ 2.9 ภาพระดับสีเทา และค่าความเข้มแสงของแต่ละจุดภาพ [11]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.1.3 ภาพสีแบบ RGB (RGB image)

ภาพสีแบบ RGB คือ ภาพที่ความเข้มสีในแต่ละจุดภาพเกิดจากการรวมตัวกันของ สาม ช่องสัญญาณสี ได้แก่ ช่องสัญญาณสีแดง (Red channel) ช่องสัญญาณสีเขียว (Green channel) และ ช่องสัญญาณสีน้ำเงิน (Blue channel) ที่มีค่าความเข้มอยู่ในช่วง 0 ถึง 255 เรียงจากมืดไปสว่าง ซึ่ง ข้อมูลภาพในแต่ละช่องสัญญาณสีใช้พื้นที่ในการจัดเก็บ 8 บิต ดังนั้น ในแต่ละจุดภาพของภาพสีจึงใช้พื้นที่ ในการจัดเก็บ 24 บิต ซึ่งสามารถแสดงสีได้ประมาณ ล้านสี หรือ 16,777,216 สี โดยสีต่าง ๆ นั้นเกิดจาก การรวมกันของสามช่องสัญญาณสีในสัดส่วนความเข้มที่แตกต่างกันตามหลักของโมเดลสีแบบบวก (Additive Color Model) ดังแสดงในรูปที่ 2.10 โดยหากในจุดภาพใดมีค่าความเข้มของทั้งสาม ช่องสัญญาณสีเท่ากับ 0 จุดภาพนั้นจะมีสีดำ แต่หากจุดภาพใดมีค่าความเข้มของทั้งสามช่องสัญญาณสี เท่ากับ 255 จุดภาพนั้นจะมีสีขาว



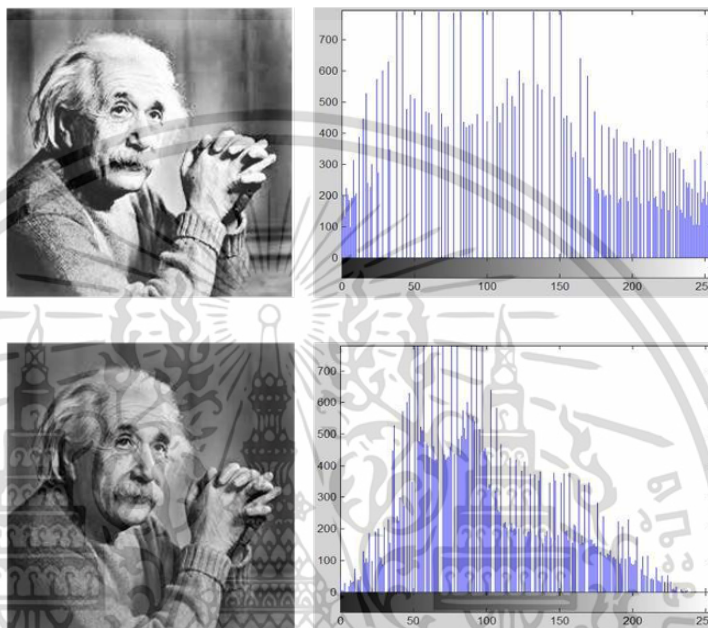
รูปที่ 2.10 ภาพสีแบบ RGB และค่าความเข้มของแต่ละช่องสัญญาณสีในหนึ่งจุดภาพ [12]

2.2.2 ฮิสโตแกรมและคอนทราสต์ของภาพ [13]

2.2.2.1 ฮิสโตแกรมของภาพ (Image Histogram)

ฮิสโตแกรม คือ เครื่องมือตัวหนึ่งที่นิยมใช้ในการนำเสนอข้อมูลในลักษณะกราฟแท่ง ซึ่งแสดง ความสัมพันธ์ของจำนวนพิกเซล ในระดับสีนั้น ๆ ใช้เพื่อวิเคราะห์ว่าการแจจจ หรือ การกระจายข้อมูล แบบใดมีผลต่อคุณภาพในทิศทางใด ซึ่งในแกนนอนของฮิสโตแกรมแสดงค่าความสว่างที่ตั้งแต่ 0-255 กล่าวคือ ค่าทางด้านซ้ายของกราฟ มีผลทำให้ภาพมีสีเข้มเข้าใกล้สีดำ ส่วนค่าทางด้านขวาของกราฟ มีผล ทำให้ภาพสว่างเข้าใกล้สีขาว ส่วนบริเวณกลางของกราฟแสดงส่วนน้ำหนักรูปกลาง และแนวแกนตั้งแสดง จำนวนพิกเซลของค่าความสว่างนั้น ๆ ถ้าหากภาพมีความมืดมาก กราฟที่ได้จะมีข้อมูลไปรวมกลุ่มกัน

ทางด้านซ้าย หรืออาจกล่าวได้ว่า ฮิสโตแกรมแสดงให้เห็นถึงความเบี่ยงเบนของข้อมูลว่า มีลักษณะการกระจายตัวของข้อมูลเป็นแบบระฆังคว่ำหรือไม่ หรือมีความเบี่ยงเบนไปทางบวกหรือลบ หรือมีลักษณะรูปแบบพื้นเลื้อยหรือแบบหน้าผา ซึ่งลักษณะของความเบี่ยงเบนแบบต่าง ๆ ทำให้ภาพที่ได้แตกต่างกันไป ตัวอย่างของกราฟฮิสโตแกรมแสดงในรูปที่ 2.11

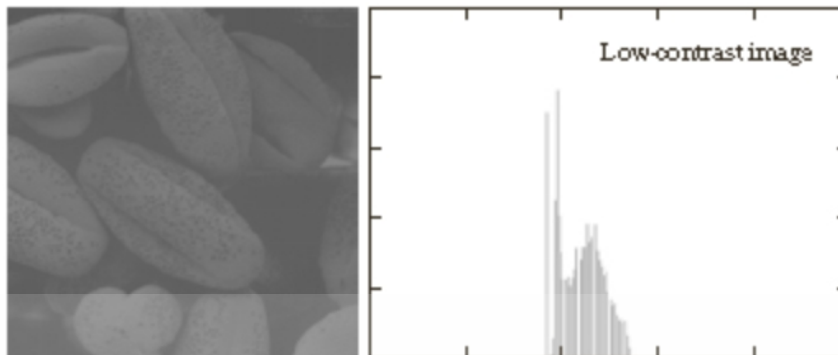


รูปที่ 2.11 กราฟฮิสโตแกรมของภาพ

2.2.2.2 คอนทราสต์ของภาพ (Contrast)

คอนทราสต์ เป็นค่าความเปรียบต่างของแสงหรือสีที่ทำให้มองเห็นวัตถุในภาพได้ อย่างชัดเจน โดยในลักษณะการมองเห็นภาพที่มีคอนทราสต์ สามารถกำหนดได้จากความต่างของความสว่างและสีของวัตถุที่มีความแตกต่างจากวัตถุอื่น ๆ ที่อยู่โดยรอบ ดังนั้นค่าความเข้มของความสว่าง และค่าความเข้มของสีของวัตถุจะส่งผลให้ภาพมีลักษณะคอนทราสต์ต่ำ และคอนทราสต์สูง แตกต่างกันไป โดยรายละเอียดของคอนทราสต์แต่ละลักษณะ มีดังนี้

1. ภาพคอนทราสต์ต่ำ (Low contrast image) เป็นภาพที่มีฮิสโตแกรมของค่าความสว่างกระจุกกันอยู่ในช่วงแคบ ๆ ซึ่งมีค่าความเปรียบต่างของจุดความสว่างสูงสุดกับจุดความสว่างต่ำสุด ห่างกันน้อย กล่าวคือ เป็นภาพที่มีความสว่างค่อนข้างสม่ำเสมอทั้งภาพ ดังแสดงในรูปที่ 2.12



รูปที่ 2.12 ตัวอย่างภาพคอนทราสต์ต่ำ

2. ภาพคอนทราสต์สูง (High contrast image) เป็นภาพที่มีฮิสโทแกรมของค่าความสว่างกระจายกันอยู่ในช่วงกว้าง หรือเป็นภาพที่มีค่าความเปรียบต่างของจุดความสว่างสูงสุดกับจุดความสว่างต่ำสุด ห่างกันมาก ดังแสดงในรูปที่ 2.13



รูปที่ 2.13 ตัวอย่างภาพคอนทราสต์สูง

2.2.3 ตัวดำเนินการสำหรับการประมวลผลภาพ (Image processing operators)

ตัวดำเนินการสำหรับการประมวลผลภาพ เป็นการใช้กระบวนการ หรือวิธีการทางคณิตศาสตร์ มาประยุกต์ใช้ในการกระทำต่อข้อมูลภาพเพื่อให้ได้ซึ่งภาพที่มีคุณลักษณะตามต้องการ อาทิเช่น การดำเนินการทางเลขคณิตของภาพ การแปลงเรขาคณิตของภาพ และการทำสังวัตนาการ เป็นต้น

2.2.3.1 การดำเนินการทางเลขคณิตของภาพ (Image arithmetic) [14]

การดำเนินการทางเลขคณิตของภาพ เป็นการนำตัวดำเนินการพื้นฐานทางคณิตศาสตร์ หรือ ตรรกะทางคณิตศาสตร์ มาประยุกต์ใช้ระหว่างภาพจำนวนสองภาพขึ้นไป หรือ ระหว่างภาพกับค่าคงที่ ซึ่ง

ข้อมูลภาพผลลัพธ์ที่ได้จากการดำเนินการทางเลขคณิตของภาพนั้นจะเป็นข้อมูลที่มีความสอดคล้องกับข้อมูลภาพขาเข้าเท่านั้น

1.) การดำเนินการทางภาพแบบบวก (Element wise addition) เป็นการนำข้อมูลภาพระหว่างสองภาพขึ้นไปที่มีขนาดเท่ากัน หรือ ระหว่างข้อมูลภาพกับค่าคงที่ มาสร้างเป็นภาพผลลัพธ์ โดยข้อมูลในแต่ละพิกเซลของภาพผลลัพธ์ เกิดจากผลรวมของข้อมูลทั้งสองในแต่ละพิกเซล โดยมีตัวอย่างการดำเนินการทางภาพแบบบวกระหว่างภาพกับค่าคงที่ ดังแสดงในรูปที่ 2.14



รูปที่ 2.14 ตัวอย่างการดำเนินการทางภาพแบบบวกระหว่างภาพกับค่าคงที่ (ก) ภาพต้นฉบับ (ข) ภาพผลลัพธ์หลังผ่านการดำเนินการทางภาพแบบบวก

2.) การดำเนินการทางภาพแบบหักล้าง (Element wise subtraction) เป็นการนำข้อมูลภาพระหว่างสองภาพขึ้นไปที่มีขนาดเท่ากัน หรือ ระหว่างข้อมูลภาพกับค่าคงที่ มาสร้างเป็นภาพผลลัพธ์ โดยข้อมูลในแต่ละพิกเซลของภาพผลลัพธ์ เกิดจากผลต่างของข้อมูลทั้งสองในแต่ละพิกเซล โดยมีตัวอย่างการดำเนินการทางภาพแบบหักล้างระหว่างภาพกับค่าคงที่ ดังแสดงในรูปที่ 2.15



รูปที่ 2.15 ตัวอย่างการดำเนินการทางภาพแบบหักล้างระหว่างภาพกับค่าคงที่ (ก) ภาพต้นฉบับ (ข) ภาพผลลัพธ์หลังผ่านการดำเนินการทางภาพแบบหักล้าง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.) การดำเนินการทางภาพแบบคูณ (Element wise multiplication) เป็นการนำข้อมูลภาพระหว่างสองภาพขึ้นไปที่มีขนาดเท่ากัน หรือ ระหว่างข้อมูลภาพกับค่าคงที่ มาสร้างเป็นภาพผลลัพธ์ โดยข้อมูลในแต่ละพิกเซลของภาพผลลัพธ์ เกิดจากผลคูณของข้อมูลทั้งสองในแต่ละพิกเซล โดยมีตัวอย่างการดำเนินการทางภาพแบบคูณระหว่างภาพกับค่าคงที่ ดังแสดงในรูปที่ 2.16



รูปที่ 2.16 ตัวอย่างการดำเนินการทางภาพแบบคูณระหว่างภาพกับค่าคงที่ (ก) ภาพต้นฉบับ (ข) ภาพผลลัพธ์หลังผ่านการดำเนินการทางภาพแบบคูณ

4.) การดำเนินการทางภาพแบบหาร (Element wise division) เป็นการนำข้อมูลภาพระหว่างสองภาพขึ้นไปที่มีขนาดเท่ากัน หรือ ระหว่างข้อมูลภาพกับค่าคงที่ มาสร้างเป็นภาพผลลัพธ์ โดยข้อมูลในแต่ละพิกเซลของภาพผลลัพธ์ เกิดจากผลหารของข้อมูลทั้งสองในแต่ละพิกเซล โดยมีตัวอย่างการดำเนินการทางภาพแบบหารระหว่างภาพกับค่าคงที่ ดังแสดงในรูปที่ 2.17



รูปที่ 2.17 ตัวอย่างการดำเนินการทางภาพแบบหารระหว่างภาพกับค่าคงที่ (ก) ภาพต้นฉบับ (ข) ภาพผลลัพธ์หลังผ่านการดำเนินการทางภาพแบบหาร

2.2.4 การปรับปรุงภาพเชิงรังสี (Radiometric image enhancement)

การปรับปรุงภาพเชิงรังสี เป็นกระบวนการในการแปลงข้อมูลค่าความสว่างของภาพ เพื่อสร้างภาพที่เน้นรายละเอียดที่ต้องการ ซึ่งพิจารณาค่าความสว่างของจุดภาพเดี่ยว (Individual pixel) ที่ไม่อ้างอิงจุดภาพข้างเคียง โดยมีวัตถุประสงค์เพื่อแบ่งระดับ หรือเพิ่มความเปรียบต่างค่าความสว่างของข้อมูลภาพที่ต้องการปรากฏให้เห็นชัดเจนมากขึ้นกว่าเดิม การปรับปรุงภาพเชิงรังสีมีหลายเทคนิคด้วยกัน ซึ่งการปรับปรุงภาพแต่ละวิธีนั้นช่วยในการแปลภาพในแง่มุมที่แตกต่างกันออกไป โดยสามารถแบ่งประเภทของวิธีการต่าง ๆ ได้ 3 ประเภท ดังต่อไปนี้

2.2.4.1 การปรับปรุงความสว่างของภาพแบบเชิงเส้น

การปรับปรุงความสว่างของภาพแบบเชิงเส้น (Linear contrast stretch) เป็นการขยายพิสัยค่าความสว่างของข้อมูลภาพเดิมให้มากขึ้นจนเต็มช่วง 0 ถึง 255 ซึ่งคำนวณได้จากสมการที่ (1) โดยมีสัดส่วนการปรับเทียบอยู่ในลักษณะกราฟเส้นตรง ดังแสดงในรูปที่ 2.18 และแสดงตัวอย่างการเน้นความคมชัดของภาพแบบเชิงเส้นได้ดังรูปที่ 2.19

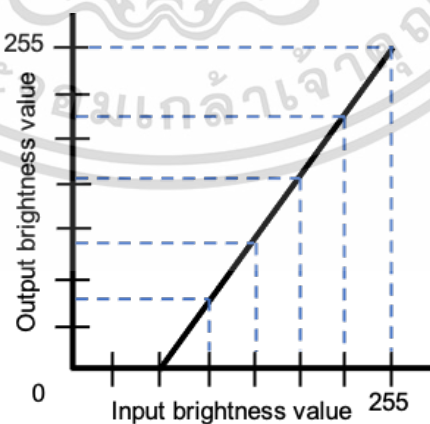
$$g(x, y) = \left(\frac{255}{DN_{\max} - DN_{\min}} \right) (f(x, y) - DN_{\min}) \quad (1)$$

เมื่อ $g(x, y)$ คือ ค่าความสว่างใหม่ของภาพที่จุด (x, y)

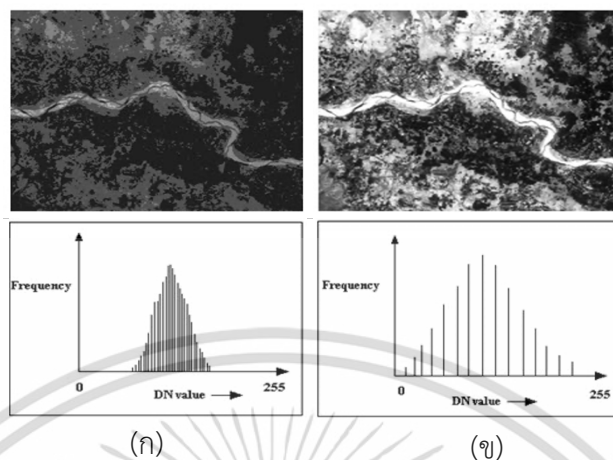
$f(x, y)$ คือ ค่าความสว่างเดิมของภาพที่จุด (x, y)

DN_{\max} คือ ค่าความสว่างสูงสุด

DN_{\min} คือ ค่าความสว่างต่ำสุด



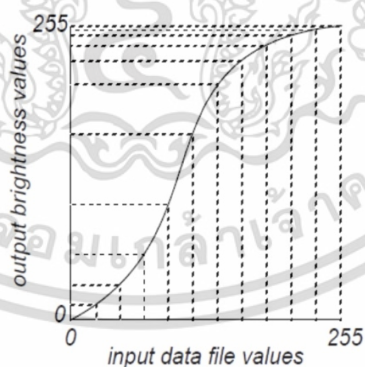
รูปที่ 2.18 กราฟความสัมพันธ์ของการขยายพิสัยค่าความสว่างแบบเชิงเส้น



รูปที่ 2.19 การปรับปรุงความสว่างของภาพแบบเชิงเส้น (ก) ภาพต้นฉบับ (ข) ภาพผลลัพธ์การปรับปรุงความสว่างแบบเชิงเส้น

2.2.4.2 การปรับปรุงความสว่างของภาพแบบไม่เป็นเชิงเส้น

การปรับปรุงความสว่างของภาพแบบไม่เป็นเชิงเส้น (Non-linear contrast stretch) เป็นการปรับระดับค่าความสว่างของข้อมูลภาพ โดยใช้ความสัมพันธ์ที่ไม่ใช่เชิงเส้นตรงดังแสดงในรูปที่ 2.20 เพื่อเปลี่ยนการกระจายข้อมูลที่ไม่แบบปกติให้เป็นแบบปกติ และปรับจำนวนจุดภาพในแต่ละค่าความเข้มให้มีจำนวนใกล้เคียงกัน ตัวอย่างเทคนิคการปรับปรุงความสว่างของภาพแบบไม่เป็นเชิงเส้นมีดังต่อไปนี้



รูปที่ 2.20 กราฟความสัมพันธ์ของการขยายค่าความสว่างของภาพแบบไม่เป็นเชิงเส้น

1.) Histogram Equalization (HE)

Histogram Equalization เป็นวิธีที่ใช้เพื่อเพิ่มคอนทราสต์ของรูปภาพ ด้วยการกระจายความเข้มของพิกเซลในภาพ โดยมีจุดประสงค์ คือ การสร้างภาพที่ความถี่ของความสว่างเสมอกันในแต่ละระดับ

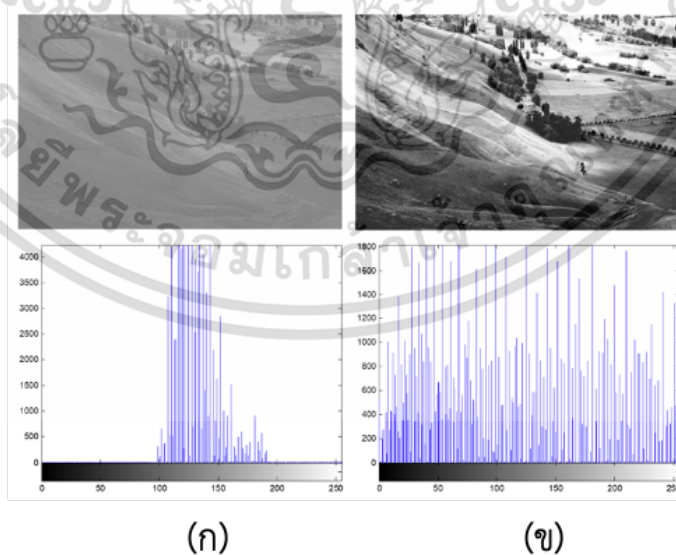
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ความสว่างของภาพ เทคนิคนี้ทำงานโดยการคำนวณฮิสโตแกรมของภาพต้นฉบับ ซึ่งเป็นกราฟที่แสดงจำนวนพิกเซลในแต่ละระดับความเข้ม จากนั้นฟังก์ชันการแจกแจงสะสม (Cumulative Distribution Function, CDF) ของฮิสโตแกรมจะถูกคำนวณ ซึ่งจะให้ความน่าจะเป็นสะสมของแต่ละค่าความเข้ม จากนั้น CDF จะถูกทำให้เป็นมาตรฐานเพื่อครอบคลุมช่วงค่าความเข้มที่เป็นไปได้ทั้งหมด ในลำดับสุดท้ายค่าพิกเซลในภาพต้นฉบับจะถูกจับคู่กับค่าใหม่โดยใช้ CDF ที่ปรับให้เป็นมาตรฐาน การจับคู่นี้ช่วยให้แน่ใจว่ารูปภาพใหม่มีการกระจายความเข้มของพิกเซลที่สมดุลมากขึ้น ซึ่งสมการคำนวณการแปลงของวิธีการนี้แสดงได้ดังสมการที่ (2) และแสดงตัวอย่างการปรับปรุงความสว่างของ Histogram Equalization ได้ดังรูปที่ 2.21

$$T(r) = \text{round}((L - 1) \times CDF(r)) \quad (2)$$

เมื่อ T คือ ฟังก์ชันการแปลง
 r คือ ระดับความเข้มของพิกเซล
 L คือ จำนวนระดับความเข้มที่เป็นไปได้
 $CDF(r)$ คือ ค่า CDF ที่เป็นมาตรฐานของความเข้ม r

ฟังก์ชันการแปลงจะจับคู่ความเข้มของภาพที่ป้อนเข้ากับค่าใหม่ในภาพผลลัพธ์ และ *round* ถูกใช้เพื่อให้แน่ใจว่าความเข้มของพิกเซลใหม่เป็นจำนวนเต็มในช่วง 0 ถึง $L - 1$

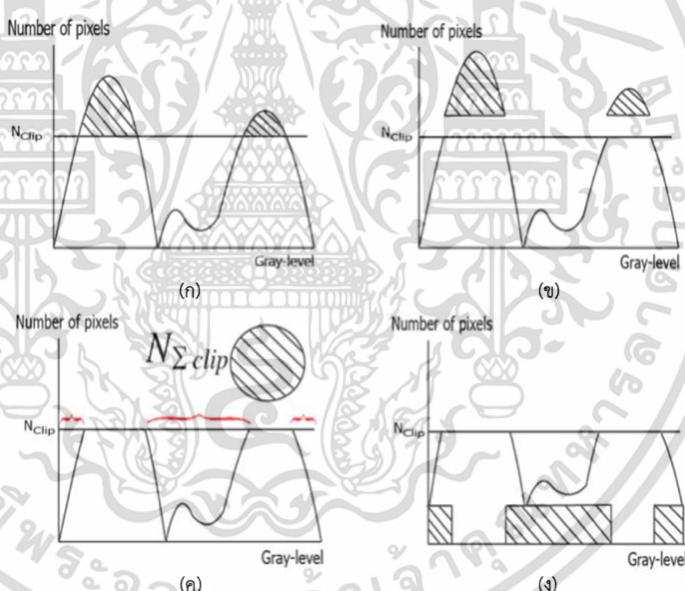


รูปที่ 2.21 การปรับปรุงความสว่างแบบ Histogram Equalization (ก) ภาพต้นฉบับ (ข) ภาพผลลัพธ์ การปรับปรุงความสว่างแบบ Histogram Equalization

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

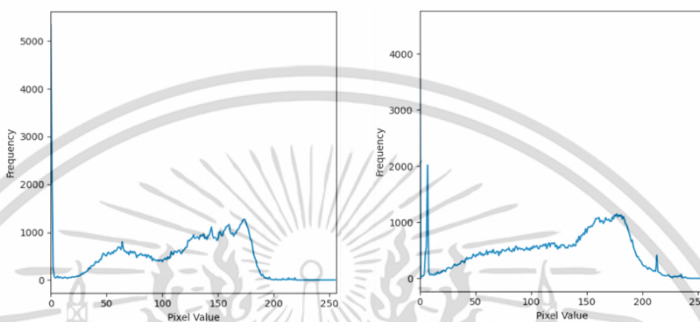
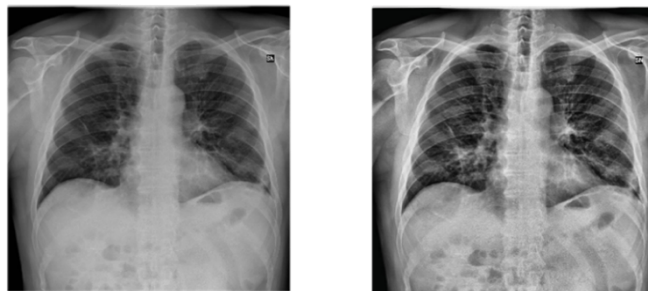
2.) Contrast Limited Adaptive Histogram Equalization (CLAHE)

Contrast Limited Adaptive Histogram Equalization เป็นรูปแบบหนึ่งของ Adaptive Histogram Equalization (AHE) ที่ใช้เพื่อเพิ่มคอนทราสต์ของภาพในขณะที่จำกัดปริมาณการขยายเพื่อป้องกันสัญญาณรบกวนที่มากเกินไป CLAHE ทำงานโดยการแบ่งภาพขาเข้าออกเป็นภาพย่อยขนาดเล็กที่ไม่ซ้อนทับกัน จากนั้นจึงทำการ Histogram Equalization ในแต่ละภาพย่อย โดยจะกระจายความสว่างของแต่ละภาพย่อยใหม่เพื่อขยายช่วงระดับความสว่างทั้งหมด ส่งผลให้คอนทราสต์ดีขึ้น อย่างไรก็ตาม เพื่อป้องกันการเพิ่มสัญญาณรบกวนที่มากเกินไป ฟังก์ชันจำกัดคอนทราสต์ถูกนำมาใช้กับจำนวนพิกเซล หลังปรับปรุง ฟังก์ชันจำกัดคอนทราสต์จะจำกัดช่วงของจำนวนพิกเซลในระดับค่าความสว่าง เพื่อป้องกันไม่ให้ถูกขยายมากเกินไป ซึ่งการจำกัดช่วงค่าความสว่างนั้นถูกควบคุมโดยพารามิเตอร์ที่เรียกว่า Clip limit (CL) หากจำนวนของค่าพิกเซลที่กำหนดให้กับระดับความสว่างเกินขีดจำกัดของ CL ค่าพิกเซลจะถูกกระจายไปยังระดับความสว่างอื่นที่ไม่เกินขีดจำกัด ดังแสดงในรูปที่ 2.22



รูปที่ 2.22 การกระจายความถี่ของพิกเซลที่เกินขีดจำกัดของ Clip limit (ก) ฮิสโตแกรมต้นฉบับ (ข) การจำกัดฮิสโตแกรมที่ระดับ N_{clip} (ค) การรวมค่าความถี่สะสม (ง) กระจายความถี่สะสมไปยังระดับที่ไม่เกินค่า N_{clip}

ตัวอย่างภาพหลังผ่านกระบวนการปรับรูปแบบ Contrast Limited Adaptive Histogram Equalization ดังแสดงในรูปที่ 2.23



รูปที่ 2.23 ตัวอย่างภาพและฮิสโตแกรมของภาพหลังผ่านกระบวนการปรับปรุงแบบ Contrast Limited Adaptive Histogram Equalization (ก) ภาพต้นฉบับ (ข) ภาพผลลัพธ์การปรับปรุง

3.) Gamma correction

Gamma correction เป็นการแปลงแบบไม่เป็นเชิงเส้นที่ดำเนินการเพื่อปรับความสว่างของภาพบนพิคเซลของภาพ และปรับสภาพความอึมตัวของภาพตามนั้น พร้อมรักษาค่าแกมมาให้คงที่ ซึ่งมีสมการคำนวณดังแสดงในสมการที่ (3)

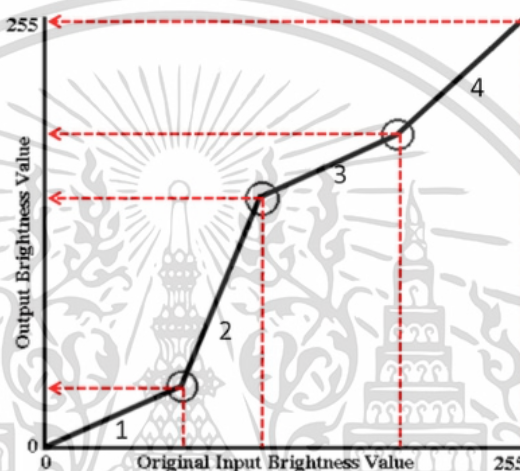
$$V_{out} = V^\gamma \quad (3)$$

เมื่อ V คือ ค่าความสว่างตั้งต้น
 V_{out} คือ ค่าความสว่างผลลัพธ์ในช่วง $[0,1]$
 γ คือ ค่ากำลังแกมมา ที่มีค่าอยู่ในช่วงลบอนันต์ ถึง อนันต์

หาก γ มีค่ามากกว่า 1 ภาพผลลัพธ์จะมีความสว่างมากขึ้น ในทางกลับกัน ถ้า γ มีค่าน้อยกว่า 1 ภาพผลลัพธ์จะมีความสว่างน้อยลง

2.2.4.3 การปรับปรุงความสว่างของภาพแบบช่วง

การปรับปรุงความสว่างของภาพแบบช่วง (Piecewise contrast stretch) เป็นการเลือกยืดค่าความสว่างของข้อมูลเป็นช่วงที่เฉพาะเจาะจง (Specific portion of data) โดยแต่ละช่วงอาจกำหนดพิสัยของการยืดแตกต่างกันไป หลักการคือ พิสัยของระดับค่าความสว่างของข้อมูลเดิมที่ต้องการขยายนั้น ถูกแบ่งออกเป็นหลาย ๆ ช่วง และในแต่ละช่วงของการขยายเป็นค่าใด ๆ ตามต้องการ แต่ต้องต่อเนื่องกันไปจนตลอดช่วง 0 ถึง 255 โดยแสดงความสัมพันธ์ได้ดังรูปที่ 2.24



รูปที่ 2.24 กราฟความสัมพันธ์ของการปรับปรุงความสว่างของภาพแบบช่วง

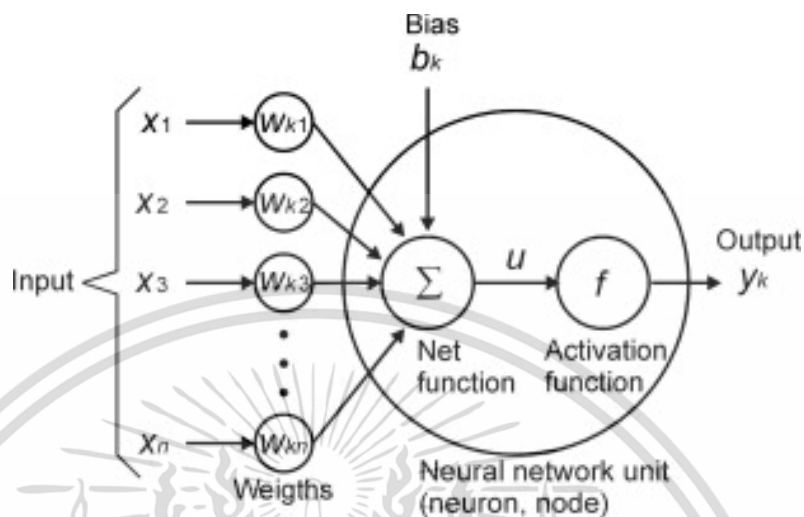
2.3 โครงข่ายประสาทเทียม (Neural network)

2.3.1 องค์ประกอบของโครงข่ายประสาทเทียม [15-16]

โครงข่ายประสาทเทียม หรือ ช่างงานประสาทเทียม คือ กระบวนการทางคณิตศาสตร์สำหรับประมวลผลข้อมูล โดยมีต้นแบบการทำงานจำลองมาจากโครงสร้างระบบประสาทของมนุษย์ เพื่อให้ได้เครือข่ายการทำงานที่สามารถจดจำรูปแบบของข้อมูล (Pattern recognition) และการอนุมานความรู้ (Knowledge deduction) จากตัวอย่างของข้อมูลที่ถูกมอบหมาย โดยจัดเก็บในรูปค่าถ่วงน้ำหนักของโครงข่าย (Weight) ที่เปลี่ยนแปลงไปตามข้อมูลที่ป้อนให้ทำการเรียนรู้ กล่าวคือ มีความสามารถในการรวบรวมความรู้ผ่านกระบวนการเรียนรู้ (Learning process) จึงทำให้โครงข่ายประสาทเทียมถูกนำมาประยุกต์ใช้ในด้านต่าง ๆ อย่างแพร่หลาย อาทิเช่น การตรวจจับผู้บุกรุก การจำแนกชนิดของวัตถุ และสนับสนุนทางการแพทย์ในการวินิจฉัยโรค

การประมวลผลต่าง ๆ ของโครงข่ายประสาทเทียมเกิดขึ้นในหน่วยประมวลผลย่อยที่เรียกว่า โหนด (Node) หรือนิวรอน (Neural) โดยมีฟังก์ชันผลรวมในการคำนวณผลรวมของข้อมูลนำเข้าและค่า

เบี่ยงเบน ฟังก์ชันกำหนดสัญญาณส่งออกที่เรียกว่า ฟังก์ชันกระตุ้น ดังแสดงในรูปที่ 2.25 โดยมีรายละเอียดดังต่อไปนี้



รูปที่ 2.25 โครงสร้างหน่วยการทำงานของโครงข่ายประสาทเทียม

- 1.) ค่าน้ำหนักการเชื่อมต่อ (Connection weights) คือ ค่าสัมประสิทธิ์ที่ปรับขนาด เพิ่ม หรือลด สัญญาณขาเข้าไปหน่วยประมวลผลในเครือข่าย
- 2.) ค่าเบี่ยงเบน (Bias) คือ ค่าสเกลาร์ที่ป้อนให้หน่วยประมวลผลโดยไม่คำนึงถึงความเข้มของสัญญาณขาเข้าซึ่งช่วยให้การเรียนรู้เกิดขึ้นในกรณีที่สัญญาณขาเข้ามีค่าต่ำ โดยค่าเบี่ยงเบนนั้นจะถูกปรับตลอดกระบวนการเรียนรู้ ขึ้นอยู่กับความเข้มของสัญญาณขาเข้า
- 3.) ฟังก์ชันผลรวม (Summation function) เป็นฟังก์ชันความสัมพันธ์ทางคณิตศาสตร์ โดยมีหน้าที่คำนวณผลรวมของข้อมูลขาเข้าและค่าเบี่ยงเบน ซึ่งคำนวณได้จากสมการที่ (4)

$$g = \sum_{i=1}^n x_i \times w_i + \beta \quad (4)$$

เมื่อ x_i คือ ข้อมูลขาเข้า
 w_i คือ ค่าน้ำหนักการเชื่อมต่อ
 β คือค่า bias

- 4.) ฟังก์ชันกระตุ้น (Activation function) เป็นฟังก์ชันที่ควบคุมผลลัพธ์ของฟังก์ชันผลรวมจากการคำนวณผลรวมของสัญญาณขาเข้า ค่าน้ำหนักการเชื่อมต่อ และค่าเบี่ยงเบน ก่อนส่งออกผลลัพธ์เพื่อ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

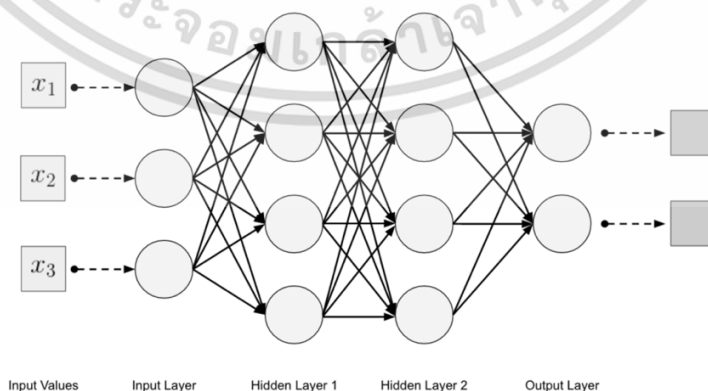
ใช้เป็นข้อมูลขาเข้าของโหนดถัดไป ซึ่งฟังก์ชันกระตุ้นมีอยู่หลายชนิด ทั้งแบบที่เป็นเชิงเส้น และแบบที่ไม่เป็นเชิงเส้น

ซึ่งค่าผลลัพธ์ที่ได้จากหน่วยการทำงานของโครงข่ายประสาทเทียมคำนวณได้จากสมการที่ (5)

$$y = F \left(\sum_{i=1}^n w_i x_i + \beta \right) \quad (5)$$

เมื่อ w_i คือ ค่าน้ำหนักการเชื่อมต่อ
 x_i คือ ข้อมูลขาเข้า
 β คือ ค่า bias
 F คือ ฟังก์ชันกระตุ้น
 y คือ ค่าผลลัพธ์ที่ได้จากจากหน่วยการทำงานของโครงข่ายประสาทเทียม

การทำงานของโครงข่ายประสาทเทียมนั้นขึ้นอยู่กับสถาปัตยกรรมของเครือข่ายที่ถูกกำหนดโดยจำนวนของนิวรอน จำนวนชั้นการทำงาน และประเภทของการเชื่อมต่อระหว่างชั้น ซึ่งสามารถยกตัวอย่างเครือข่ายที่นิยมนำมาใช้พัฒนา ได้แก่ โครงข่ายประสาทเทียมหลายชั้นแบบป้อนไปข้างหน้า (Multilayer feed-forward neural network) เป็นโครงข่ายที่ประกอบไปด้วยชั้นการทำงานทั้งหมด 3 ชั้น คือ ชั้นนำเข้าข้อมูล ชั้นแอบแฝง และชั้นส่งออกข้อมูล โดยมีชั้นนำเข้าข้อมูล และชั้นส่งออกข้อมูลเพียงหนึ่งชั้น แต่สามารถมีหลายชั้นแอบแฝง ซึ่งประกอบไปด้วยจำนวนนิวรอนที่แตกต่างกันได้ โดยมีการส่งสัญญาณเชื่อมต่อกันระหว่างชั้น ดังแสดงในรูปที่ 2.26 โดยมีรายละเอียดขององค์ประกอบต่าง ๆ ดังต่อไปนี้



รูปที่ 2.26 โครงข่ายประสาทเทียมหลายชั้นแบบป้อนไปข้างหน้า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.) ชั้นนำเข้าข้อมูล (Input layer) เป็นชั้นที่มีหน้าที่รับข้อมูลขาเข้าในรูปของเวกเตอร์ ที่ป้อนเข้าสู่เครือข่ายจำนวนนิเวรอนในชั้นนี้ โดยทั่วไปจะเป็นจำนวนเดียวกับคุณสมบัติการป้อนข้อมูลของเครือข่ายเทียม ซึ่งหากเป็นข้อมูลเชิงคุณภาพต้องผ่านกระบวนการแปลงให้อยู่ในรูปเชิงปริมาณที่โครงข่ายยอมรับได้ ชั้นนำเข้าข้อมูลอินพุตจะตามด้วยชั้นแอมแปงอยู่อย่างน้อยหนึ่งเลเยอร์ที่มีการเชื่อมต่อระหว่างกันอย่างสมบูรณ์ในโครงข่ายประสาทเทียมหลายชั้นแบบป้อนไปข้างหน้า

2.) ชั้นแอมแปง (Hidden layer) เป็นชั้นที่มีจำนวนตั้งแต่หนึ่งชั้นขึ้นไปในโครงข่ายประสาทเทียมหลายชั้นแบบป้อนไปข้างหน้า โดยในแต่ละชั้นของ Hidden layer สามารถมีจำนวนนิเวรอน ซึ่งจำนวนชั้นและจะส่งผลต่อการทำงานของโครงข่ายที่แตกต่างกัน ค่าน้ำหนักในการเชื่อมต่อระหว่างชั้น คือ วิธีการเข้ารหัสข้อมูลการเรียนรู้ที่ได้จากข้อมูลตั้งต้น นอกจากนี้ Hidden layer เป็นส่วนสำคัญในการสร้างแบบจำลองฟังก์ชันที่ไม่เป็นเชิงเส้นของโครงข่ายประสาท

3.) ชั้นส่งออกข้อมูล (Output layer) เป็นชั้นที่ทำหน้าที่ส่งออกผลลัพธ์ที่ได้จากการคำนวณของชั้นก่อนการทำงานหน้า เพื่อนำไปใช้เป็นข้อมูลนำเข้าไปแก่หน่วยการทำงานอื่นในโครงข่าย โดยผลลัพธ์ที่ส่งออกไปนั้นจะมีสอดคล้องกับข้อมูลนำเข้าไปที่ได้รับที่ถูกควบคุมโดยฟังก์ชันกระตุ้น ซึ่งอาจเป็นผลลัพธ์ในการวิเคราะห์การถดถอย (Regression) หรือการจัดหมวดหมู่ของข้อมูล (Classification) ทั้งนี้ขึ้นอยู่กับ การตั้งค่าเครือข่าย

4.) การเชื่อมต่อระหว่างชั้น (Fully connected) เป็นการเชื่อมต่ออย่างสมบูรณ์ระหว่างชั้นการทำงานของโครงข่ายประสาทเทียมหลายชั้นแบบป้อนไปข้างหน้า เพื่อส่งค่าของสัญญาณข้อมูล โดยนิเวรอนในชั้นก่อนหน้าถูกเชื่อมต่อไปยังนิเวรอนทั้งหมดในชั้นถัดไปในการปรับปรุงเวกเตอร์พารามิเตอร์ให้เหมาะสม เพื่อลดข้อผิดพลาดให้เหลือน้อยที่สุด

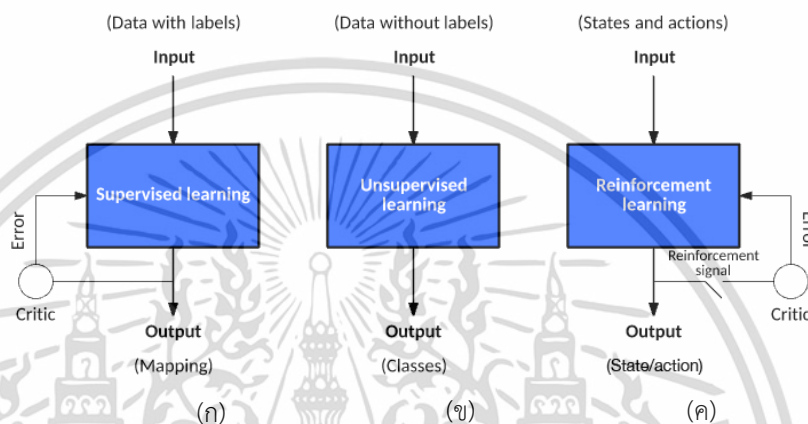
2.3.2 การเรียนรู้ของโครงข่ายประสาทเทียม [15-17]

การเรียนรู้ข้อมูลของโครงข่ายประสาทเทียม เป็นการป้อนข้อมูลเพื่อให้โครงข่ายทำการเรียนรู้ และหาค่าน้ำหนักที่เหมาะสม โดยอาศัยกระบวนการทำซ้ำ (Iterative) ซึ่งการเรียนรู้ข้อมูลของโครงข่ายประสาทเทียม สามารถแบ่งได้เป็น 3 ประเภท ดังแสดงในรูปที่ 2.27 โดยมีรายละเอียดดังต่อไปนี้

1.) การเรียนรู้แบบ Supervised learning เป็นการเรียนรู้ที่มีการกำหนดค่าเป้าหมายของผลลัพธ์ให้กับข้อมูลป้อนเข้า และหาความคลาดเคลื่อนจากผลลัพธ์การเรียนรู้เทียบกับค่าเป้าหมาย โดยค่าความคลาดเคลื่อนที่ได้ถูกนำมาใช้ในการปรับค่าถ่วงน้ำหนักของโครงข่าย เพื่อให้ได้ซึ่งผลลัพธ์ที่ใกล้เคียงกับเป้าหมายมากที่สุด

2.) การเรียนรู้แบบ Unsupervised learning เป็นการเรียนรู้ที่ไม่มีการกำหนดค่าเป้าหมายของผลลัพธ์ให้กับข้อมูลป้อนเข้า โดยใช้การหาความสัมพันธ์ของข้อมูลป้อนเข้าที่มีรูปแบบคล้ายกันเพื่อปรับค่าถ่วงน้ำหนักของโครงข่าย จนได้ซึ่งโครงข่ายที่มีค่าถ่วงน้ำหนักในการจัดหมวดหมู่ของข้อมูลที่เหมาะสม

3.) การเรียนรู้แบบ Reinforcement learning เป็นการเรียนรู้ที่อาศัยทั้งหลักการเรียนรู้แบบ Supervised learning และ Unsupervised learning โดยในการเรียนรู้แบบนี้ไม่จำเป็นต้องอาศัยการกำหนดค่าเป้าหมายของผลลัพธ์ให้กับข้อมูลป้อนเข้า แต่สามารถจับคู่ข้อมูลขาเข้า และผลลัพธ์ได้เอง โดยประเมินจากความเป็นไปได้ของสภาพแวดล้อมในสถานการณ์ที่กำหนด ในระหว่างกระบวนการเรียนรู้ อัลกอริทึมจะสุ่มสำรวจสถานะภายในสภาพแวดล้อมต่าง ๆ และเลือกผลลัพธ์ที่เหมาะสมที่สุด



รูปที่ 2.27 แผนภาพรูปแบบการเรียนรู้ของโครงข่ายประสาทเทียม (ก) Supervised learning (ข) Unsupervised learning (ค) Reinforcement learning

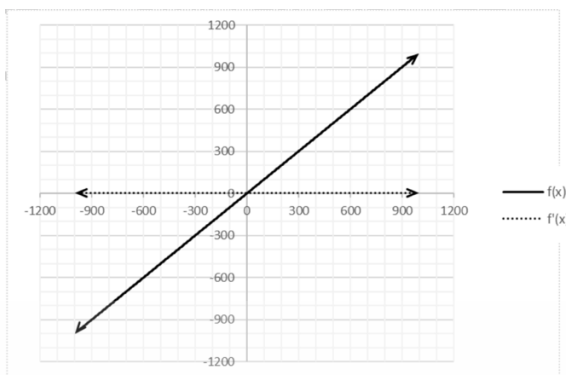
2.3.3 ฟังก์ชันกระตุ้น (Activation Functions)

ฟังก์ชันกระตุ้น เป็นฟังก์ชันแบบสเกลาร์ถึงสเกลาร์ (Scalar-to-scalar function) ที่ใช้ควบคุมผลลัพธ์ก่อนส่งออกค่าผลลัพธ์ไปเป็นข้อมูลขาเข้าของโหนดถัดไป จนกว่าจะถึง Output layer โดยพิจารณาจากค่าผลรวมของข้อมูลเข้า และค่าความเข้มสัญญาณ ว่าควรส่งออกค่าผลลัพธ์เท่าไร ซึ่งฟังก์ชันกระตุ้นมีอยู่หลายชนิด โดยแบ่งออกเป็น 2 กลุ่ม คือ ฟังก์ชันแบบเชิงเส้น และฟังก์ชันแบบไม่เป็นเชิงเส้น

2.3.3.1 ฟังก์ชันกระตุ้นแบบเชิงเส้น (Linear activation function)

ฟังก์ชันกระตุ้นแบบเชิงเส้น โดยพื้นฐานแล้วคือฟังก์ชันเอกลักษณ์ โดยที่ตัวแปรตามมีความสัมพันธ์โดยตรง และเป็นสัดส่วนกับตัวแปรอิสระ ซึ่งหมายถึงฟังก์ชันที่ส่งผ่านสัญญาณโดยไม่เปลี่ยนแปลง โดยแสดงกราฟความสัมพันธ์ของฟังก์ชันกระตุ้นแบบเชิงเส้น ได้ดังรูปที่ 2.28 และคำนวณได้ตามสมการที่ (6)

$$f(x) = x \quad (6)$$



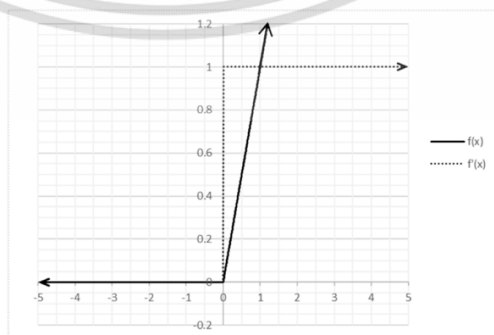
รูปที่ 2.28 กราฟฟังก์ชันกระตุ้นแบบเชิงเส้น

2.3.3.2 ฟังก์ชันกระตุ้นแบบไม่เป็นเชิงเส้น (Non-linear activation function)

ฟังก์ชันกระตุ้นแบบไม่เป็นเชิงเส้น คือ ฟังก์ชันทางคณิตศาสตร์ที่ใช้กับผลลัพธ์ของนิวรอนในโครงข่ายประสาทเทียมเพื่อสร้างความไม่เป็นเชิงเส้นของผลลัพธ์ ซึ่งความไม่เป็นเชิงเส้นหมายถึงความสัมพันธ์ระหว่างข้อมูลขาเข้าและผลลัพธ์ไม่ได้อยู่ในรูปกราฟเส้นตรง และมีรูปร่างกราฟความสัมพันธ์ต่างกันได้ จุดประสงค์ของฟังก์ชันแบบไม่เป็นเชิงเส้นคือเพื่อให้โครงข่ายประสาทเทียมเรียนรู้ และสร้างแบบจำลองความสัมพันธ์ที่ซับซ้อนระหว่างข้อมูลขาเข้าและผลลัพธ์ โดยอย่างฟังก์ชันกระตุ้นแบบไม่เป็นเชิงเส้นได้ดังต่อไปนี้

1.) Rectified linear unit (ReLU)

Rectified linear unit เป็นฟังก์ชันที่เปิดใช้งานเมื่อข้อมูลนำเข้ามีค่ามากกว่า 0 เท่านั้น โดยแทนที่ข้อมูลที่มีค่าลบด้วย 0 เหมาะสำหรับข้อมูลที่มีลักษณะเป็นเชิงปริมาณที่ไม่เป็นค่าลบ โดยมีค่าที่เป็นไปได้ในช่วง 0 ถึงอนันต์ ซึ่งการที่อนุพันธ์ของฟังก์ชันนี้มีค่าเป็น 0 ในแกนลบ ช่วยให้การคำนวณในโครงข่ายเป็นไปได้อย่างรวดเร็ว และมีประสิทธิภาพสูงขึ้น โดยในปี 2009 Zou และคณะได้ทำการพิสูจน์ว่า ReLU มีประสิทธิภาพที่ดีกว่า Hyperbolic tangent ถึง 6 เท่า โดยมีสมการในการคำนวณ และกราฟของฟังก์ชัน ReLU ดังแสดงในสมการที่ (7) และรูปที่ 2.29 ตามลำดับ



รูปที่ 2.29 กราฟฟังก์ชันกระตุ้น Rectified linear unit

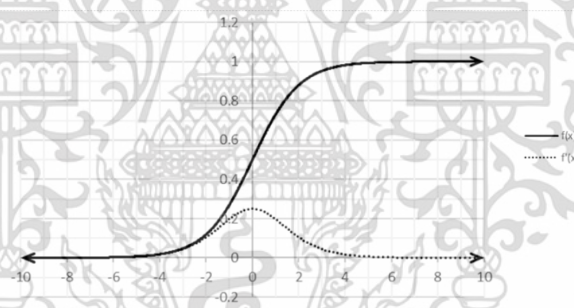
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$ReLU(x) = \max(0, x) \quad (7)$$

ข้อเสียของฟังก์ชัน ReLU คือ ไม่สามารถเรียนรู้จากตัวอย่างที่การเปิดใช้งานเป็นศูนย์ มักเกิดขึ้นหากกำหนดค่าเริ่มต้นของโครงข่ายทั้งหมดด้วยศูนย์ อีกสาเหตุหนึ่งคือเมื่อเกรเดียนต์ขนาดใหญ่ไหลผ่านฟังก์ชัน ReLU จะอัปเดตค่าน้ำหนัก ที่อาจเป็นค่าน้ำหนักเชิงลบ และค่าความเอนเอียงจำนวนมาก ในกรณีนี้ฟังก์ชัน ReLU จะส่งผลลัพธ์เป็นค่า 0 เสมอในระหว่างการแพร่เดินหน้า ส่งผลให้เกรเดียนต์ที่ส่งผ่านฟังก์ชันนี้จะมีผลลัพธ์เป็นศูนย์เสมอโดยไม่คำนึงถึงค่าที่ถูกป้อนเข้า โดยเรียกปัญหานี้ว่า Dead neuron [18]

2.) Sigmoid

Sigmoid เป็นฟังก์ชันคณิตศาสตร์ที่แปลงตัวแปรอิสระที่มีช่วงใดก็ได้เป็นความน่าจะเป็นที่มีค่าระหว่าง 0 ถึง 1 และผลลัพธ์ส่วนใหญ่จะใกล้เคียงกับ 0 หรือ 1 มาก โดยฟังก์ชัน Sigmoid จัดกลุ่มของตัวแปรจากการเลือกจุดแบ่งของผลลัพธ์ โดยมีแสดงกราฟได้ดังรูปที่ 2.30 และสมการคำนวณดังแสดงในสมการที่ (8)



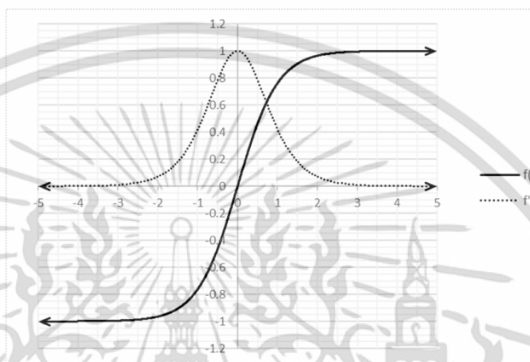
รูปที่ 2.30 กราฟฟังก์ชันกระตุ้น Sigmoid

$$\text{sigmoid}(x) = \frac{1}{1+e^{-x}} \quad (8)$$

ปัญหาของฟังก์ชัน Sigmoid คือ เมื่อพิจารณาจากกราฟของฟังก์ชัน Sigmoid จะเห็นว่าค่าผลลัพธ์ ในแกนตั้ง เปลี่ยนแปลงไปน้อย มากเมื่อเทียบกับค่าของตัวแปรนำเข้าไปในแกนนอน ซึ่งค่าผลลัพธ์ จะเกิดการเปลี่ยนแปลงสูงเฉพาะที่ค่า ของตัวแปรนำเข้าไปในแกนนอน อยู่ในช่วง -2 ถึง 2 เท่านั้น อีกทั้งยังมีค่าอนุพันธ์น้อย และง่ายต่อการลู่เข้าสู่ศูนย์ ส่งผลให้ค่าความคลาดเคลื่อนตกอยู่ที่ค่าต่ำสุดสัมพัทธ์ไม่ใช่ค่าต่ำสุดสัมบูรณ์ ส่งผลให้ประสิทธิภาพของการทำนายของโครงข่ายประสาทเทียมน้อยลง

3.) Hyperbolic tangent

Hyperbolic tangent มีลักษณะคล้ายกับฟังก์ชัน Sigmoid แต่มีค่าที่เป็นไปได้ อยู่ในช่วง -1 ถึง 1 และค่าของอนุพันธ์ของฟังก์ชัน Hyperbolic tangent จะมีความชันมากกว่า ฟังก์ชัน Sigmoid ซึ่งแสดงให้เห็นว่าฟังก์ชันนี้ มีประสิทธิภาพมากกว่าฟังก์ชัน Sigmoid เนื่องจากสามารถสร้างค่าที่คำนวณได้ในช่วงที่กว้างกว่า และเกิดการเปลี่ยนแปลงสูงกว่า โดยมีกราฟและ สมการคำนวณของฟังก์ชัน Hyperbolic tangent ดังแสดงในรูปที่ 2.31 และสมการที่ (9) ตามลำดับ



รูปที่ 2.31 กราฟฟังก์ชันกระตุ้น Hyperbolic tangent

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (9)$$

4.) Softmax

Softmax เป็นฟังก์ชันที่มักถูกวางไว้ที่ชั้นสุดท้ายของโครงข่ายเพื่อใช้หาผลลัพธ์ ซึ่งมีผลลัพธ์เป็นค่าความน่าจะเป็นของแต่ละหมวดหมู่ (Class) โดยผลรวมของความน่าจะเป็นของทุกหมวดหมู่จะมีค่าเท่ากับ 1 สามารถคำนวณได้จากสมการที่ (10)

$$\text{softmax}(z_i) = \frac{\exp(z_i)}{\sum_{k=1}^K \exp(z_k)} \quad (10)$$

เมื่อ z_i คือ คะแนนความน่าจะเป็นของหมวดหมู่ในเวกเตอร์ Z

z_k คือ คะแนนความน่าจะเป็นของหมวดหมู่ทั้งหมด

ในกรณีหากผลลัพธ์ความน่าจะเป็นสูงสุดมีค่าเท่ากันมากกว่าหนึ่งหมวดหมู่ การเลือกกว่าหมวดหมู่ใดเป็นผลลัพธ์ขึ้นอยู่กับวิธีการที่ใช้พิจารณา เช่น การเลือกหมวดหมู่แรกที่พบว่ามีค่าความน่าจะเป็นสูงสุดมาเป็นผลลัพธ์ การสุ่มเลือกหมวดหมู่ผลลัพธ์จากหมวดหมู่ที่มีค่าความน่าจะเป็นสูงสุดเท่ากันทั้งหมด

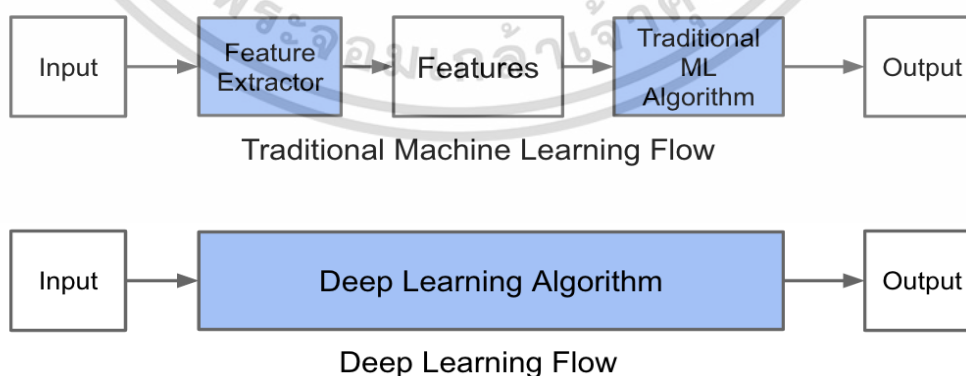
นอกจากนี้ยังมีวิธีการใช้ข้อมูลเพิ่มเติมเพื่อพิจารณา อาทิเช่น การประเมินจากความน่าจะเป็นในการปรากฏของข้อมูล (Prior probability) เป็นต้น

2.4 การเรียนรู้เชิงลึก (Deep learning)

2.4.1 องค์ประกอบของการเรียนรู้เชิงลึก [19-21]

การเรียนรู้เชิงลึก คือ วิธีการเรียนรู้แบบอัตโนมัติด้วยการเลียนแบบการทำงานของโครงข่ายประสาทของมนุษย์ (Neurons) โดยนำระบบโครงข่ายประสาทเทียม มาสร้างสถาปัตยกรรมในระดับสูงที่ประกอบไปด้วยองค์ประกอบย่อยที่ไม่เป็นเชิงเส้นหลายประการ เพื่อทำการเรียนรู้ข้อมูลตัวอย่าง ซึ่งข้อมูลดังกล่าวจะถูกนำไปใช้ในการตรวจจ็รูปแบบ (Pattern) หรือจำแนกหมวดหมู่ของข้อมูล (Classify) โดยทั่วไปวิธีโครงข่ายประสาทเทียมเกิดการเรียนรู้ได้เพียงไม่กี่ชั้น (Layer) และเพื่อที่จะทำให้โครงข่ายประสาทเทียมสามารถประมวลผลข้อมูลที่ซับซ้อนได้มากขึ้นนั้น จึงจำเป็นต้องมีชั้นที่เกิดการประมวลผลและส่งต่อข้อมูลสูงอย่าง Hidden layer มากขึ้น โดยหากยังมีชั้นการทำงานของโครงข่ายมากขึ้น ก็ส่งผลให้มีการประมวลผลในเชิงลึกมากขึ้นไปด้วยจึงเป็นที่มาของคำว่า การเรียนรู้เชิงลึก หรือ Deep Learning

การเรียนรู้เชิงลึก อาจมองได้ว่าเป็นวิธีการหนึ่งของการเรียนรู้ของเครื่อง (Machine learning) ที่พยายามเรียนรู้วิธีการแทนข้อมูล และส่งข้อมูลต่อกันระหว่างโครงข่ายหลาย ๆ ชั้นอย่างมีประสิทธิภาพ กล่าวคือ คุณลักษณะ (Feature) ในชั้นที่อยู่ลึกกว่าจะได้รับมาจากในชั้นที่อยู่ก่อนหน้า ซึ่งสามารถนำมาประยุกต์ใช้เพื่อช่วยลดขั้นตอนในการจำแนกภาพแบบดั้งเดิม (Traditional) ที่จำเป็นต้องสกัดคุณลักษณะ (Feature extraction) เพื่อหาเส้น (edge) สี (color) รูปทรง (shape) ฯลฯ ด้วยตัวเองก่อน (hand-crafted features) เพื่อให้ได้ซึ่งคุณลักษณะที่เหมาะสมกับปัญหา หรือชุดข้อมูลนั้น ๆ โดย Deep learning สามารถอาศัยการหาความสัมพันธ์ และการเรียนรู้ข้อมูลเพื่อดำเนินการตั้งแต่ต้นจนจบกระบวนการได้ ดังแสดงในรูปที่ 2.32



รูปที่ 2.32 ความแตกต่างของการจำแนกภาพระหว่าง Deep learning และ Machine learning แบบดั้งเดิม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สถาปัตยกรรมการเรียนรู้เชิงลึกมีองค์ประกอบที่มีส่วนช่วยในการเรียนรู้ และประมวลผลของโครงข่าย ดังต่อไปนี้

2.4.1.1 ไฮเปอร์พารามิเตอร์ (Hyperparameters)

ไฮเปอร์พารามิเตอร์ คือ พารามิเตอร์ที่ใช้ปรับแต่งเพื่อช่วยให้โครงข่ายทำการเรียนรู้ได้ดีขึ้น หรือรวดเร็วขึ้น รวมถึงควบคุมฟังก์ชันการปรับปรุงประสิทธิภาพให้ทำงานเหมาะสมในระหว่างการฝึกอบรม ซึ่งส่งผลถึงการปรับค่าน้ำหนักของโครงข่ายเพื่อให้เกิดความผิดพลาดน้อยที่สุด ไฮเปอร์พารามิเตอร์มุ่งเน้นไปที่การทำให้มั่นใจว่าโมเดลไม่พอดิ (Underfit) หรือเกินพอดิ (Overfit) ต่อชุดข้อมูลการเรียนรู้ โดยมีตัวอย่างของไฮเปอร์พารามิเตอร์ ดังต่อไปนี้

1.) อัตราการเรียนรู้ (Learning rate)

อัตราการเรียนรู้ คือ ไฮเปอร์พารามิเตอร์ที่ควบคุมขนาดของการปรับค่าน้ำหนักในโครงข่าย โดยในระหว่างกระบวนการเรียนรู้ อัตราการเรียนรู้ เป็นส่วนสำคัญที่ส่งผลต่อประสิทธิภาพของโครงข่าย โดยเป็นตัวกำหนดว่าโครงข่าย หรือแบบจำลองจะไวต่อการเรียนรู้จากข้อมูลป้อนเข้าเพียงใด ซึ่งหากอัตราการเรียนรู้สูงเกินไป อาจส่งผลให้ไม่เหมาะสมกับแบบจำลอง ซึ่งนำไปสู่การฝึกเรียนรู้ที่ไม่เสถียร และมีประสิทธิภาพต่ำ ในทางกลับกัน อัตราการเรียนรู้ที่ต่ำเกินไปอาจส่งผลให้แบบจำลองใช้เวลานานในการหาความสัมพันธ์ของข้อมูลสำหรับการเรียนรู้ การหาค่าที่เหมาะสมของอัตราการเรียนรู้ขึ้นขึ้นอยู่กับปัจจัยหลายประการ อาทิเช่น loss function โครงสร้างของแบบจำลอง และข้อมูลในการเรียนรู้ จึงเป็นไปได้ยากที่จะมีค่าใดค่าหนึ่งที่สามารถใช้ได้เหมาะสมกับทุก ๆ กรณี แต่โดยทั่วไปแล้ว อัตราการเรียนรู้จะตั้งอยู่ในช่วงประมาณ 0.01 – 0.1

2.) อัตราการสุ่มทิ้ง (Dropout)

อัตราการสุ่มทิ้ง เป็นเทคนิคที่ใช้กันทั่วไปในการเรียนรู้เชิงลึกเพื่อลดการเกิดปัญหา Overfitting โดยวิธีการสุ่มนำนิวรอนของโครงข่ายบางส่วนออกระหว่างกระบวนการเรียนรู้ เพื่อป้องกันไม่ให้นิวรอนในโครงข่ายปรับร่วมกับข้อมูลการเรียนรู้มากเกินไป ซึ่งโครงข่ายจะถูกบังคับให้เรียนรู้คุณลักษณะที่มีประสิทธิภาพมากขึ้นแม้ว่านิวรอนบางส่วนจะหายไป ในระหว่างกระบวนการเรียนรู้ Dropout จะถูกนำไปใช้กับแต่ละชั้นการทำงานของโครงข่ายด้วยความน่าจะเป็น p ตัวอย่างเช่น ถ้า $p=0.5$ นิวรอนแต่ละโหนดในชั้นนั้นมีโอกาส 50% ที่จะถูกตัดออก โดยทั่วไป Dropout มักถูกกำหนดให้อยู่ไม่เกิน 0.5 ขึ้นอยู่กับความจำเพาะของข้อมูล และความซับซ้อนของโครงข่าย

2.4.1.2 ฟังก์ชันการสูญเสีย (Loss function)

ฟังก์ชันการสูญเสีย เป็นฟังก์ชันทางคณิตศาสตร์ที่มีค่าอยู่ระหว่าง 0 ถึง 1 ใช้สำหรับประเมินว่าโครงข่ายมีประสิทธิภาพการทำนายได้ดีเพียงใดในชุดข้อมูลที่กำหนด เพื่อหาปริมาณความแตกต่าง

ระหว่างผลลัพธ์ที่คาดการณ์ของแบบจำลองกับค่าความเป็นจริง เพื่อใช้เป็นตัวแปรในการปรับค่าถ่วงน้ำหนักที่เหมาะสมที่สุด โดยหากค่าที่ได้จากฟังก์ชันการสูญเสียเท่ากับ 0 แสดงว่าโครงข่ายมีประสิทธิภาพในการทำงานได้ดีเยี่ยม โดยมีตัวอย่างฟังก์ชันการสูญเสีย ได้แก่ Logistic loss และ Cross-entropy loss เป็นต้น

2.4.1.3 อัลกอริทึมปรับปรุงประสิทธิภาพ (Optimization algorithm)

อัลกอริทึมปรับปรุงประสิทธิภาพ เป็นส่วนสำคัญสำหรับการเรียนรู้ของโครงข่าย ทำหน้าที่ปรับค่าให้แบบจำลองมีการเรียนรู้ที่เหมาะสม หาดำแหน่งที่ทำให้มีข้อผิดพลาดน้อยที่สุด รวมถึงการปรับค่าพารามิเตอร์ของโครงข่าย เพื่อปรับปรุงแบบจำลองให้มีประสิทธิภาพดีขึ้น โดยวิธีการที่นิยมใช้กัน คือ การเคลื่อนลงตามความชัน (Gradient descent algorithm) ซึ่งคำนวณได้จากสมการที่ (11)

$$\nabla g(w) = \left(\frac{\partial g}{\partial w_1}(w), \dots, \frac{\partial g}{\partial w_K}(w) \right) \quad (11)$$

จากสมการที่ (11) เกรเดียนต์ $\nabla g(w)$ คือ เวกเตอร์ของอนุพันธ์ย่อย (Partial derivatives) ทั้งหมดซึ่งบอกถึงความเปลี่ยนแปลงของ $g(w)$ เมื่อ w เปลี่ยน โดยแต่ละองค์ประกอบในเวกเตอร์เกรเดียนต์จะเป็นอนุพันธ์ย่อยของ g ตามตัวแปรแต่ละตัวใน w โดยสามารถคำนวณค่าถ่วงน้ำหนักใหม่ได้จากสมการที่ (12)

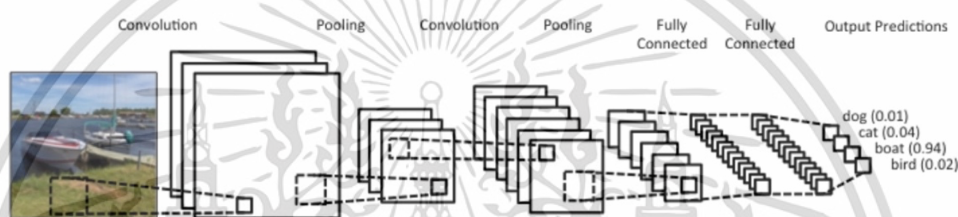
$$\Delta w = w - \alpha * \frac{\partial L}{\partial w} \quad (12)$$

เมื่อ Δw คือ ค่าถ่วงน้ำหนักใหม่
 w คือ ค่าถ่วงน้ำหนักเดิม
 α คือ อัตราการเรียนรู้ (learning rate)
 $\frac{\partial L}{\partial w}$ คือ เกรเดียนต์ของฟังก์ชันของฟังก์ชันสูญเสีย L ต่อพารามิเตอร์ w

เทคนิคการปรับปรุงประสิทธิภาพมีอยู่ด้วยกันหลายเทคนิค ซึ่งหนึ่งในเทคนิคที่นิยมใช้กันอย่างแพร่หลายได้แก่ Adam optimizer (Adaptive moment estimation) เป็นเทคนิคที่ใช้ Gradient โดยผสมระหว่างแนวคิดของ Momentum และการปรับปรุงของอัตราการเรียนรู้ โดยการเพิ่มเศษส่วนของ Gradient ที่ผ่านมาเข้ากับ Gradient ในปัจจุบัน โดยมีจุดเด่นที่ช่วยแก้ไขปัญหาดังกล่าว เช่น การลดอัตราการเรียนรู้ตามเวลาที่ช่วยให้โมเดลไม่หยุดเรียนรู้ได้ และมีการลู่เข้าที่เร็วกว่า Gradient descent

2.4.2 โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional neural network) [19-25]

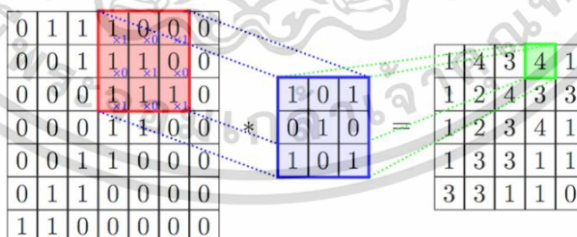
โครงข่ายประสาทเทียมแบบคอนโวลูชัน เป็นโครงข่ายถูกพัฒนาขึ้นมาเพื่อวัตถุประสงค์ใช้ในด้านข้อมูลภาพ ด้วยองค์ประกอบของโครงข่ายที่มีปัจจัยเด่นในการเรียนรู้ และสกัดคุณลักษณะของภาพ โดยใช้หลักการเข้ารหัสข้อมูลความสว่างในแต่ละพิกเซลของภาพให้อยู่ในรูปแบบตัวเลข โครงสร้างของโครงข่ายประสาทเทียมแบบคอนโวลูชัน ประกอบไปด้วย ชั้นคอนโวลูชัน (Convolution layer) และ ชั้นพูลลิ่ง (Pooling layer) ซ้อนกันหลายชั้น ซึ่งเรียกส่วนนี้ว่า Convolution base และมีชั้นสุดท้ายของโครงข่าย คือ ชั้นเชื่อมโยงสมบูรณ์ (Fully connected layer) ทำหน้าที่เชื่อมโยงข้อมูลทั้งหมดเข้าด้วยกัน ก่อนส่งออกเป็นหมวดหมู่ผลลัพธ์ ดังแสดงในรูปที่ 2.33 โดยมีรายละเอียดดังต่อไปนี้



รูปที่ 2.33 ชั้นการทำงานของโครงข่ายประสาทเทียมแบบคอนโวลูชัน

1.) ชั้นคอนโวลูชัน (Convolution layer)

ชั้นคอนโวลูชัน เป็นชั้นที่ทำการรวมภาพข้อมูลเข้ากับตัวกรอง (Filter) ที่มีขนาดเล็กกว่าภาพ ข้อมูลเข้า ในรูปที่ 2.34 เป็นตัวอย่างการใช้ตัวกรองขนาด 3x3 ทาบลงบนภาพ และทำการเข้ารหัสก่อนเลื่อนตัวกรอง (Stride) ไปที่ตำแหน่งถัดไปจนกว่าจุดภาพทั้งหมดจะถูกประมวลผล



รูปที่ 2.34 การทำงานของชั้นคอนโวลูชัน

โดยผลลัพธ์ที่ได้จากการคอนโวลูชันเรียกว่า Feature map ซึ่งชุดข้อมูลใหม่ที่ผ่านการสกัดคุณลักษณะสำคัญจากชุดข้อมูลเดิม ดังแสดงในรูปที่ 2.35



รูปที่ 2.35 ผลลัพธ์การคอนโวลูชัน (ก) ภาพขาเข้า (ข) Feature map ที่ได้จาก Convolution layer

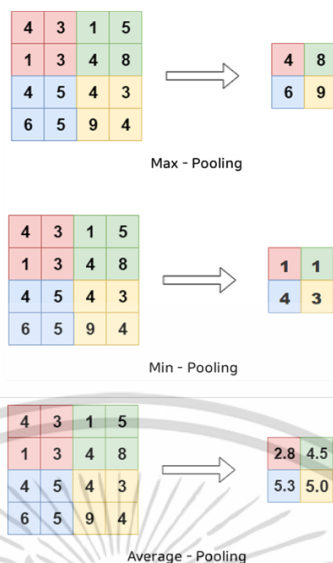
ซึ่งโดยปกติขนาดของข้อมูล Feature map ที่ได้จากการคอนโวลูชันจะมีขนาดเล็กกว่าขนาดของข้อมูลขาเข้า จึงมีเทคนิคที่ใช้เพิ่มข้อมูลที่มีค่าเท่ากับศูนย์เข้าไปยังบริเวณขอบของข้อมูลขาเข้า โดยเรียกเทคนิคนี้ว่า Zero-padding เพื่อรักษาขนาดของผลลัพธ์ให้ใกล้เคียงกับข้อมูลขาเข้า โดยหากต้องการให้ขนาดระหว่างข้อมูลขาเข้า และข้อมูลผลลัพธ์มีขนาดเท่ากัน (Same padding) นั้นสามารถคำนวณหาจำนวน padding เมื่อกำหนดให้จำนวนการเลื่อนตัวกรองมีค่าเท่ากับ 1 ได้ตามสมการที่ (13)

$$P = \frac{(F-1)}{2} \quad (13)$$

เมื่อ P คือ จำนวน padding ที่เป็นไปได้ และ F คือ ขนาดของตัวกรอง

2.) ชั้นพูลลิ่ง (Pooling layer)

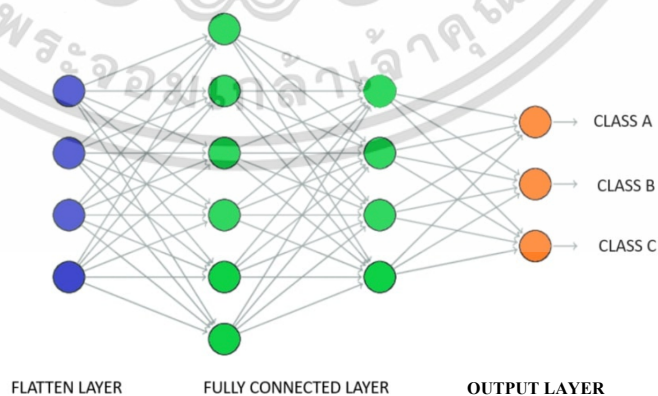
ชั้นพูลลิ่ง เป็นชั้นที่โดยทั่วไปถัดจะแทรกอยู่ระหว่างชั้นคอนโวลูชัน มีไว้เพื่อลดขนาดเชิงพื้นที่ (Down sampling) ของ Feature map ซึ่งช่วยควบคุมปัญหาการจำข้อมูลมากเกินไปของแบบจำลอง (Over fitting) โดยมีแบ่งออกเป็น 3 เทคนิค ได้แก่ เทคนิคการใช้ฟังก์ชันค่าเฉลี่ย (Average pooling) ซึ่งเป็นการเลือกค่าเฉลี่ยของกลุ่มข้อมูลที่กำหนด เทคนิคการใช้ฟังก์ชันค่าต่ำสุด (Min pooling) ที่เป็นการเลือกค่าต่ำสุดจากกลุ่มข้อมูลที่กำหนด และเทคนิคการใช้ฟังก์ชันค่าสูงสุด (Max pooling) ที่เป็นการเลือกค่าสูงสุดจากกลุ่มข้อมูลที่กำหนด หากเลือกใช้ฟังก์ชันค่าสูงสุด ชั้นนี้จะเรียกว่า Max pooling นั้นหมายถึง Feature Map ถูกแบ่งออกเป็น Local region ก่อนนำมาแบ่งออกเป็นขนาด $p \times p$ โดยกำหนดให้ p อยู่ในช่วงระหว่าง 2 ถึง 5 ก่อนเลือกค่าที่มากที่สุดของแต่ละส่วนการแบ่งมาใช้เป็นตัวแทนจนครบทุก Local region ของ Feature Map โดยหากเลือกใช้ขนาดของตัวกรองเป็น 2×2 จะส่งผลให้ขนาดความกว้างและความสูงของข้อมูลลดลง 2 เท่าในมิติเชิงพื้นที่ โดยมีตัวอย่างการพูลลิ่งในแต่ละเทคนิคฟังก์ชันดังแสดงในรูปที่ 2.36



รูปที่ 2.36 ตัวอย่างการพูลลิ่งในแต่ละเทคนิคฟังก์ชัน

3.) ชั้นเชื่อมโยงสมบูรณ์ (Fully connected layer)

ชั้นเชื่อมโยงสมบูรณ์ เป็นชั้นที่อยู่ในส่วนท้ายของโครงข่ายที่ประกอบไปด้วยพารามิเตอร์ปกติของชั้น และไฮเปอร์พารามิเตอร์ โดยมีหน้าที่ในการคำนวณความน่าจะเป็นของหมวดหมู่ที่ใช้เป็นผลลัพธ์ของโครงข่าย ทุกนิรอนในชั้นนี้จะถูกเชื่อมต่อกับจะเชื่อมต่อกับ Feature map ก่อนหน้าแบบ Fully-Connected โดยจะทำการแปลงข้อมูลให้เป็นเวกเตอร์หนึ่งมิติ (Flatten) ที่มีขนาดของปริมาณผลลัพธ์คือ $[1 \times 1 \times N]$ โดยที่ N คือจำนวนหมวดหมู่ที่กำลังประเมิน ดังแสดงในรูปที่ 2.37 ซึ่ง Fully connected layer จะทำการแปลงปริมาณข้อมูลขาเข้าที่เป็นฟังก์ชันของการเปิดใช้งานเป็นน้ำหนัก และค่าความเอนเอียงของนิรอน ก่อนส่งออกข้อมูลไปยังตัวจำแนกหมวดหมู่ (Classifier)



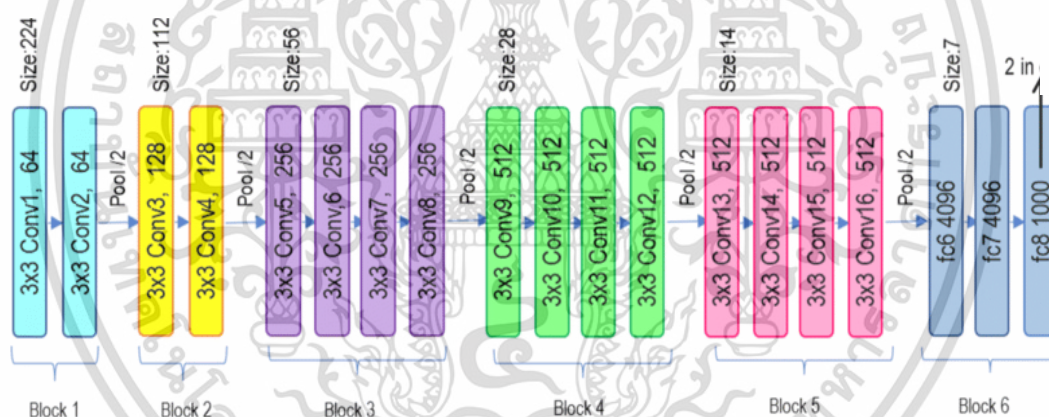
รูปที่ 2.37 ชั้นเชื่อมโยงสมบูรณ์ในโครงข่าย CNN

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4.3 แบบจำลองการถ่ายโอนการเรียนรู้เชิงลึก (Deep transfer learning models)

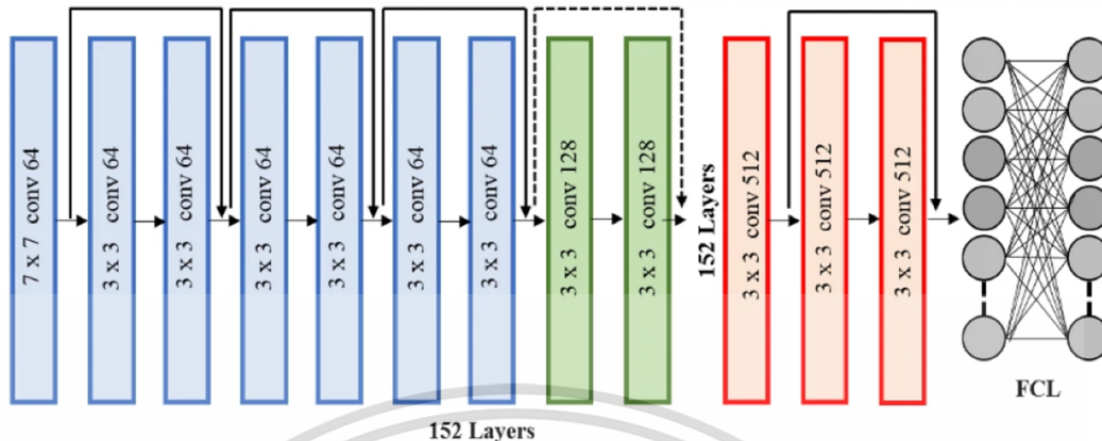
แบบจำลองการถ่ายโอนการเรียนรู้เชิงลึก เป็นแนวทางที่เกี่ยวกับการถ่ายโอนความรู้จากน้ำหนักของแบบจำลองที่ใช้กับงานหนึ่งไปเป็นน้ำหนักของแบบจำลองที่ใช้กับอีกงานหนึ่ง โดยทั่วไปจะใช้เมื่อโครงข่ายประกอบด้วยพารามิเตอร์จำนวนมาก หรือจำนวนตัวอย่างที่รวบรวมไม่เพียงพอต่อการกำหนดคุณลักษณะเฉพาะ โดยการใช้งานการถ่ายโอนการเรียนรู้จะช่วยลดจำนวนข้อมูลที่ต้องป้อนให้แบบจำลองทำการเรียนรู้เพื่อปรับน้ำหนักของโครงข่ายให้เข้ากับปัญหาเฉพาะ โดยมีตัวอย่างดังต่อไปนี้

1.) VGG19 [26] เป็นโครงข่ายประสาทเทียมที่เสนอโดย A. Zisserman และ K. Simonyan ซึ่งเป็นแบบจำลองที่มีจำนวน Hidden layer ทั้งหมด 19 ชั้น แบ่งเป็น convolution layers 16 ชั้น และ fully-connected layers 3 ชั้น แบบจำลองนี้รับข้อมูลนำเข้าที่มีขนาด $224 \times 224 \times 3$ โดยใช้ Convolutional kernel ขนาด 3×3 ที่มี stride เท่ากับ 1 พิกเซล 2×2 max-pooling และ Rectified linear unit (ReLU) ถูกนำมาใช้เพื่อปรับปรุงการเรียนรู้ในการจำแนกของแบบจำลอง โครงสร้างของแบบจำลอง VGG19 แสดงดังรูปที่ 2.38



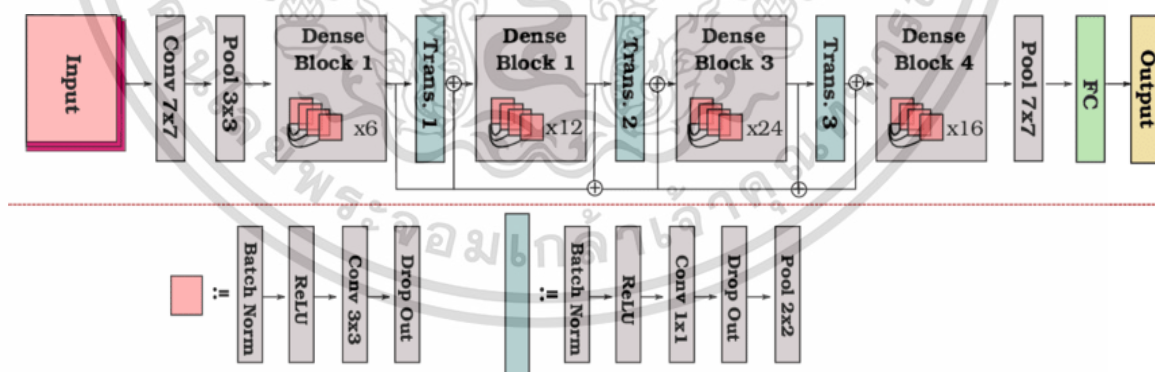
รูปที่ 2.38 โครงสร้างแบบจำลอง VGG19 [27]

2.) ResNet152V2 [28] เป็นเวอร์ชันหนึ่งของเครือข่าย Residual network (ResNet) ที่ชนะการแข่งขัน ILSVRC2015 ซึ่งโครงข่ายนี้ประกอบไปด้วย ชั้นโครงข่ายประสาทเทียมทั้งหมด 152 ชั้น ที่ใช้ skip-connection ซึ่งช่วยให้สามารถ backpropagate และฝึกรวมเครือข่ายที่ลึกขึ้นได้โดยใช้การไล่ระดับ gradient องค์ประกอบหลักของโครงข่ายนี้ได้แก่ Identity blocks และ convolutional blocks นอกจากนี้ ResNetV2 ได้มีการนำเทคนิค batch normalization มาใช้เพื่อปรับค่าข้อมูลของแต่ละ layer ให้มีความเหมาะสมมากยิ่งขึ้น โครงสร้างของแบบจำลอง ResNet152V2 แสดงดังรูปที่ 2.39



รูปที่ 2.39 โครงสร้างแบบจำลอง ResNet152V2 [29]

3.) DenseNet121 เป็นหนึ่งในเครือข่าย Dense convolutional networks [30] เสนอโดย Huang และคณะ เพื่อใช้ในการจำแนกรูปภาพ โดยมีการเชื่อมต่อระหว่าง layer ผ่าน Dense Blocks แบบจำลองนี้ประกอบไปด้วย Hidden layer ทั้งหมด 121 ชั้น แบ่งออกเป็น convolution layer ขนาด 7x7 จำนวน 1 ชั้น convolution layer ขนาด 3x3 จำนวน 58 ชั้น convolution layer ขนาด 1x1 จำนวน 61 ชั้น และ fully-connected layer 1 ชั้น โดยองค์ประกอบภายในโครงข่ายจะแบ่งเป็นส่วน ๆ ที่เรียกว่า DenseBlock โดยมี layer ระหว่าง block เรียกว่า Transition layer โครงสร้างของแบบจำลอง DenseNet121 แสดงดังรูปที่ 2.40



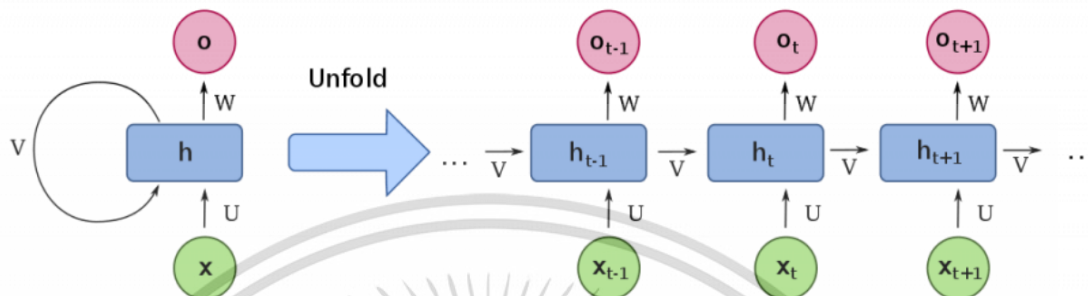
รูปที่ 2.40 โครงสร้างแบบจำลอง DenseNet121 [31]

2.4.4 โครงข่ายประสาทเทียมแบบวนกลับ (Recurrent neural network)

โครงข่ายประสาทเทียมแบบวนกลับ เป็นโครงข่ายที่ถูกพัฒนาขึ้นเพื่อจัดการข้อมูลที่เป็น อนุกรมเวลา (Time series) เนื่องจากโครงข่ายมีการนำผลลัพธ์ที่ได้จากการคำนวณใน state ก่อนหน้า มาร่วมใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เป็นข้อมูลขาเข้าอีกคู่กับข้อมูลขาแบบปกติอีกครั้ง ที่เรียกว่า Hidden state หรือ Internal state ช่วยให้โครงข่ายสามารถรู้จำรูปแบบ (Pattern) ของลำดับข้อมูลขาเข้า (Input Sequence) ได้ โดยมีโครงสร้างของ RNN ดังแสดงรูปที่ 2.41

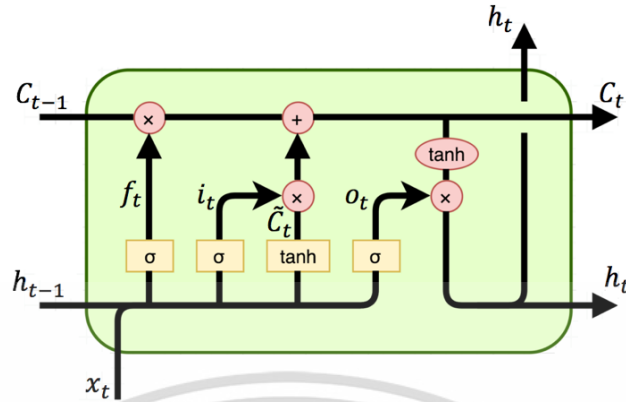


รูปที่ 2.41 โครงสร้างของโครงข่ายประสาทเทียมแบบวนกลับ

1.) Long short-term memory (LSTM) [32] เป็น Recurrent Neural Network ประเภทหนึ่งที่มีความสามารถในการเรียนรู้การพึ่งพาลำดับของข้อมูล ข้อได้เปรียบหลักของ LSTM คือสามารถแยกแยะหน่วยความจำระยะสั้น และระยะยาวเพื่อจัดเก็บไว้ในหน่วยความจำ และปรับปรุงหรือเรียกใช้เมื่อจำเป็น ตลอดจนแก้ปัญหา vanishing gradient ซึ่ง LSTM แต่ละเซลล์ มีองค์ประกอบหลักอยู่ทั้งหมด 4 หน่วยคือ input gate, forget gate, memory cell state gate และ output gate ดังแสดงในรูปที่ 2.42 โดยในแต่ละหน่วยมีหน้าที่ดังต่อไปนี้

- Input gate เป็นหน่วยย่อยในการกำหนดข้อมูลที่น่าเข้ามาวิเคราะห์ในโครงข่าย โดยรับข้อมูลเข้ามาเพื่อเขียนค่าลงไปในแต่ละ Cell
- Forget gate เป็นหน่วยย่อยที่ใช้ในการกำหนดข้อมูลที่น่าเข้ามาวิเคราะห์ในโครงข่าย โดยทำการกำหนดว่าข้อมูลนั้นควรที่จะถูกบันทึกหรือถูกนำออกจากโครงข่าย
- Memory cell state gate เป็นหน่วยย่อยในการกำหนดข้อมูลที่น่าเข้ามาวิเคราะห์ใน Cell และทำการคำนวณค่าสถานะ เพื่อใช้ในการคำนวณครั้งถัดไป
- Output gate เป็นหน่วยย่อยสำหรับการคำนวณ output ของ Cell ซึ่งผลลัพธ์ที่ได้จาก Cell นี้จะมีอยู่ 2 อย่าง ได้แก่ output และ hidden state สำหรับใช้ในการคำนวณครั้งถัดไป

นิพจน์การเปลี่ยนแปลงของ LSTM ถูกกำหนดไว้ในสมการที่ (14) – (19) ตามลำดับ



รูปที่ 2.42 โครงสร้างของโครงข่าย Long short-term memory [33]

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + V_i c_{t-1}) \quad (14)$$

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + V_f c_{t-1}) \quad (15)$$

$$O_t = \sigma(W_o x_t + U_o h_{t-1} + V_o c_{t-1}) \quad (16)$$

$$\tilde{C}_t = \tanh(W_c x_t + U_c h_{t-1}) \quad (17)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{C}_t \quad (18)$$

$$h_t = O_t \odot \tanh(c_t) \quad (19)$$

เมื่อ i_t คือ input gate f_t คือ forget gate O_t คือ output gate c_t หมายถึง memory cell state gate h_t หมายถึง hidden state ณ เวลา t x_t หมายถึง ข้อมูลขาเข้า ณ ปัจจุบัน σ และ \odot หมายถึง ฟังก์ชัน sigmoid และการคูณแบบตำแหน่งต่อตำแหน่ง ตามลำดับ

2.) Gated recurrent unit (GRU) [34] เป็นโครงข่ายที่ถูกปรับปรุงมาจาก Recurrent neural network ปกติ ถูกนำเสนอในปี 2014 โดย Cho และคณะ GRU มีลักษณะการทำงานคล้ายกับ LSTM ตรงที่มี forget gate แต่ใช้พารามิเตอร์น้อยกว่า เนื่องจากไม่มี output gate และ discrete cell แต่ประกอบไปด้วย update gate และ reset gate ดังต่อไปนี้โดยมีโครงสร้างดังแสดงในรูปที่ 2.43 ทำให้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

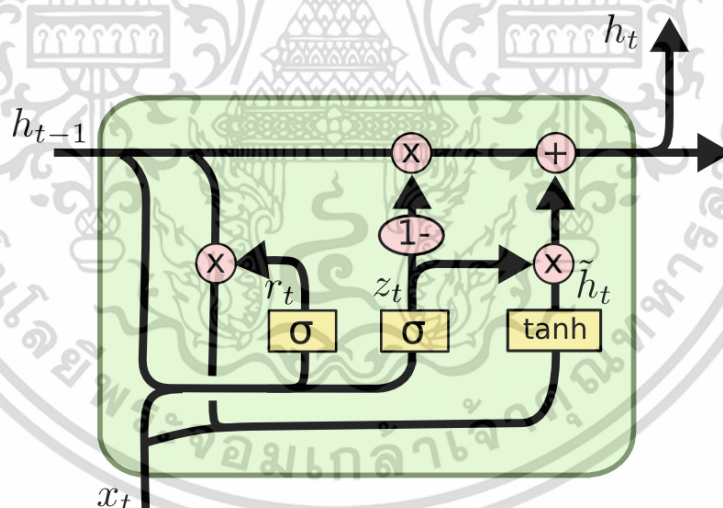
GRU ใช้เวลาในการเรียนรู้ และทำนายผลได้รวดเร็วกว่า LSTM เนื่องจากมีความเรียบง่ายของสถาปัตยกรรมมากกว่า ซึ่งในแต่ละหน่วยมีหน้าที่ดังต่อไปนี้

- Update gate เป็นหน่วยย่อยที่ทำการนำข้อมูลเข้า (input) และสถานะซ่อน (hidden state) ปัจจุบันไปคำนวณเพื่อกำหนดสถานะของ Cell สำหรับใช้ในการคำนวณในขั้นถัดไป การคำนวณในทุกๆ รอบที่มีข้อมูลเข้ามา Update Gate จะช่วยควบคุมปริมาณข้อมูลใหม่ที่ต้องการเก็บไว้ในสถานะซ่อน และข้อมูลเก่าที่จะถูกลบออก

- Reset gate เป็นหน่วยย่อยที่ใช้ในการกำหนดข้อมูลว่าควรที่จะเก็บค่าสถานะที่ได้จากการคำนวณในครั้งที่แล้วมากน้อยเพียงใด Reset Gate จะทำหน้าที่ในการล้างสถานะซ่อนบางส่วน หรือทั้งหมด

สำหรับการคำนวณค่า output และ hidden state ของ GRU นั้น หน่วยย่อยที่ใช้ในการคำนวณเป็นแบบคงที่ โดยจะคำนวณค่าผลลัพธ์จากทั้งสอง gate ก่อนนำมาผ่านฟังก์ชัน Tanh จากนั้นค่าที่ได้จาก reset gate จะกำหนดว่าจะเก็บหรือทำการล้างค่าเดิมออกไป และควบคุมปริมาณข้อมูลด้วยค่าจาก update gate

นิพจน์การเปลี่ยนแปลงของ GRU ดังสมการที่ (20) – (23) ตามลำดับ



รูปที่ 2.43 โครงสร้างของโครงข่าย Gated recurrent unit [35]

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \quad (20)$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (21)$$

$$\hat{h}_t = \tanh(W_h x_t + U_h (r_t \odot h_{t-1}) + b_h) \quad (22)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \hat{h}_t \quad (23)$$

เมื่อ z_t คือ update gate r_t คือ reset gate \hat{h}_t คือ candidate activation vector h_t หมายถึง เวกเตอร์ผลลัพธ์ x_t หมายถึง เวกเตอร์ข้อมูลขาเข้า σ หมายถึง ฟังก์ชัน sigmoid \odot หมายถึง การคูณแบบตำแหน่งต่อตำแหน่ง W, U และ b คือ พารามิเตอร์เมทริกซ์ และเวกเตอร์

2.5 ทบทวนวรรณกรรม

ในปัจจุบันแบบจำลองการเรียนรู้เชิงลึก (Deep learning) ถูกนำมาประยุกต์ใช้กับทางการแพทย์ เพื่อวินิจฉัยโรคเนื่องจากมีประสิทธิภาพสูงและช่วยอำนวยความสะดวกต่อบุคคลากรทางการแพทย์ เนื่องจากแบบจำลองการเรียนรู้เชิงลึกนั้นสามารถกำหนดข้อมูลที่ต้องการให้เรียนรู้และประมวลผลลัพธ์ตามข้อมูลตัวอย่างที่ถูกรู้ได้เป็นอย่างดี ซึ่งการเลือกนำภาพเอกซเรย์ และภาพ CT มาใช้ร่วมกับโครงข่ายแบบจำลองการเรียนรู้เชิงลึกนั้นเป็นอีกหนึ่งทางเลือกสำหรับการตรวจคัดกรองที่สามารถทราบผลได้ในทันที ในการศึกษาก่อนหน้านี้ Zhang และคณะ [36] ใช้ Deep learning ที่มีจำนวนของชุดข้อมูลจำกัดในการเรียนรู้เพื่อระบุภาวะติดเชื้อไวรัสโคโรนาจากภาพเอกซเรย์ทรวงอก อันประกอบไปด้วยภาพเอกซเรย์ของผู้ป่วยโรค COVID-19 135 ราย และผู้ป่วยโรคปอดอักเสบ 320 ราย ผลลัพธ์ที่ได้จากการทดลองนี้ให้ค่าความแม่นยำ (Accuracy) อยู่ที่ 91.24% เมื่อใช้แบบจำลอง Pre-trained VGG-19 และ ResNet50 เมื่อปี ค.ศ. 2020 Hemdan และคณะ [37] ทดลองใช้แบบจำลอง Pre-trained CNN ทั้งหมด 7 แบบเพื่อวินิจฉัยโรค COVID-19 จากภาพเอกซเรย์ทรวงอกโดยใช้ชุดข้อมูลที่มีภาพเอกซเรย์ผู้ติดเชื้อ COVID-19 และภาวะปกติอย่างละ 25 ภาพ ผลลัพธ์ที่ได้พบว่าแบบจำลอง VGG-19 และ DenseNet-201 ให้ประสิทธิภาพสูงที่สุด โดยมีค่าความแม่นยำอยู่ที่ 90% และ F1-score เท่ากับ 91% ต่อมา Islam และคณะ [38] ได้พัฒนาโครงข่าย CNN ที่ทำงานร่วมกับ LSTM เพื่อตรวจหาผู้ติดเชื้อ COVID-19 จากภาพเอกซเรย์ทั้งหมด 4,575 ภาพ ซึ่ง 1,525 ภาพ เป็นของผู้ป่วยที่ติดเชื้อ COVID-19 ผลลัพธ์ที่ได้จากโครงข่ายที่นำเสนอมีมีประสิทธิภาพที่สูง โดยมีค่าความแม่นยำสูงถึง 99.4% และ Recall 99.3% ถัดมาในปี 2021 Rahman และคณะ [39] ระบุการติดเชื้อ COVID-19 จากภาพเอกซเรย์ โดยเปรียบเทียบประสิทธิภาพจากการใช้เทคนิคการปรับปรุงคุณภาพสีที่กระทำต่อข้อมูลภาพก่อนนำไปประมวลผลโดยแบบจำลอง CNN ซึ่งผลลัพธ์ที่ได้แสดงให้เห็นว่าการใช้แบบจำลอง ChexNet ร่วมกับการใช้เทคนิค Gamma correction ในการปรับปรุงข้อมูลภาพให้ค่าประสิทธิภาพความแม่นยำสูงสุด คือ 96.2%

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กลุ่มงานวิจัยที่สำคัญอีกหนึ่งกลุ่มคือกลุ่มที่ใช้ภาพ CT ในการระบุผู้ติดเชื้อ COVID-19 เมื่อปี ค.ศ. 2020 Wu และคณะ [40] ใช้สถาปัตยกรรมของโครงข่าย ResNet50 ในการวินิจฉัยผู้ป่วยโรค COVID-19 โดยใช้แนวคิด Multi-View Fusion ซึ่งในงานวิจัยดังกล่าวใช้ภาพ CT จำนวน 495 ภาพ โดยวิธีการที่พวกเขาแนะนำเสนอนั้นให้ค่าประสิทธิภาพความแม่นยำอยู่ที่ 76% ค่าความไว (Sensitivity) 81.1% และค่าความจำเพาะ (Specificity) 61.5% นอกจากนี้ Xu และคณะ [41] ใช้แบบจำลอง Pre-trained ResNet18 ในการตรวจหาโรค COVID-19 จากภาพ CT ทรวงอกทั้งหมด 618 ภาพ ซึ่งให้ค่าประสิทธิภาพความแม่นยำ เท่ากับ 86.7% และ F1-score 83.9%

อีกหนึ่งงานวิจัยก่อนหน้านี้ Perumal และคณะ [42] เสนอการตรวจหาโรค COVID-19 จากทั้งภาพเอกซเรย์ และภาพ CT โดยใช้วิธีการถ่ายโอนการเรียนรู้ของแบบจำลอง และคุณสมบัติ Haralick ซึ่งให้ผลลัพธ์ความแม่นยำ 93% ค่า Precision 91% และค่า Recall 90% โดยการใช้การถ่ายโอนการเรียนรู้ของแบบจำลอง VGG-16

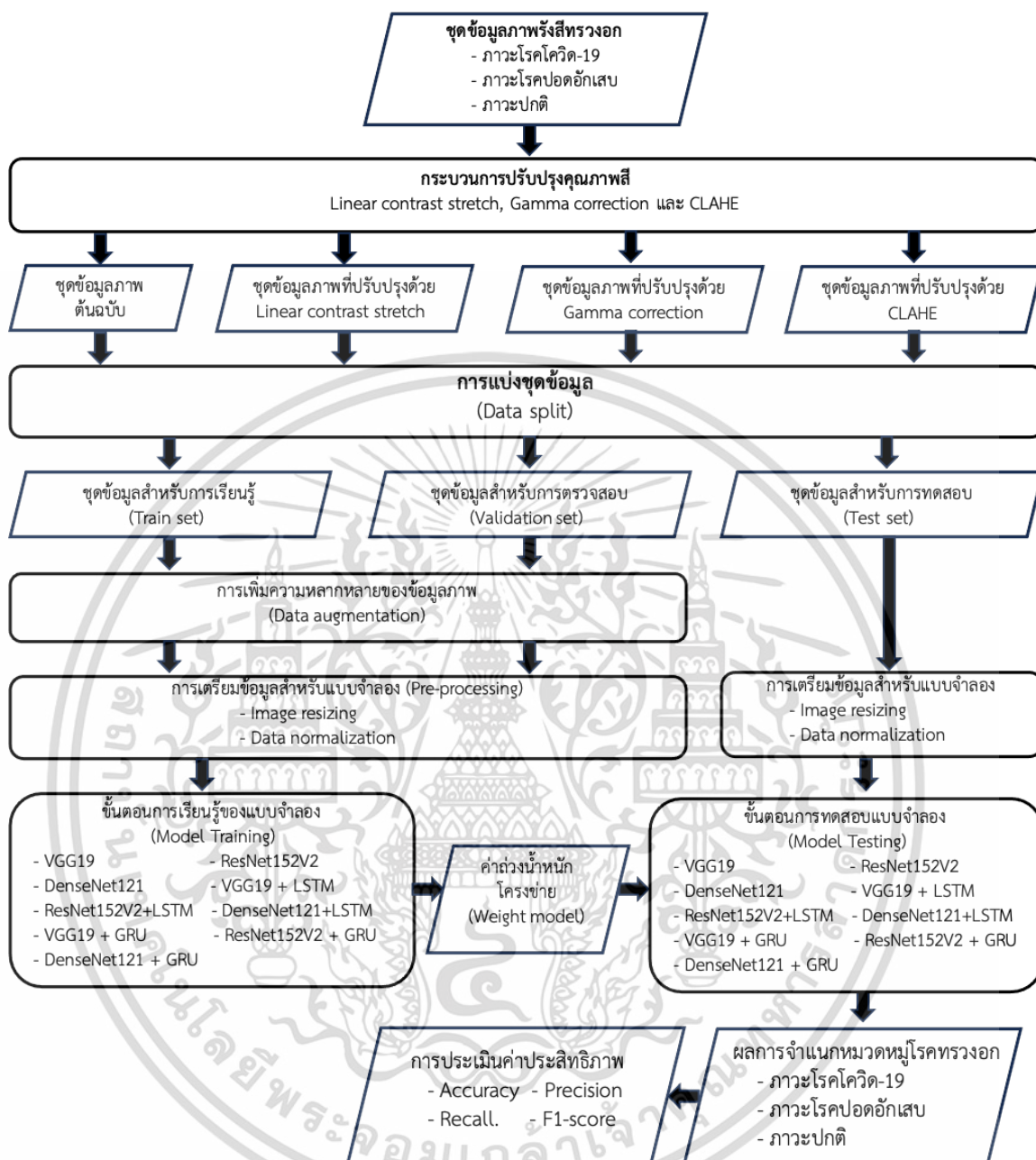
ในงานวิจัยนี้โครงข่าย CNN ที่ทำงานร่วมกับ RNN ถูกพัฒนาเพื่อใช้จำแนกโรคปอดสามภาวะ ได้แก่ โรค COVID-19 โรคปอดอักเสบ และภาวะปกติ จากทั้งภาพเอกซเรย์ และภาพ CT ทรวงอกที่ได้รับการปรับปรุงคุณภาพสัด้วยเทคนิคต่าง ๆ ได้แก่ Normalization, Gamma Correction และ CLAHE โดยการนำแบบจำลอง Pre-trained VGG19, ResNet152V2 และ DenseNet121 ซึ่งเป็นแบบจำลอง CNN มารวมกับแบบจำลอง RNN ได้แก่ LSTM และ GRU ซึ่งค่าประสิทธิภาพของวิธีการที่นำเสนอถูกประเมินโดยใช้ตัวชี้วัดทั้งหมดสี่ค่า คือ ค่าความแม่นยำ (Accuracy, ACC) Precision Recall และ F1-score

บทที่ 3

วิธีดำเนินการทดลอง

งานวิจัยนี้เป็นการหาโครงข่ายแบบจำลอง CNN-RNN ที่มีประสิทธิภาพสูงที่สุดในการคัดกรองโรคทรวงอกของผู้ป่วยจากภาพเอกซเรย์ และภาพ CT ทรวงอก โดยชุดข้อมูลภาพถ่ายรังสีทรวงอกที่ใช้ในการทดลองแบ่งออกเป็น 3 หมวดหมู่ได้แก่ ภาพทรวงอกของผู้ป่วยโรคโควิด-19 ภาพทรวงอกในภาวะปกติ และภาพทรวงอกของผู้ป่วยโรคปอดอักเสบ ที่มีขนาด และคุณภาพของภาพที่แตกต่างกัน จึงได้มีการนำเทคนิคการปรับปรุงคุณภาพสีของภาพ (Image enhancement) ได้แก่ Linear contrast stretch Gamma correction และ Contrast limited adaptive histogram equalization (CLAHE) มาใช้ในการปรับปรุงคุณภาพสีของชุดข้อมูลเพื่อร่วมใช้ในการวิเคราะห์ผลการจำแนกของโครงข่ายที่ได้รับการพัฒนาโดยการรวมแบบจำลองประเภท CNN เข้ากับแบบจำลองประเภท RNN ในการศึกษาครั้งนี้ ได้นำแบบจำลอง Pre-trained CNN ที่ผ่านกระบวนการเรียนรู้ด้วยชุดข้อมูล ImageNet แล้ว ซึ่งเป็นที่นิยมใช้กันอย่างแพร่หลายในด้านการจำแนกหมวดหมู่ภาพ ได้แก่ ResNet152V2 VGG19 และ DenseNet121 มาใช้งานร่วมกับแบบจำลองประเภท RNN คือ LSTM และ GRU นอกจากนี้ทางผู้จัดทำยังได้นำแบบจำลองเฉพาะ Pre-trained CNN ที่กล่าวไว้ข้างต้นโดยไม่ถูกรวมเข้ากับแบบจำลองประเภท RNN มาใช้ในการทดลอง เพื่อเปรียบเทียบค่าประสิทธิภาพอีกด้วย ซึ่งค่าประสิทธิภาพของโครงข่ายที่พัฒนาถูกประเมินโดยใช้ตัวชี้วัดทั้งหมดสี่ค่า คือ ค่าความแม่นยำ (Accuracy, ACC) Precision Recall และ F1-score ในงานวิจัยนี้สามารถแสดงแผนภาพลำดับขั้นตอนการทำงานของวิธีการที่นำเสนอ ดังแสดงในรูปที่ 3.1 และแบ่งลำดับวิธีการวิจัยออกได้เป็น 6 ขั้นตอนหลักดังต่อไปนี้

- 1.) การรวบรวมชุดข้อมูลภาพรังสีทรวงอกเพื่อใช้ในการวิจัย
- 2.) การปรับปรุงคุณภาพสีของชุดข้อมูลภาพรังสีทรวงอก
- 3.) การพัฒนาโครงข่าย CNN และ CNN-RNN สำหรับจำแนกโรคทรวงอกจากภาพรังสี
- 4.) การจัดเตรียมข้อมูลเพื่อใช้ในกระบวนการเรียนรู้ และทดสอบโครงข่าย
- 5.) กระบวนการเรียนรู้ข้อมูลในการจำแนกโรคทรวงอกจากภาพถ่ายรังสีของแบบจำลอง
- 6.) การจำแนกหมวดหมู่โรคทรวงอก และการประเมินค่าประสิทธิภาพของแบบจำลอง



รูปที่ 3.1 แผนภาพลำดับขั้นตอนการทำงานของวิธีการที่นำเสนอ

3.1 การรวบรวมชุดข้อมูลภาพรังสีทรวงอก

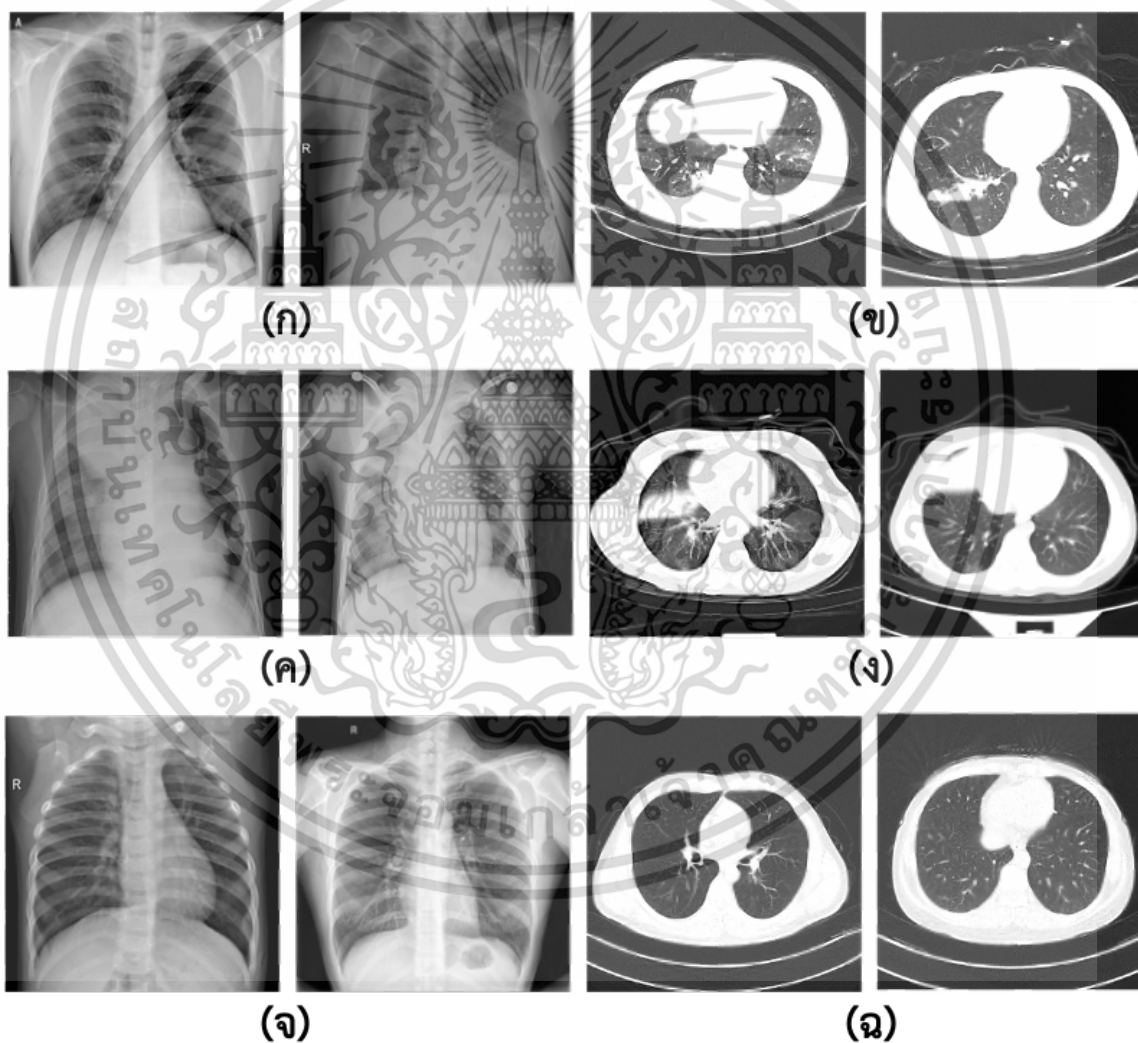
ในงานวิจัยนี้ ผู้จัดทำได้รวบรวมภาพเอกซเรย์ และภาพ CT ทรวงอก จากสี่ฐานข้อมูลที่เผยแพร่อยู่บนสาธารณะเพื่อนำมาใช้สำหรับการทดลอง โดยข้อมูลภาพทั้งหมดเป็นภาพที่อยู่ในปริภูมิสี RGB ที่มีจำนวน 8 บิต อันประกอบไปด้วยภาพทรวงอกของผู้ป่วยโรคโควิด-19 โรคปอดอักเสบ และภาพทรวงอก

ในภาวะปกติ ลำดับแรกเป็นฐานข้อมูลที่ถูกเผยแพร่โดย Cohen และคณะ [43] เป็นฐานข้อมูลที่ประกอบไปด้วยภาพเอกซเรย์ และภาพ CT ที่มีขนาด $224 \times 224 \times 3$ รวม 930 ภาพ ซึ่งถูกนำมาใช้ในการทดลองนี้ 422 ภาพ แบ่งออกเป็นภาพเอกซเรย์จำนวน 342 ภาพ และภาพ CT 80 ภาพ โดยเลือกใช้เฉพาะภาพที่เป็นมุมมองถ่ายภาพจากด้านหน้าของร่างกายไปยังด้านหลังของร่างกายเท่านั้น ถัดมาเป็นฐานข้อมูลที่ถูกรวบรวม และเผยแพร่โดย Chowdhury และคณะ [44] เพื่อใช้สำหรับวิจัยการตรวจคัดกรองโรคโควิด-19 จากภาพถ่ายเอกซเรย์ ขนาด $256 \times 256 \times 3$ ซึ่งชุดข้อมูลนี้ประกอบไปด้วยภาพเอกซเรย์ของทรวงอกของผู้ป่วยโรคโควิด-19 โรคปอดติดเชื้อ โรคปอดผิดปกติจากสภาวะน้ำในปอด หรือเนื้องอกในปอด และภาพทรวงอกของผู้ตรวจคัดกรองที่เป็นปกติ โดยในงานวิจัยนี้ผู้เลือกใช้เฉพาะภาพเอกซเรย์ที่เป็นโรคโควิด-19 และภาพทรวงอกปกติ จำนวน 5,140 ภาพ โดยแบ่งเป็น ภาพเอกซเรย์ที่เป็นโรคโควิด-19 2,595 ภาพ จากทั้งหมด 3,616 และ ภาพทรวงอกปกติ 2,545 จากทั้งหมด 10,192 ภาพ เพื่อให้เกิดความสมดุลของจำนวนในแต่ละหมวดหมู่ภาพ ฐานข้อมูลถัดมาเป็นฐานข้อมูลภาพ CT ทรวงอก ที่เผยแพร่โดย Kang และคณะ [45] เพื่อใช้ในการศึกษาการตรวจคัดกรองโรคทรวงอกจากภาพ CT โดยในชุดข้อมูลนี้ประกอบไปด้วยภาพ CT ทรวงอก สามหมวดหมู่ ได้แก่ ภาพทรวงอกในภาวะโรคโควิด-19 ภาพทรวงอกในภาวะโรคปอดอักเสบ และภาพทรวงอกที่เป็นปกติ จำนวนรวม 104,009 ภาพ ที่มีขนาดต่างกัน ดังนี้ $256 \times 256 \times 3$, $512 \times 512 \times 3$ และ $1,024 \times 1,024 \times 3$ โดยในงานวิจัยนี้ผู้เลือกใช้ 6,859 ภาพ จากทั้งหมดเพื่อรักษาสมดุลของจำนวนภาพเอกซเรย์ และ CT แบ่งเป็นทรวงอกในภาวะโรคโควิด-19 2,235 ภาพ ทรวงอกในภาวะโรคปอดอักเสบ 2,365 ภาพ และทรวงอกที่เป็นปกติ 2,259 ภาพ ฐานข้อมูลสุดท้ายที่ใช้ในงานวิจัยนี้เป็นฐานข้อมูลภาพเอกซเรย์ทรวงอกในภาวะโรคปอดอักเสบ ที่เผยแพร่โดย Kermany และคณะ [46] ซึ่งผู้จัดทำผู้เลือกใช้ 3,789 ภาพ จากภาพในชุดข้อมูลทั้งหมด 4,273 ภาพ ที่มีขนาดรูปภาพระหว่าง $400 \times 138 \times 3$ และ $2,772 \times 2,098 \times 3$ เพื่อรักษาสมดุลของจำนวนข้อมูลในแต่ละหมวดหมู่ภาพทรวงอก

ในงานวิจัยนี้ใช้ภาพรังสีทรวงอกรวมทั้งหมด 16,210 ภาพ แบ่งหมวดหมู่ตามกลุ่มโรคทรวงอกได้เป็น ภาพทรวงอกในภาวะโรคโควิด-19 จำนวน 5,252 ภาพ แบ่งเป็นภาพเอกซเรย์ 2,937 ภาพ และภาพ CT 2,315 ภาพ ภาพทรวงอกในภาวะโรคปอดอักเสบ 6,154 ภาพ แบ่งเป็นภาพเอกซเรย์ 3,789 ภาพ และภาพ CT 2,365 ภาพ และทรวงอกที่เป็นปกติ 4,803 ภาพ แบ่งเป็นภาพเอกซเรย์ 2,545 ภาพ และภาพ CT 2,259 ภาพ ซึ่งสามารถแสดงจำนวนชุดข้อมูลที่ใช้ในงานวิจัยนี้ได้ดังตารางที่ 3.1 และตัวอย่างภาพที่รวบรวมได้ดังรูปที่ 3.2

ตารางที่ 3.1 จำนวนข้อมูลที่ใช้ในงานวิจัย

ฐานข้อมูล	จำนวนภาพเอกซเรย์ทรวงอก			จำนวนภาพ CT ทรวงอก			จำนวนทั้งหมด
	โควิด-19	ปอดอักเสบ	ปกติ	โควิด-19	ปอดอักเสบ	ปกติ	
[3.1]	342	-	-	80	-	-	422
[3.2]	2,595	-	2,545	-	-	-	5,140
[3.3]	-	-	-	2,235	2,365	2,259	6,859
[3.4]	-	3,789	-	-	-	-	3,789
จำนวนทั้งหมด	2,937	3,789	2,545	2,315	2,365	2,259	16,210



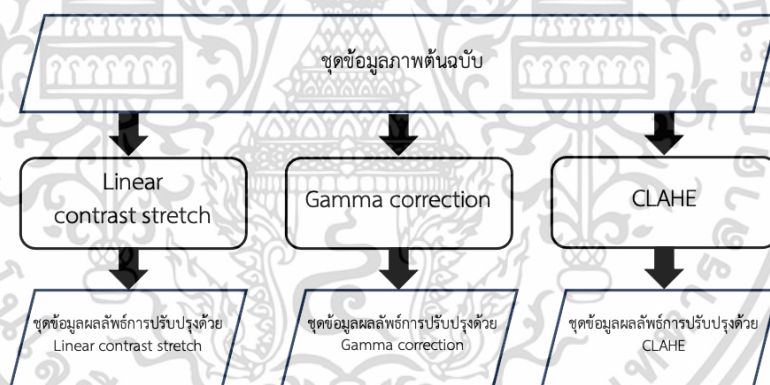
รูปที่ 3.2 ตัวอย่างข้อมูลภาพรังสีทรวงอกที่ใช้ในการทดลอง (ก) ภาพเอกซเรย์ทรวงอกภาวะโรคโควิด-19 (ข) ภาพ CT ทรวงอกภาวะโรคโควิด-19 (ค) ภาพเอกซเรย์ทรวงอกภาวะโรคปอดอักเสบ (ง) ภาพ CT ทรวงอกภาวะโรคปอดอักเสบ (จ) ภาพเอกซเรย์ทรวงอกภาวะปกติ (ฉ) ภาพ CT ทรวงอกภาวะปกติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากตัวอย่างข้อมูลภาพรังสีทรวงอกที่แสดงในรูปที่ 3.2 จะเห็นได้ว่าภาพถ่ายรังสีทรวงอกภาวะโรคโควิด-19 มีบริเวณที่เป็นฝ้าขาวกระจุกตัวอยู่ในภาพ (Ground-glass pattern) แตกต่างจากภาพทรวงอกในภาวะปกติ เนื่องจากความหนาแน่นของปอดที่เพิ่มขึ้น [47] โดยที่ภาพทรวงอกภาวะโรคปอดอักเสบ จะมีลักษณะของฝ้าขาวที่หนากว่า และมีการกระจายตัวมากกว่า [48]

3.2 การปรับปรุงคุณภาพสีของชุดข้อมูลภาพรังสีทรวงอก

ในงานวิจัยนี้ ผู้จัดทำได้นำเทคนิคการปรับปรุงคุณภาพสี 3 เทคนิค ได้แก่ Linear contrast stretch Gamma correction และ CLAHE มาประยุกต์ใช้ในการปรับปรุงชุดข้อมูลก่อนนำไปป้อนเป็นข้อมูลขาเข้าของแบบจำลองในการประมวลผล โดยมีสมมุติฐานว่าการใช้เทคนิคการปรับปรุงคุณภาพสีนั้น ช่วยเพิ่มความเด่นชัดของรอยโรคภายในภาพ และลดสัญญาณรบกวนของชุดข้อมูลภาพ ซึ่งอาจส่งผลให้แบบจำลองประมวลผลข้อมูลภาพได้ดียิ่งขึ้น เมื่อชุดข้อมูลภาพต้นฉบับผ่านขั้นตอนการปรับปรุงคุณภาพสีแล้วจำนวนชุดข้อมูลจะเพิ่มขึ้น จำนวน 3 ชุด ตามเทคนิคที่ใช้ โดยในแต่ละชุดมีจำนวนภาพเท่ากับชุดข้อมูลภาพต้นฉบับ ขั้นตอนการปรับปรุงคุณภาพสีของชุดข้อมูลภาพแสดงดังรูปที่ 3.3



รูปที่ 3.3 ขั้นตอนการปรับปรุงคุณภาพสีของภาพรังสีทรวงอก

3.3 การพัฒนาโครงข่าย CNN และ CNN-RNN สำหรับจำแนกรโรคทรวงอกจากภาพรังสี

ในงานวิจัยนี้ผู้จัดทำได้พัฒนาโครงข่ายแบบจำลอง CNN-RNN เพื่อใช้ในการจำแนกหมวดหมู่โรคทรวงอกจากภาพถ่ายรังสี ทั้งภาพเอกซเรย์ และภาพ CT โดยได้หยิบยกแบบจำลอง Pre-trained CNN ที่ผ่านกระบวนการเรียนรู้ด้วยชุดข้อมูล ImageNet ที่มีประสิทธิภาพในการจำแนกหมวดหมู่ภาพ ได้แก่ ResNet152V2 VGG19 และ DenseNet12 มาประยุกต์ใช้ในการสกัดคุณลักษณะของรูปภาพ และรวมเข้ากับแบบจำลองประเภท RNN ในการหาความสัมพันธ์ของคุณลักษณะประเภทของข้อมูลภาพ โดยแบบจำลองประเภท RNN ที่ผู้จัดทำเลือกนำมาใช้ คือ แบบจำลอง LSTM และ GRU ซึ่งเป็นแบบจำลองที่

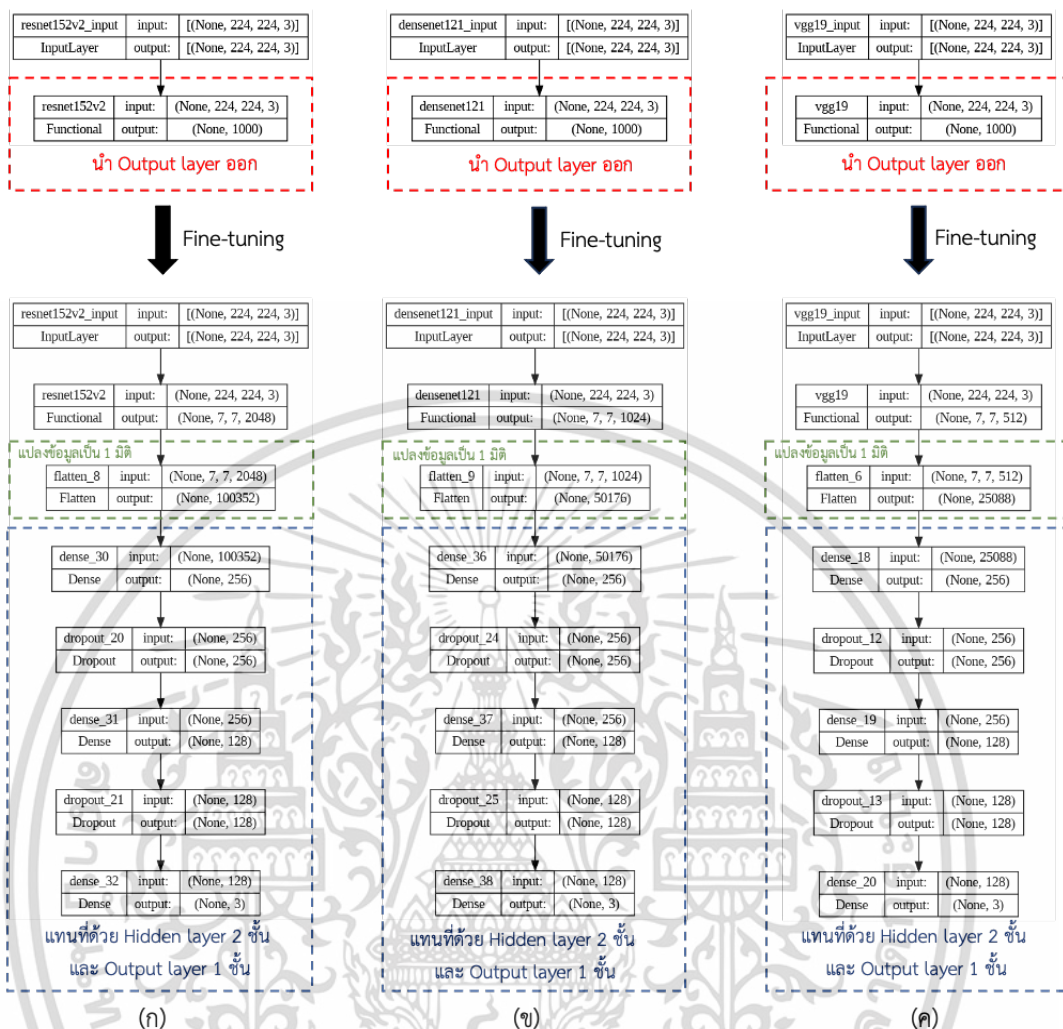
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ถูกพัฒนามาจาก RNN แบบดั้งเดิม โดยในการทดลองแบบจำลอง CNN จะถูกรวมเข้ากับหนึ่งแบบจำลองประเภท RNN ก่อนนำมาเข้าสู่กระบวนการเรียนรู้ด้วยชุดข้อมูลภาพถ่ายรังสีที่รวบรวมได้โดยใช้หลักการถ่ายโอนข้อมูลการเรียนรู้ และทำการเปรียบเทียบประสิทธิภาพของโครงข่ายทั้งหมด เพื่อให้ได้ซึ่งแบบจำลอง CNN-RNN ที่มีประสิทธิภาพสูงสุดในการจำแนกภาวะโรคทรวงอกแบบหลายหมวดหมู่ นอกจากนี้ผู้จัดทำยังได้นำเฉพาะแบบจำลอง Pre-trained CNN ที่กล่าวไว้ข้างต้นมาทำการทดลองเพื่อเปรียบเทียบค่าประสิทธิภาพระหว่างการใช้เพียงแบบจำลอง CNN และแบบจำลอง CNN-RNN อีกด้วยการพัฒนาโครงข่ายทั้งหมดในงานวิจัยนี้ดำเนินการด้วยภาษา Python3 [49] โดยโครงข่ายแบบจำลองเหล่านี้สามารถเข้าถึงได้จาก Keras [50] ที่เปิดให้สามารถใช้งานฟังก์ชันการถ่ายโอนค่าถ่วงน้ำหนักของโครงข่าย [51] การประยุกต์ใช้ และพัฒนาโครงข่ายที่น่าเสนอมีรายละเอียดดังต่อไปนี้

3.3.1 การพัฒนาโครงข่าย CNN สำหรับจำแนกโรคทรวงอกจากภาพรังสี

ในการนำ Pre-trained CNN ผ่านกระบวนการเรียนรู้ด้วยชุดข้อมูล ImageNet มาประยุกต์ใช้ในการจำแนกหมวดหมู่โรคทรวงอก นั้นต้องทำการปรับแต่งโครงสร้างของแบบจำลอง (Fine-tuning) ซึ่งสามารถทำได้โดยการเพิ่มชิ้นการทำงานของโครงข่าย หรือ แทนที่บางชิ้นการทำงาน เนื่องจาก Pre-trained CNN ผ่านกระบวนการเรียนรู้ด้วยชุดข้อมูล ImageNet ที่นำเข้ามาจาก Keras นั้นมีหมวดหมู่ในการจำแนกมากถึง 1,000 หมวดหมู่ ซึ่งทำให้แบบจำลองไม่เหมาะสมกับการจำแนกหมวดหมู่โรคทรวงอกที่มีความจำเพาะเพียง 3 หมวดหมู่ ดังนั้นการปรับแต่งโครงสร้างของแบบจำลองจึงมีความสำคัญ เพื่อให้โครงข่ายมีความเหมาะสมต่อการประมวลผลกับชุดข้อมูลภาพรังสีทรวงอกที่ใช้ และให้ผลลัพธ์ที่จำเพาะตามจำนวนหมวดหมู่โรคของชุดข้อมูล

โดยในงานวิจัยนี้ได้นำเข้าแบบจำลอง Pre-trained CNN โดยกำหนดให้ Input layer รับขนาดของข้อมูลภาพขาเข้าเป็น $224 \times 224 \times 3$ และมีโครงสร้างของแบบจำลองตามโครงข่าย CNN นั้น ๆ ที่เลือกนำมาประยุกต์ ก่อนทำการปรับแต่งแบบจำลอง โดยตัด Output layer ของ Pre-trained CNN ที่มีจำนวน 1,000 โหนดออก ก่อนทำการแปลงข้อมูลผลลัพธ์ที่ได้จากโครงข่าย CNN ที่มีหลายมิติ ให้เป็น 1 มิติ เพื่อเตรียมข้อมูลให้อยู่ในรูปแบบที่พร้อมสำหรับเป็นข้อมูลขาเข้าของ Hidden layer แล้วแทนที่ด้วย Hidden layer จำนวน 2 ชั้นที่มี ReLU เป็น Activation function เพื่อเพิ่มหน่วยการประมวลผลความจำเพาะของโรคทรวงอก โดยกำหนดให้มีจำนวนโหนดใน Hidden layer ที่เพิ่มขึ้นมาเป็น 256 และ 128 ตามลำดับ โดย ในลำดับ และลำดับสุดท้ายเป็น Output layer ที่มีจำนวน 3 โหนด ตามจำนวนหมวดหมู่ของภาพรังสีทรวงอกโดยใช้ Softmax เป็น Activation function ซึ่งในการส่งผ่านข้อมูลระหว่างชั้นใช้อักรากการสุ่มทิ้งข้อมูล (Drop out) เท่ากับ 0.5 เพื่อป้องกันปัญหา Overfitting การปรับแต่งโครงสร้างแบบจำลอง CNN ที่นำเสนอมีแผนผังสถาปัตยกรรมแสดงได้ดังรูปที่ 3.4



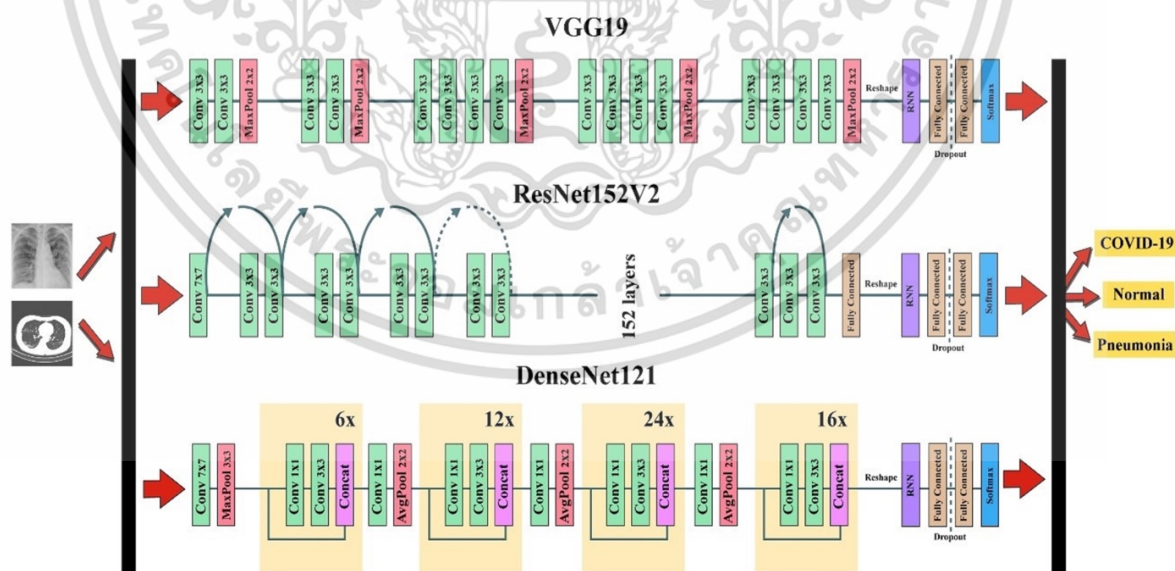
รูปที่ 3.4 แผนผังสถาปัตยกรรมแบบจำลอง CNN ที่นำเสนอ (ก) โครงสร้างแบบจำลอง ResNet152V2 (ข) โครงสร้างแบบจำลอง DenseNet121 (ค) โครงสร้างแบบจำลอง VGG19

3.3.2 การพัฒนาโครงข่าย CNN-RNN สำหรับจำแนกโรคตรวจออกจากภาพรังสี

โครงข่ายแบบจำลอง CNN-RNN เป็นสถาปัตยกรรมที่มีความสามารถในการจำแนกหมวดหมู่ภาพที่มีความซับซ้อนได้ โดยอาศัยการพึ่งพากันของ CNN ที่มีคุณสมบัติเด่นในการสกัดคุณลักษณะสำคัญของรูปภาพ มารวมเข้ากับแบบจำลอง RNN ในการหาความสัมพันธ์ของคุณลักษณะภายในภาพ เพื่อปรับคุณสมบัติตามผลลัพธ์การประมวลผลคุณลักษณะก่อนหน้า โดยมีแนวคิด คือ การปรับพื้นที่ความสนใจภายในภาพ เพื่อให้ CNN สามารถมุ่งความสนใจไปที่บริเวณต่าง ๆ ของข้อมูลภาพ เมื่อคาดเดาหมวดหมู่ของข้อมูลที่แตกต่างกัน โดย RNN เรียนรู้จากคุณลักษณะที่สกัดได้โดย CNN เพื่อสร้างความเข้าใจเชิงลึกเกี่ยวกับความสัมพันธ์ระหว่างข้อมูลภาพ และหมวดหมู่ของข้อมูลได้ดีขึ้น ทำให้สามารถจำแนกหมวดหมู่ที่มีความละเอียดทางคุณลักษณะได้ดียิ่งขึ้น [52]

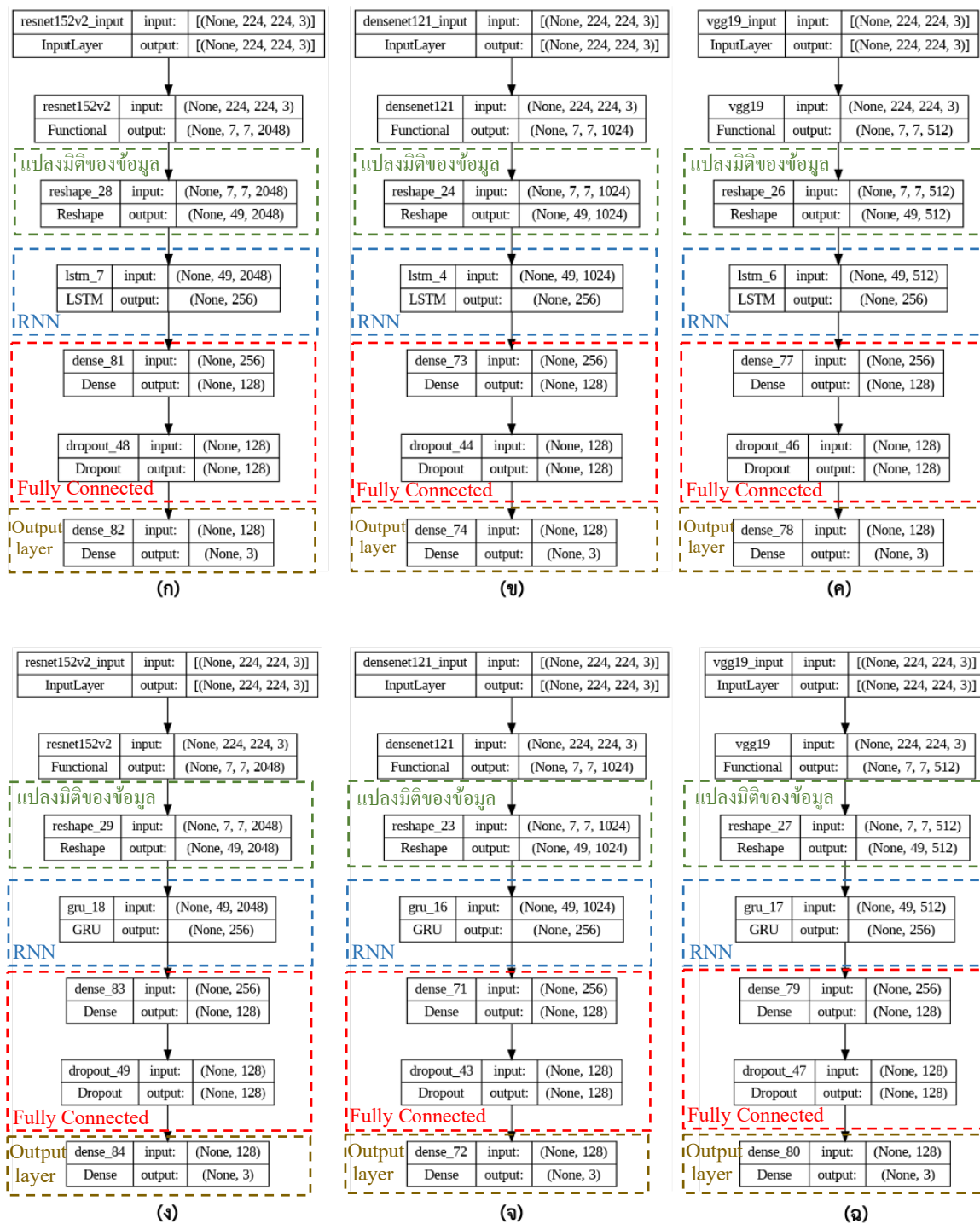
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในการพัฒนาโครงข่าย CNN-RNN ซึ่งแบบจำลอง CNN ที่ใช้เป็นโครงข่ายที่ผ่านกระบวนการเรียนรู้แล้วด้วยชุดข้อมูล ImageNet จำเป็นต้องปรับแต่งแบบจำลองดังกล่าวไว้ในหัวข้อ การพัฒนาโครงข่าย CNN สำหรับจำแนกโรคทรวงอกจากภาพรังสี โดยในการพัฒนาโครงข่าย CNN-RNN สำหรับจำแนกโรคทรวงอก 3 หมวดหมู่นี้ ผู้วิจัยได้ทำการนำ Output layer ของ Pre-trained CNN ที่มีจำนวน 1,000 โหนดออก ก่อนทำการปรับมิติข้อมูลผลลัพธ์ที่ได้จากโครงข่าย CNN นั้น ๆ ให้มีรูปแบบมิติ ที่พร้อมสำหรับป้อนเป็นข้อมูลขาเข้าของ RNN ซึ่งผลลัพธ์ของ CNN จะอยู่ในรูป 4 มิติ ที่แสดงถึง (ขนาดของกลุ่มข้อมูล, ความกว้างของคุณลักษณะ, ความยาวของคุณลักษณะ, จำนวนคุณลักษณะทั้งหมด) โดยต้องถูกปรับให้อยู่ในรูป 3 มิติ ที่แสดงถึง (ขนาดของกลุ่มข้อมูล, ความกว้างของคุณลักษณะ, จำนวนคุณลักษณะทั้งหมด) โดยเกิดจากการนำเอาความยาวของคุณลักษณะมาเรียงต่อกันเป็นลำดับชุดข้อมูลในแนวนอน โดยผลลัพธ์ของ ResNet152V2 DenseNet121 และ VGG19 ถูกปรับมิติผลลัพธ์จาก (None, 7, 7, 2048) (None, 7, 7, 1024) และ (None, 7, 7, 512) เป็น (None, 49, 2048) (None, 49, 1024) และ (None, 49, 512) ตามลำดับ ก่อนนำมาประมวลผลด้วยแบบจำลอง RNN ซึ่งคือ LSTM และ GRU แล้วจึงส่งผลลัพธ์ที่ได้ไปยัง Fully connected layer ที่ประกอบไปด้วย Hidden layer ที่มีโหนดจำนวน 128 โหนดที่มี ReLU เป็น Activation function และ Output layer ที่มีจำนวน 3 โหนด ตามจำนวนหมวดหมู่ของภาพรังสีทรวงอกโดยใช้ Softmax เป็น Activation function โดยมีโครงสร้างแบบจำลอง CNN-RNN สำหรับจำแนกโรคทรวงอกจากภาพรังสีที่นำเสนอแสดงในรูปที่ 3.5 และแผนผังสถาปัตยกรรมแบบจำลอง CNN-RNN ที่พัฒนาขึ้นในงานวิจัยนี้แสดงดังรูปที่ 3.6



รูปที่ 3.5 โครงสร้างแบบจำลอง CNN-RNN สำหรับจำแนกโรคทรวงอกจากภาพรังสีที่นำเสนอ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.6 แผนผังสถาปัตยกรรมแบบจำลอง CNN-RNN ที่นำเสนอ (ก) ResNet152V2-LSTM (ข) DenseNet121-LSTM (ค) VGG19-LSTM (ง) ResNet152V2-GRU (จ) DenseNet121-GRU (ฉ) VGG19-GRU

3.4 การจัดเตรียมข้อมูลเพื่อใช้ในกระบวนการเรียนรู้และทดสอบโครงข่าย

การจัดเตรียมข้อมูล (Data preparation) เพื่อใช้ในกระบวนการเรียนรู้ และทดสอบโครงข่าย แสดงแผนผังลำดับขั้นตอนการดำเนินการได้ดังรูปที่ 3.7 ซึ่งแบ่งออกได้เป็น 2 ขั้นตอนดังต่อไปนี้

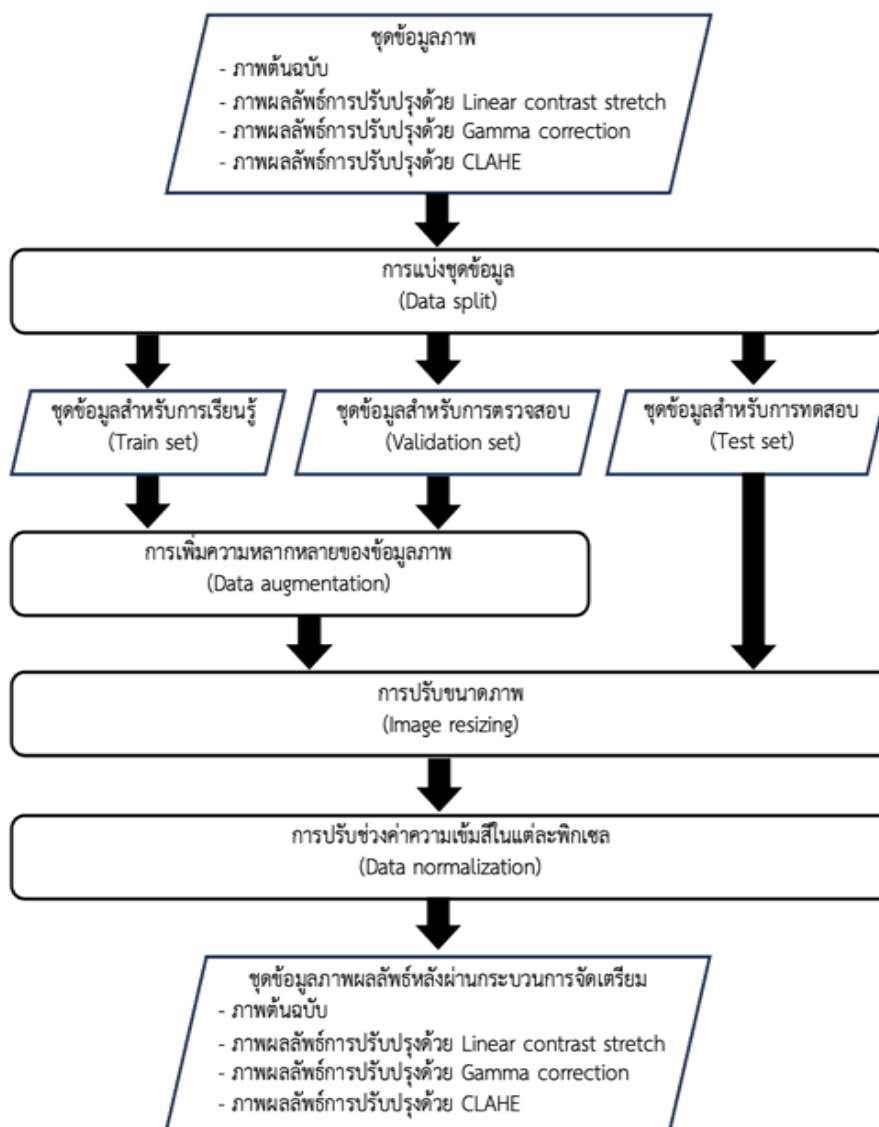
1.) การแบ่งชุดข้อมูลสำหรับกระบวนการเรียนรู้ และทดสอบโครงข่าย (Data split) เป็นขั้นตอนที่นำข้อมูลภาพทั้งหมดจากชุดข้อมูลนั้น ๆ มาทำการสุ่มแบ่งเพื่อใช้ในกระบวนการต่าง ๆ ของแบบจำลอง ซึ่งในงานวิจัยนี้ชุดข้อมูลภาพทั้งหมดที่รวบรวมได้ถูกแบ่งออกเป็น 3 ชุด ได้แก่ ชุดสำหรับการเรียนรู้ (Train set) ชุดสำหรับการตรวจสอบระหว่างการเรียนรู้ (Validation set) และชุดสำหรับการทดสอบ (Test set) โดย 80% ของภาพทั้งหมดใช้สำหรับการเรียนรู้ของแบบจำลอง และอีก 20% ที่เหลือใช้สำหรับการทดสอบ ในส่วนของข้อมูลสำหรับการตรวจสอบระหว่างการเรียนรู้ ได้จากการนำเอา 20% ของภาพในชุดสำหรับการเรียนรู้มาใช้ โดยแสดงจำนวนภาพของแต่ละชุดข้อมูลได้ดังตารางที่ 3.2

ตารางที่ 3.2 จำนวนภาพของชุดข้อมูลที่ใช้ในกระบวนการต่าง ๆ ของแบบจำลอง

ชุดข้อมูล	จำนวนภาพเอกซเรย์ทรวงอก			จำนวนภาพ CT ทรวงอก			จำนวนทั้งหมด
	โควิด-19	ปอดอักเสบ	ปกติ	โควิด-19	ปอดอักเสบ	ปกติ	
Train set	1,750	2,713	1,556	1,452	1,452	1,452	10,375
Test set	750	398	600	500	550	444	3,242
Validation set	437	678	389	363	363	363	2,593
จำนวนทั้งหมด	2,937	3,789	2,545	2,315	2,365	2,259	16,210

2.) การเตรียมข้อมูลก่อนป้อนใช้เป็นข้อมูลขาเข้าของแบบจำลอง (Pre-processing) เป็นกระบวนการที่ใช้ในการปรับชุดข้อมูลให้มีความเหมาะสม และตรงตามเงื่อนไขการรับข้อมูลขาเข้าของแบบจำลองในการประมวลผลต่าง ๆ ซึ่งในงานวิจัยนี้แบ่งออกได้เป็น 3 ขั้นตอน ได้แก่ การเพิ่มความหลากหลายของข้อมูลภาพ (Data augmentation) การปรับขนาดภาพ (Image resizing) และการปรับช่วงค่าของพิกเซลของข้อมูลภาพ (Data normalization) เริ่มจากการนำชุดข้อมูลภาพสำหรับการเรียนรู้มาผ่านการเพิ่มความหลากหลายของข้อมูลภาพ ด้วยเทคนิคการสุ่ม หมุนภาพตั้งแต่ -40 ถึง 40 องศา การเลื่อนตำแหน่งภาพในแนวนอน และแนวตั้ง โดยสุ่มตำแหน่งที่จะเลื่อนอยู่ในช่วงอัตราส่วน -0.2 ถึง 0.2 ของขนาดภาพ เพื่อช่วยหลีกเลี่ยงการเกิดปัญหา Overfitting [53] ถัดมาเนื่องจากในงานวิจัยนี้ชุดข้อมูลภาพรังสีทรวงอกถูกรวบรวมจากหลายแหล่งที่มา ซึ่งมีขนาดของภาพที่แตกต่างกันจึงจำเป็นต้องปรับขนาดของภาพทั้งหมดให้มีขนาดตรงตามเงื่อนไขการรับข้อมูลของแบบจำลอง คือ $224 \times 224 \times 3$ ก่อนนำทั้งชุดข้อมูลภาพสำหรับการเรียนรู้ และข้อมูลสำหรับการทดสอบ มาทำการปรับช่วงค่าความเข้มสีในแต่ละพิกเซลของภาพให้อยู่ในช่วงระหว่าง 0 ถึง 1 ด้วยการนำความเข้มสีในแต่ละพิกเซลของภาพมา

หารด้วย 255 เนื่องจากเป็นภาพจำนวน 8 บิต ที่อยู่ในปริภูมิสี RGB เพื่อช่วยลดความแตกต่างในการกระจายของข้อมูลภาพ

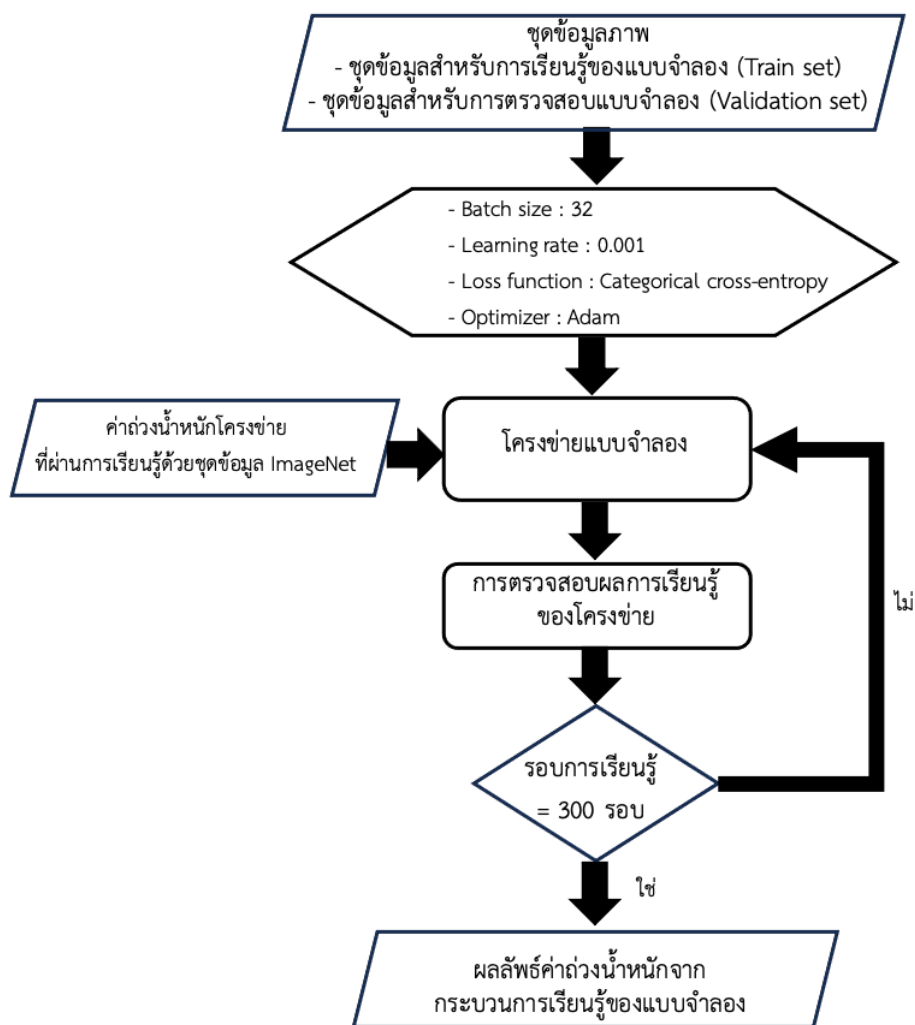


รูปที่ 3.7 ขั้นตอนการจัดเตรียมข้อมูลเพื่อใช้ในกระบวนการเรียนรู้ และทดสอบโครงข่าย

3.5 กระบวนการเรียนรู้ข้อมูลในการจำแนกโรคทรวงอกจากภาพถ่ายรังสีของแบบจำลอง

การเรียนรู้ข้อมูลการจำแนกโรคทรวงอกจากภาพถ่ายรังสีของแบบจำลอง เป็นกระบวนการป้อนข้อมูล เพื่อให้โครงข่ายแบบจำลองทำการเรียนรู้คุณลักษณะของรอยโรคภายในภาพถ่ายรังสีในแต่ละหมวดหมู่ มีขั้นตอนดังแสดงในรูปที่ 3.8 เริ่มจากนำชุดข้อมูลภาพสำหรับการเรียนรู้ และชุดข้อมูลภาพ

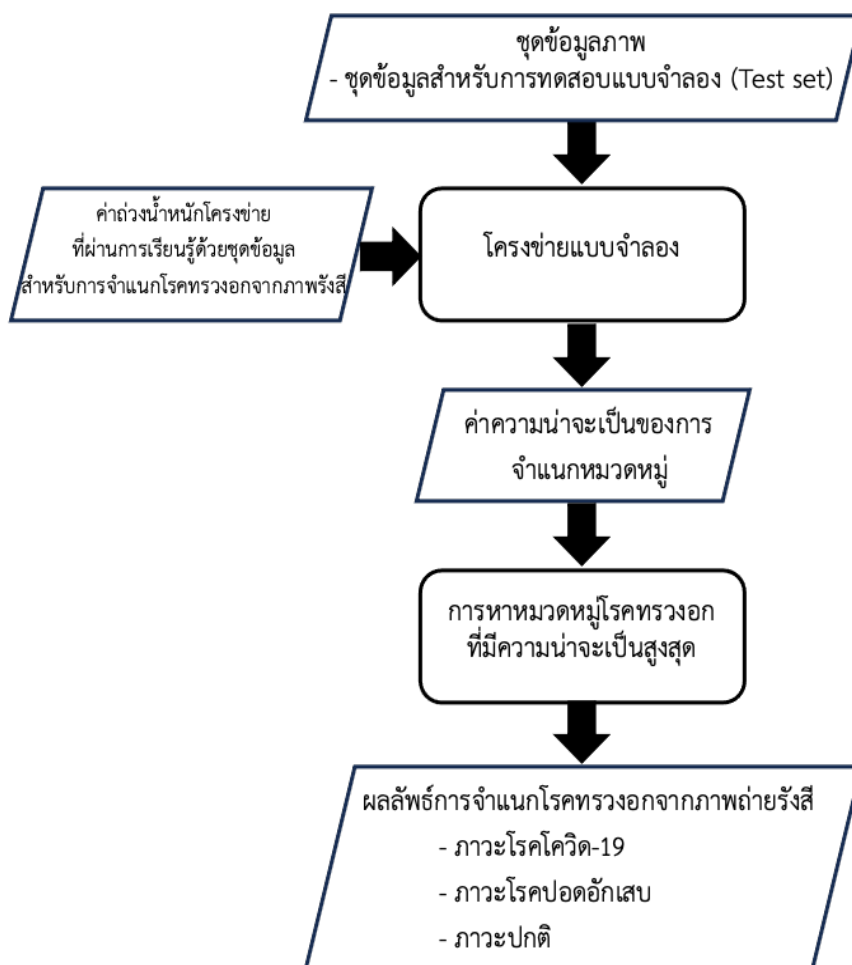
สำหรับการตรวจสอบที่ได้จากกระบวนการจัดเตรียม มาทยอยป้อนเข้าสู่แบบจำลองที่ผ่านกระบวนการเรียนรู้ด้วยชุดข้อมูล ImageNet แล้ว (Pre-trained model) เพื่อปรับค่าถ่วงน้ำหนักของโครงข่ายให้เข้ากับการจำแนกโรคทรวงอกจากภาพถ่ายรังสี ด้วยขนาดของกลุ่มข้อมูล (Batch size) เท่ากับ 32 ภาพ โดยดำเนินการเรียนรู้ทั้งหมด (Epoch) 300 รอบ ด้วยอัตราการเรียนรู้ (Learning rate) เป็น 0.001 โดยใช้ Adam optimizer เป็นเครื่องมือเพิ่มประสิทธิภาพ และคำนวณค่าความคลาดเคลื่อนจากผลลัพธ์การตรวจสอบในกระบวนการเรียนรู้แต่ละรอบด้วยฟังก์ชันการสูญเสีย (Loss function) Categorical cross-entropy เนื่องจากในงานวิจัยนี้เป็นการจำแนกภาพแบบหลายหมวดหมู่ กระบวนการทั้งหมดในการเรียนรู้ถูกประมวลผลด้วย Keras และ TensorFlow [54] บนแพลตฟอร์ม Google Colab pro [55] ที่มี Tesla P100-PCIE-16GB GPU และ RAM 25 GB



รูปที่ 3.8 ขั้นตอนการการเรียนรู้ของโครงข่ายแบบจำลอง

3.6 การจำแนกหมวดหมู่โรคทรวงอก และการประเมินค่าประสิทธิภาพของแบบจำลอง

การทำนายผลการจำแนกโรคทรวงอกจากภาพถ่ายรังสีของโครงข่าย เป็นการจำแนกหมวดหมู่ของข้อมูลภาพรังสีทรวงอกที่ป้อนให้โครงข่ายประมวลผลว่าภาพนั้นจัดอยู่ในภาวะใดของหมวดหมู่โรค มีขั้นตอนดังแสดงในรูปที่ 3.9 เริ่มจากทยอยนำชุดข้อมูลภาพสำหรับการทดสอบที่ได้จากกระบวนการจัดเตรียม มาป้อนเข้าสู่โครงข่ายแบบจำลองที่ผ่านกระบวนการเรียนรู้ข้อมูลสำหรับจำแนกโรคทรวงอก จากภาพถ่ายรังสีแล้ว ด้วยขนาดของกลุ่มข้อมูล (Batch size) เท่ากับ 32 ภาพ โดยผลลัพธ์ที่ได้จากการทำนาย คือ ความน่าจะเป็นของภาพที่ถูกจัดอยู่ในหมวดหมู่โรคนั้น ๆ ก่อนนำหมวดหมู่ที่มีค่าความน่าจะเป็นที่สูงสุดมาใช้เป็นผลลัพธ์การจำแนกโรคทรวงอกจากภาพถ่ายรังสี กระบวนการทั้งหมดในการทำนายผลนี้ถูกประมวลผลด้วย Keras และ TensorFlow บนแพลตฟอร์ม Google Colab pro ที่มี Tesla P100-PCI-E-16GB GPU และ RAM 25 GB



รูปที่ 3.9 ขั้นตอนการการทำนายผลของโครงข่ายแบบจำลอง

ในการประเมินประสิทธิภาพผลลัพธ์การจำแนกโรคทรวงอกจากภาพถ่ายรังสีของโครงข่าย นั้นใช้ตัวชี้วัดทั้งหมด 4 ตัว ในการเปรียบเทียบเพื่อหาโครงข่ายแบบจำลองที่เหมาะสมที่สุดในการจำแนกโรคทรวงอกจากภาพถ่ายรังสี ซึ่งคำนวณได้จากเมทริกซ์วัดประสิทธิภาพ (Confusion matrix) ได้แก่

1.) ค่าความแม่นยำ (Accuracy, ACC) คือ ค่าความสามารถในการจำแนกหมวดหมู่ของภาวะโรคทรวงอกได้ถูกต้องตามความเป็นจริง ซึ่งคำนวณได้จากสมการที่ (24)

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (24)$$

2.) ค่าการทำนายเชิงบวก (Precision) คือ ค่าที่แสดงถึงการทำนายข้อมูลว่าเป็นผลบวกในหมวดหมู่โรค นั้น ๆ ได้ตรงตามความเป็นจริงจากการทำนายข้อมูลว่าเป็นผลบวกทั้งหมด คำนวณได้จากสมการที่ (25)

$$Precision = \frac{TP}{TP+FP} \quad (25)$$

3.) อัตราบวกที่แท้จริง (Recall) คือ ค่าที่บ่งบอกความสามารถในการระบุข้อมูลว่าเป็นผลบวกในหมวดหมู่โรคนั้น ๆ ได้ตรงตามความเป็นจริงจากข้อมูลที่เป็นผลบวกจริงทั้งหมด ซึ่งคำนวณได้จากสมการที่ (26)

$$Recall = \frac{TP}{TP+FN} \quad (26)$$

4.) ค่าคะแนน F1 (F1-score) คือ เป็นค่าเฉลี่ย harmonic ระหว่าง Precision และ Recall ซึ่งหากค่า F1-score มีค่าสูง แสดงถึงแบบจำลองมีความสามารถในการทำนายที่ดี โดยให้ความสำคัญกับทั้งความถูกต้องในการทำนายหมวดหมู่โรค (Precision) และความสามารถในการพบความผิดพลาดในการทำนายหมวดหมู่โรคทั้งหมด (Recall) ซึ่งคำนวณได้จากสมการที่ (27)

$$F1 - score = \frac{2 \times TP}{2 \times TP + (FP + FN)} \quad (27)$$

เมื่อ TP คือ True positive TN คือ True negative FP และ FN หมายถึง False positive และ False negative ตามลำดับ

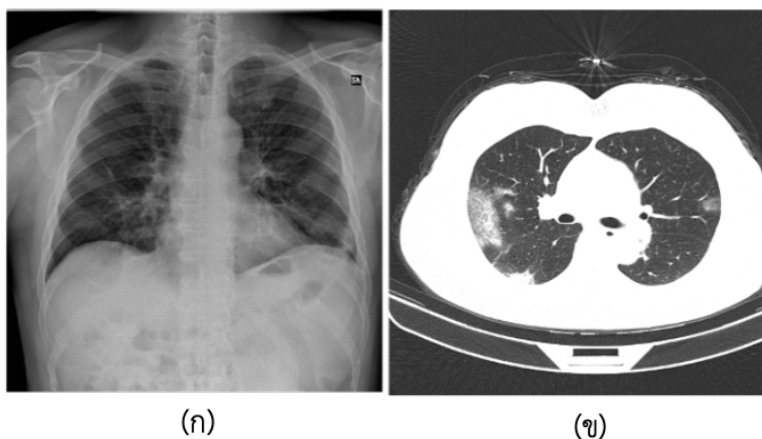
บทที่ 4

ผลการทดลอง

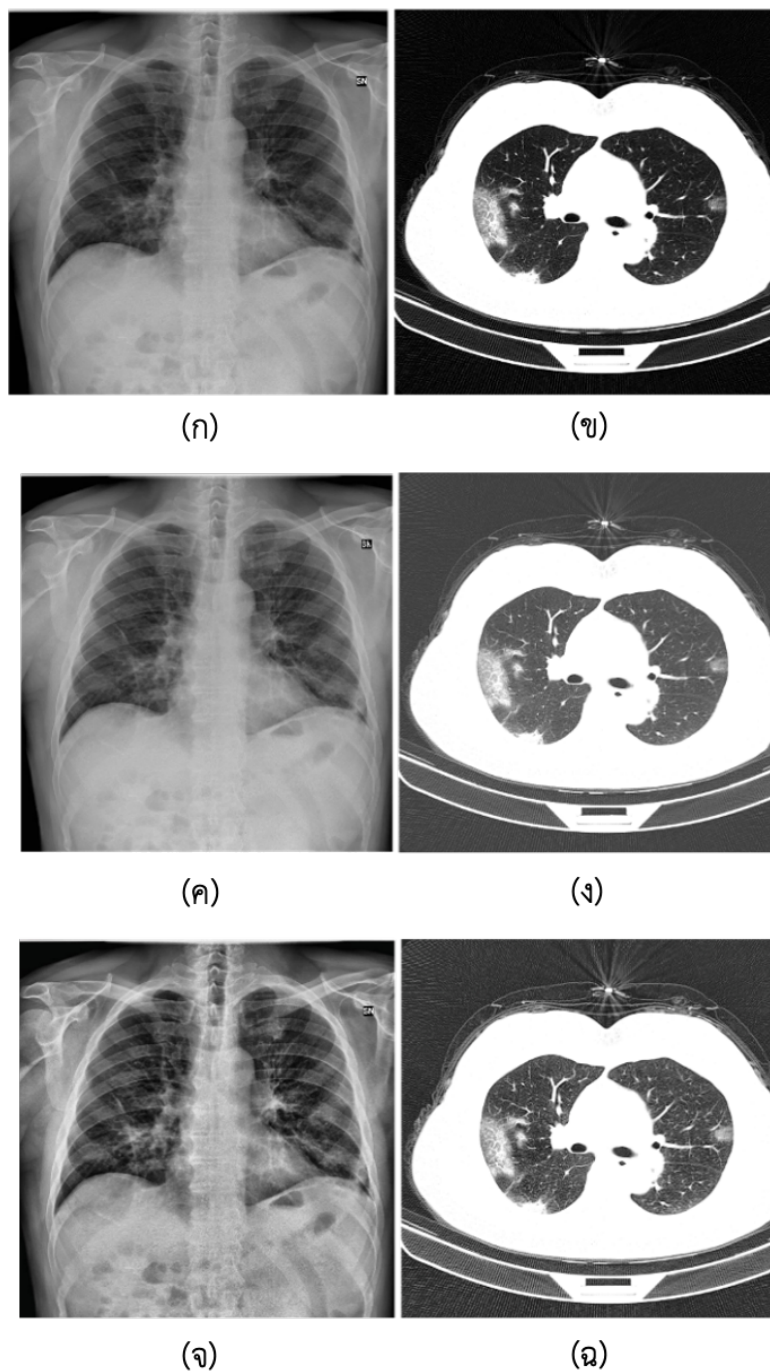
ในบทนี้เป็นการกล่าวถึงผลลัพธ์ในแต่ละขั้นตอนของการจำแนกหมวดหมู่โรคจากภาพเอกซเรย์และภาพ CT ทรวงอก โดยมีผลการทดลองตามลำดับขั้นตอนของวิธีการที่นำเสนอ ดังนี้ อันดับแรกคือ ผลลัพธ์การปรับปรุงคุณภาพสีของชุดข้อมูลภาพรังสีทรวงอกที่รวบรวมได้เพื่อนำมาใช้ในการทดลอง อันดับต่อมา คือ การพัฒนาโครงข่ายแบบจำลองทั้ง CNN และ CNN-RNN สำหรับจำแนกโรคทรวงอกจากภาพถ่ายรังสี ต่อมาเป็นการกล่าวถึงการเตรียมข้อมูลสำหรับการเรียนรู้ และผลลัพธ์ประสิทธิภาพการเรียนรู้ของโครงข่าย จำแนกโรคทรวงอกจากภาพถ่ายรังสี และอันดับสุดท้ายเป็นการกล่าวถึงผลลัพธ์การจำแนกหมวดหมู่ของภาพรังสีทรวงอกตามภาวะโรค และหาค่าประสิทธิภาพการจำแนกของวิธีการที่นำเสนอ รวมถึงเปรียบเทียบค่าประสิทธิภาพกับวิธีการอื่น ๆ

4.1 ผลลัพธ์การปรับปรุงคุณภาพสีของชุดข้อมูลภาพรังสีทรวงอก

ผลการปรับปรุงคุณภาพสีของภาพถ่ายรังสีทรวงอกที่รวบรวมจากชุดข้อมูลต่าง ๆ ช่วยเพิ่มความเด่นชัดของรอยโรคภายในภาพ และลดสัญญาณรบกวนของชุดข้อมูลภาพ ก่อนนำไปเป็นข้อมูลนำเข้าทั้งในส่วนการเรียนรู้ และทำนายผลของโครงข่ายแบบจำลอง โดยรูปที่ 4.1 เป็นภาพเอกซเรย์ และภาพ CT ทรวงอกต้นฉบับก่อนนำมาผ่านกระบวนการปรับปรุงคุณภาพสี รูปที่ 4.2 เป็นภาพเอกซเรย์ และภาพ CT ทรวงอกหลังผ่านกระบวนการปรับปรุงคุณภาพสีด้วยเทคนิค Linear contrast stretch Gamma correction และ Contrast limited adaptive histogram equalization ตามลำดับ



รูปที่ 4.1 ตัวอย่างภาพรังสีทรวงอกต้นฉบับ (ก) ภาพเอกซเรย์ (ข) ภาพ CT



รูปที่ 4.2 ตัวอย่างภาพรังสีทรวงอกหลังผ่านกระบวนการปรับปรุงคุณภาพสี (ก) ภาพเอกซเรย์ที่ผ่านการปรับปรุงด้วยวิธี Linear contrast stretch (ข) ภาพ CT ผ่านการปรับปรุงด้วยวิธี Linear contrast stretch (ค) ภาพเอกซเรย์ที่ผ่านการปรับปรุงด้วยวิธี Gamma correction (ง) ภาพ CT ผ่านการปรับปรุงด้วยวิธี Gamma correction (จ) ภาพเอกซเรย์ที่ผ่านการปรับปรุงด้วยวิธี CLAHE (ฉ) ภาพ CT ที่ผ่านการปรับปรุงด้วยวิธี CLAHE

4.2 ผลลัพธ์การพัฒนาโครงข่าย CNN และ CNN-RNN สำหรับจำแนกโรคทรวงอกจากภาพถ่ายรังสี

โครงข่ายแบบจำลองที่ถูกพัฒนาขึ้นเพื่อเปรียบเทียบประสิทธิภาพการจำแนกหมวดหมู่โรคทรวงอกจากภาพถ่ายรังสีในงานวิจัยนี้แบ่งออกเป็น 2 ประเภท อันดับแรก คือ แบบจำลองที่ใช้เฉพาะ CNN ซึ่งอันประกอบไปด้วย แบบจำลอง ResNet152V2 แบบจำลอง DenseNet121 และแบบจำลอง VGG19 อีกประเภทเป็นแบบจำลอง CNN-RNN ประกอบไปด้วย แบบจำลอง ResNet152V2-LSTM แบบจำลอง ResNet152V2-GRU แบบจำลอง DenseNet121-LSTM แบบจำลอง DenseNet121-GRU แบบจำลอง VGG19-LSTM และแบบจำลอง VGG19-GRU โดยมีข้อมูลโครงสร้างดังแสดงในตารางที่ 4.1 โดย

- จำนวนชั้นการทำงาน (Layers) เป็นจำนวนชั้นการทำงานทั้งหมดของโครงข่าย
- พารามิเตอร์ทั้งหมด คือ จำนวนพารามิเตอร์ในโมเดลที่ประกอบด้วย ค่าถ่วงน้ำหนักของโครงข่าย (weights) และค่าไบแอส (biases)
- พารามิเตอร์ที่ได้รับการฝึก เป็นพารามิเตอร์ที่ถูกปรับค่าในระหว่างกระบวนการเรียนรู้
- พารามิเตอร์ที่ไม่ได้รับการฝึก เป็นพารามิเตอร์ที่ถูกกำหนดค่าคงที่ไว้ล่วงหน้าและไม่เปลี่ยนแปลงระหว่างการเรียนรู้

ตารางที่ 4.1 ข้อมูลโครงสร้างแบบจำลองที่พัฒนาขึ้นสำหรับจำแนกโรคทรวงอกจากภาพถ่ายรังสี

แบบจำลอง CNN	แบบจำลอง RNN	จำนวนชั้นการทำงาน	พารามิเตอร์ทั้งหมด	พามิเตอร์ที่ได้รับการฝึก	พามิเตอร์ที่ไม่ได้รับการฝึก
ResNet152V2	-	570	84,055,299	83,911,555	143,744
ResNet152V2	LSTM	572	73,308,163	73,164,419	143,744
ResNet152V2	GRU	571	69,573,123	69,429,379	143,744
DenseNet121	-	433	19,916,099	19,832,451	83,648
DenseNet121	LSTM	435	14,673,987	14,590,339	83,648
DenseNet121	GRU	434	12,773,955	12,690,307	83,648
VGG19	-	28	26,480,451	26,480,451	0
VGG19	LSTM	30	23,990,851	23,990,851	0
VGG19	GRU	39	23,008,323	23,008,323	0

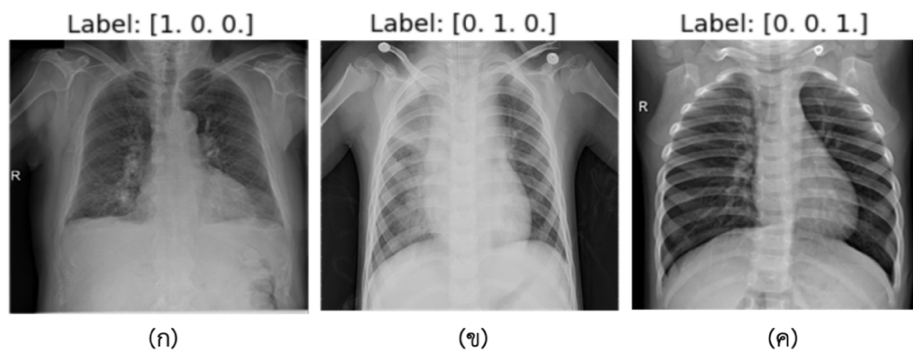
จากตารางที่ 4.1 แสดงให้เห็นว่าการใช้เฉพาะแบบจำลอง CNN ที่ถูกพัฒนาขึ้นในงานวิจัยนี้มีจำนวนพารามิเตอร์ทั้งหมด และพารามิเตอร์ที่ได้รับการฝึกมากกว่า เพราะมีชั้น Fully connected layer ที่มีการเชื่อมโยงกันทุกโหนดมากกว่า

4.3 การเรียนรู้ข้อมูลในการจำแนกโรคทรวงอกจากภาพถ่ายรังสีของแบบจำลอง

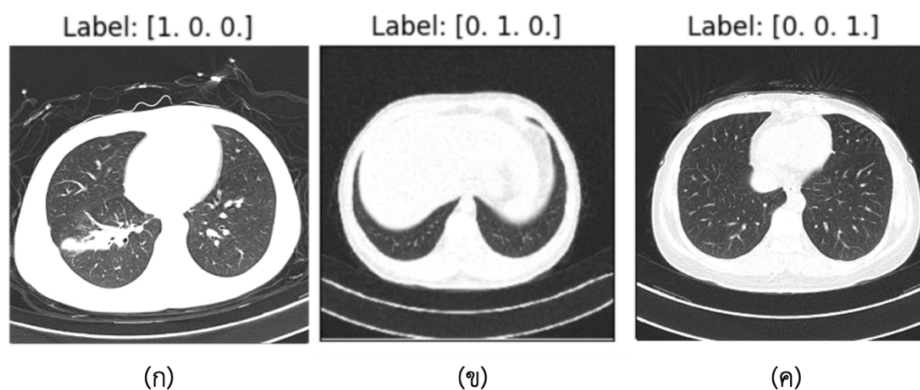
กระบวนการเรียนรู้ข้อมูลการจำแนกโรคทรวงอกจากภาพถ่ายรังสีของแบบจำลองที่พัฒนาขึ้น ใช้ชุดข้อมูลภาพรังสีทรวงอกทั้งภาพเอกซเรย์ และภาพ CT ในชุดข้อมูลสำหรับการเรียนรู้จำนวน 10,375 ภาพ และชุดข้อมูลสำหรับการตรวจสอบจำนวน 2,593 ภาพ มาป้อนให้แบบจำลอง Pre-trained CNN ที่ผ่านกระบวนการเรียนรู้ด้วยชุดข้อมูล ImageNet แล้ว เพื่อเรียนรู้ข้อมูลจำเพาะในการจำแนกโรคทรวงอก โดยดำเนินการเรียนรู้ทั้งหมด (Epoch) 300 รอบ ด้วยขนาดของกลุ่มข้อมูล (Batch size) เท่ากับ 32 ภาพ อัตราการเรียนรู้ (Learning rate) เป็น 0.001 โดยใช้ Adam optimizer เป็นเครื่องมือเพิ่มประสิทธิภาพ และคำนวณค่าความคลาดเคลื่อนจากผลลัพธ์การตรวจสอบในกระบวนการเรียนรู้แต่ละรอบด้วยฟังก์ชันการสูญเสีย (Loss function) Categorical cross-entropy โดยแบ่งการเรียนรู้ของแบบจำลองแต่ละโครงข่าย CNN-RNN ออกเป็น 4 การเรียนรู้ตามชุดข้อมูลภาพที่ใช้ ได้แก่ ชุดข้อมูลภาพต้นฉบับ ชุดข้อมูลภาพที่ผ่านการปรับปรุงด้วยวิธี Linear contrast stretch ชุดข้อมูลภาพที่ผ่านการปรับปรุงด้วยวิธี Gamma correction และชุดข้อมูลภาพที่ผ่านการปรับปรุงด้วยวิธี CLAHE แต่ในส่วนของโครงข่าย CNN ทดลองโดยใช้เพียงชุดข้อมูลภาพต้นฉบับเท่านั้น

4.3.1 ข้อมูลสำหรับการเรียนรู้

ข้อมูลที่ใช้ในการเรียนรู้การจำแนกโรคทรวงอกจากภาพถ่ายรังสีของแบบจำลองที่พัฒนาขึ้นนั้น ประกอบไปด้วยภาพรังสีทรวงอก 3 ภาวะโรคทรวงอก และป้ายกำกับภาวะโรคทรวงอกของแต่ละภาพ (Label) ซึ่งใช้เป็นเป้าหมายการเรียนรู้ เพื่อให้แบบจำลองปรับค่าถ่วงน้ำหนักของโครงข่าย จากค่าความคลาดเคลื่อนที่ได้จากการคำนวณด้วย Loss function โดยป้ายกำกับภาวะโรคทรวงอกอยู่ในรูปแบบของ One-Hot Encoding ที่มีสมาชิก 3 ตัว โดยแทนหมวดหมู่ของภาวะโรคโควิด-19 ด้วย [1, 0, 0] ภาวะโรคปอดอักเสบด้วย [0, 1, 0] และภาวะปกติด้วย [0, 0, 1] ตัวอย่างภาพเอกซเรย์ และภาพ CT พร้อมป้ายกำกับภาวะโรคทรวงอก ดังแสดงดังรูปที่ 4.3 และ 4.4

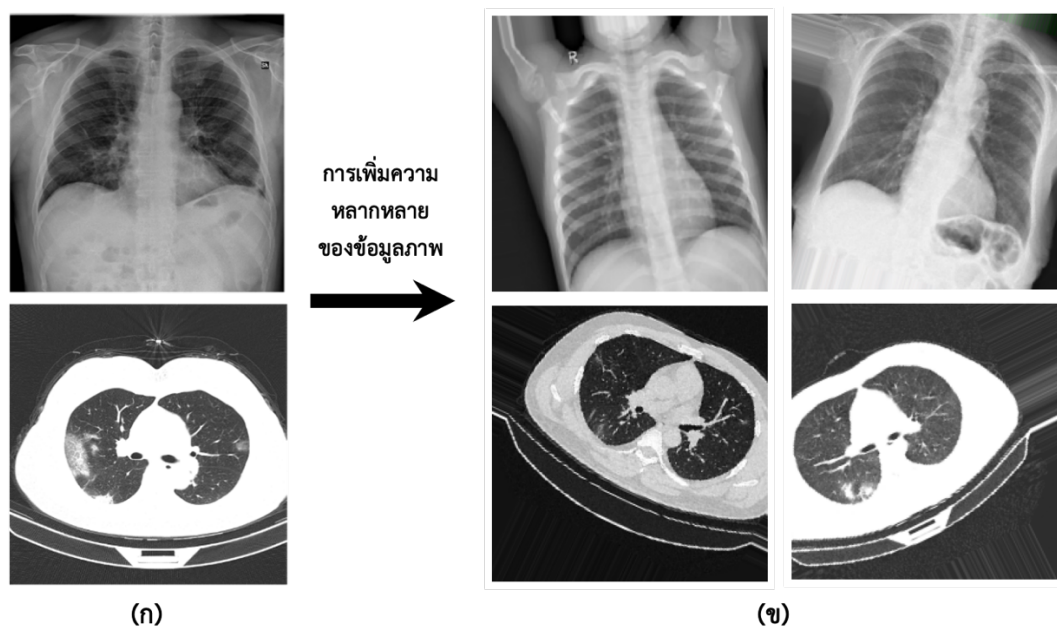


รูปที่ 4.3 ตัวอย่างข้อมูลภาพเอกซเรย์ และป้ายกำกับที่ใช้ในการเรียนรู้ (ก) ภาวะโรคโควิด-19 (ข) ภาวะโรคปอดอักเสบ (ค) ภาวะปกติ



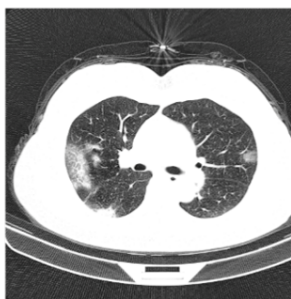
รูปที่ 4.4 ตัวอย่างข้อมูลภาพ CT และป้ายกำกับที่ใช้ในการเรียนรู้ (ก) ภาวะโรคโควิด-19 (ข) ภาวะโรคปอดอักเสบ (ค) ภาวะปกติ

ถัดมาชุดข้อมูลภาพสำหรับการเรียนรู้ และชุดข้อมูลภาพสำหรับการตรวจสอบ ถูกมาผ่านกระบวนการเพิ่มความหลากหลายของข้อมูลภาพ (Data augmentation) เพื่อช่วยหลีกเลี่ยงการเกิดปัญหา Overfitting ด้วยเทคนิคการสุ่ม หมุนภาพตั้งแต่ -40 ถึง 40 องศา การเลื่อนตำแหน่งภาพในแนวนอน และแนวตั้ง โดยสุ่มตำแหน่งที่จะเลื่อนอยู่ในช่วงอัตราส่วน -0.2 ถึง 0.2 ของขนาดภาพ และจัดการพิกเซลที่ว่างหลังจากกระบวนการด้วยการใช้ค่าจากพิกเซลที่อยู่ใกล้เคียง (Nearest interpolation) ดังแสดงในรูปที่ 4.5



รูปที่ 4.5 ตัวอย่างการเพิ่มข้อมูลเพิ่มความหลากหลายของข้อมูลภาพเอกซเรย์ และภาพ CT (ก) ภาพต้นฉบับ (ข) ภาพผลลัพธ์หลังจากกระบวนการเพิ่มความหลากหลาย

หลังจากผ่านกระบวนการเพิ่มความหลากหลายของข้อมูลภาพแล้ว ชุดข้อมูลภาพถูกนำไปปรับขนาดเป็น $224 \times 224 \times 3$ เพื่อให้สามารถนำไปใช้กับแบบจำลองได้ตามมาตรฐานที่กำหนดไว้ก่อนนำไปทำการปรับช่วงค่าความเข้มสีในแต่ละพิกเซลของภาพให้อยู่ในช่วงระหว่าง 0 ถึง 1 ด้วยการนำความเข้มสีในแต่ละพิกเซลของภาพมาหารด้วย 255 ก่อนนำไปใช้เป็นข้อมูลขาเข้าของแบบจำลองในกระบวนการเรียนรู้ ดังแสดงในรูปที่ 4.6



(ก)

```
[ 0 0 0] [33 33 33] [36 36 36] ... [39 39 39] [ 0 0 0] [ 0 0 0]
[ 0 0 0] [32 32 32] [36 36 36] ... [56 56 56] [ 0 0 0] [ 0 0 0]
[ 0 0 0] [50 50 50] [35 35 35] ... [43 43 43] [ 0 0 0] [ 0 0 0]
```

(ข)

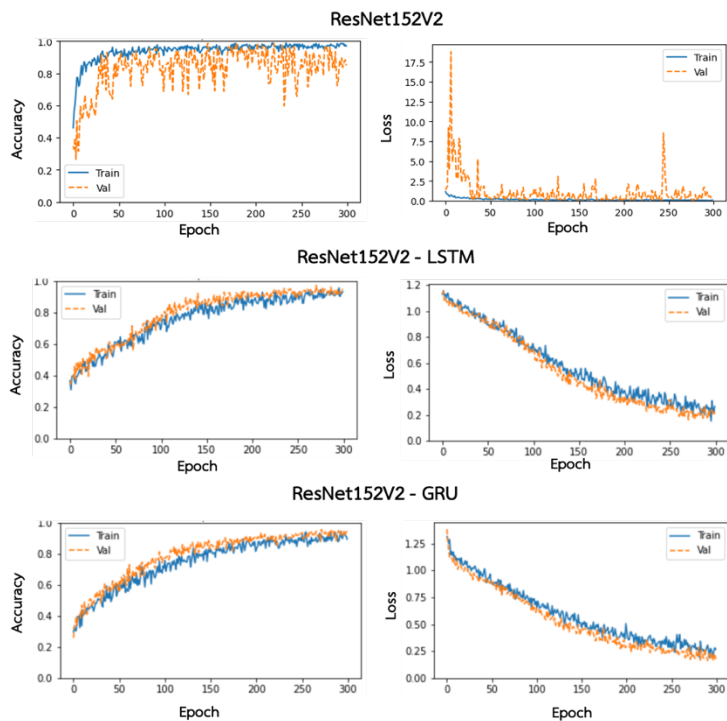
```
[0. 0. 0. ] [0.129 0.129 0.129] [0.141 0.141 0.141] ... [0.152 0.152 0.152] [0. 0. 0.] [0. 0. 0.]
[0. 0. 0. ] [0.125 0.125 0.125] [0.141 0.141 0.141] ... [0.219 0.219 0.219] [0. 0. 0.] [0. 0. 0.]
[0. 0. 0. ] [0.199 0.199 0.199] [0.137 0.137 0.137] ... [0.168 0.168 0.168] [0. 0. 0.] [0. 0. 0.]
```

(ค)

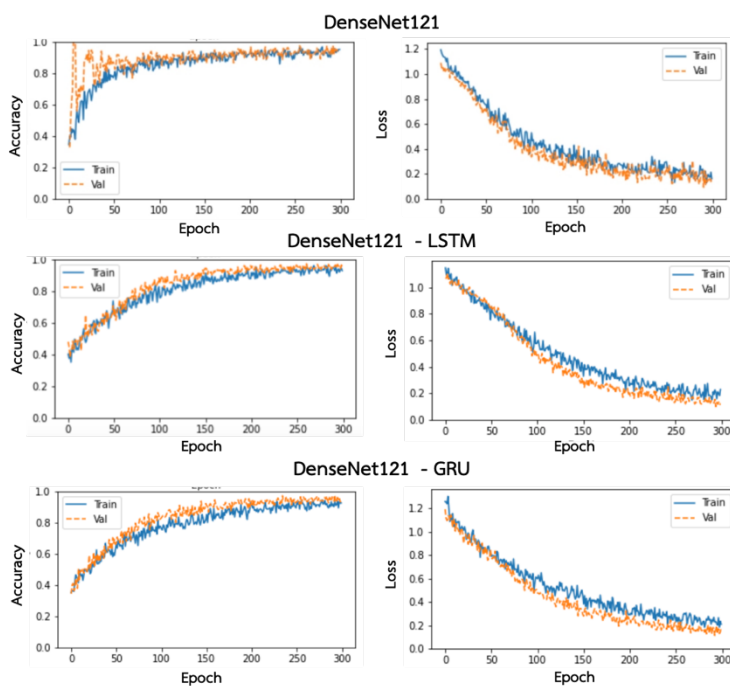
รูปที่ 4.6 การปรับช่วงค่าความเข้มสีในแต่ละพิกเซลของภาพให้อยู่ในช่วงระหว่าง 0 ถึง 1 (ก) ภาพจากชุดข้อมูล (ข) ค่าความเข้มสีของแต่ละพิกเซลในภาพ (ค) ค่าความเข้มสีของแต่ละพิกเซลในภาพหลังผ่านการปรับช่วงค่าความเข้มสี

4.3.2 ผลการเรียนรู้การจำแนกโรคทรวงอกจากภาพถ่ายรังสีของแบบจำลอง

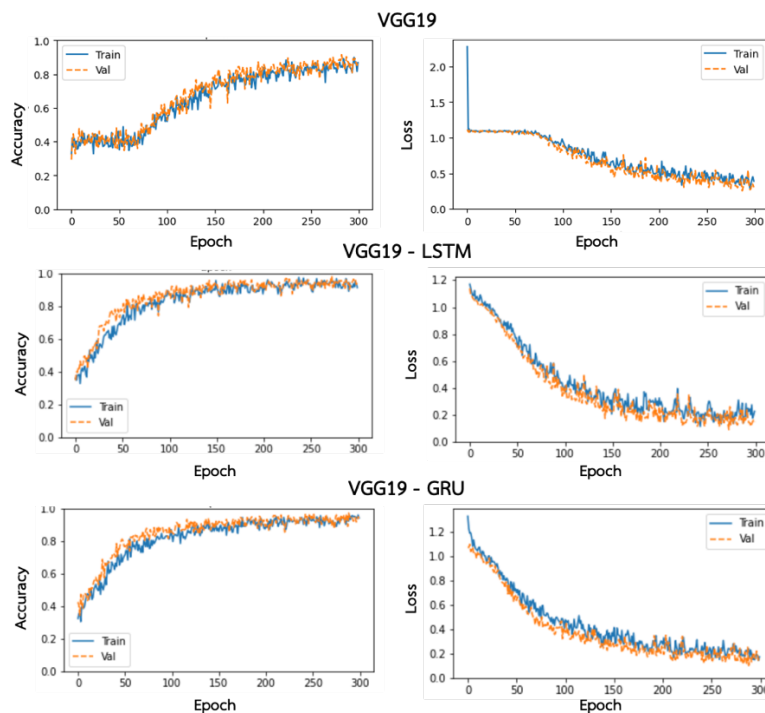
การเรียนรู้ข้อมูลการจำแนกโรคทรวงอกจากภาพถ่ายรังสีของแบบจำลองที่พัฒนาขึ้นในงานวิจัยนี้เป็นการป้อนชุดข้อมูลภาพรังสีทรวงอกสำหรับการเรียนรู้ให้แบบจำลองทำการเรียนรู้คุณลักษณะของข้อมูลภาพ และความสัมพันธ์ระหว่างข้อมูลภาพและป้ายกำกับหมวดหมู่โรคทรวงอก และใช้ชุดข้อมูลตรวจสอบในการตรวจสอบผลการเรียนรู้แต่ละรอบเพื่อปรับค่าถ่วงน้ำหนักของโครงข่าย โดยมีตัวอย่างประวัติค่าความแม่นยำ (Accuracy) และค่าความคลาดเคลื่อน (Loss) ในกระบวนการเรียนรู้ของโครงข่ายแต่ละรอบด้วยชุดข้อมูลภาพรังสีทรวงอกต้นฉบับ ดังแสดงในรูปที่ 4.7 ถึง 4.9 โดยแบ่งการแสดงผลตามแบบจำลอง CNN ที่ใช้ในโครงข่าย ได้แก่ ResNet152v2 DenseNet121 และ VGG19 ตามลำดับ



รูปที่ 4.7 กราฟประวัติค่าความแม่นยำ (Accuracy) และค่าความคลาดเคลื่อน (Loss) ในกระบวนการเรียนรู้ของโครงข่ายที่ใช้แบบจำลอง CNN เป็น ResNet152V2 ด้วยชุดข้อมูลภาพต้นฉบับ



รูปที่ 4.8 กราฟประวัติค่าความแม่นยำ (Accuracy) และค่าความคลาดเคลื่อน (Loss) ในกระบวนการเรียนรู้ของโครงข่ายที่ใช้แบบจำลอง CNN เป็น DenseNet121 ด้วยชุดข้อมูลภาพต้นฉบับ



รูปที่ 4.9 กราฟประวัติค่าความแม่นยำ (Accuracy) และค่าความคลาดเคลื่อน (Loss) ในกระบวนการเรียนรู้ของโครงข่ายที่ใช้แบบจำลอง CNN เป็น VGG19 ด้วยชุดข้อมูลภาพต้นฉบับ

จากกราฟประวัติการเรียนรู้ของแบบจำลองที่แสดงในรูปที่ 4.7 ถึง 4.9 แสดงให้เห็นว่าค่าประสิทธิภาพผลลัพธ์ในแต่ละรอบของกระบวนการเรียนรู้มีค่าความแม่นยำในการเรียนรู้ (Train accuracy) และการตรวจสอบ (Validate accuracy) มีแนวโน้มที่เพิ่มขึ้นเข้าใกล้ 100% ในขณะเดียวกันค่าความคลาดเคลื่อนในการเรียนรู้ (Train loss) และการตรวจสอบ (Validate loss) มีแนวโน้มลดลงเข้าใกล้ 0 ซึ่งบ่งบอกว่าแบบจำลองที่หยิบยกนำมาใช้ในการทดลองสามารถเรียนรู้ข้อมูลนำเข้าที่ป้อนให้ได้ดี

ค่าผลลัพธ์การเรียนรู้ และค่าผลลัพธ์การตรวจสอบระหว่างกระบวนการเรียนรู้ของแต่ละโครงข่ายเฉพาะแบบจำลอง CNN ที่ได้จากกระบวนการเรียนรู้รอบสุดท้ายแสดงดังตารางที่ 4.2

ตารางที่ 4.2 ค่าผลลัพธ์การเรียนรู้และการตรวจสอบของโครงข่ายแบบจำลอง CNN

แบบจำลอง CNN	แบบจำลอง RNN	ชุดข้อมูล	Final accuracy (%)		Final loss	
			Train	Validate	Train	Validate
ResNet152V2	-	ต้นฉบับ	96.84	86.61	0.0961	0.1667
DenseNet121			88.99	89.18	0.2375	0.2568
VGG19			82.61	83.10	0.3786	0.4636

ค่าผลลัพธ์การเรียนรู้ และค่าผลลัพธ์การตรวจสอบระหว่างกระบวนการเรียนรู้ของแต่ละโครงข่ายแบบจำลอง CNN-RNN ที่ได้จากกระบวนการเรียนรู้รอบสุดท้ายแสดงดังตารางที่ 4.3

ตารางที่ 4.3 ค่าผลลัพธ์การเรียนรู้และการตรวจสอบของโครงข่ายแบบจำลอง CNN-RNN

แบบจำลอง CNN	แบบจำลอง RNN	ชุดข้อมูล	Final accuracy (%)		Final loss	
			Train	Validate	Train	Validate
ResNet152V2	LSTM	ต้นฉบับ	94.82	94.98	0.2617	0.2342
		Linear contrast stretch	94.39	94.09	0.2681	0.2742
		Gamma correction	95.16	94.98	0.2335	0.2668
		CLAHE	91.05	91.88	0.3390	0.3219
	GRU	ต้นฉบับ	94.14	95.56	0.3214	0.2621
		Linear contrast stretch	93.08	93.44	0.3258	0.2990
		Gamma correction	94.30	94.12	0.2706	0.2821
		CLAHE	89.46	89.58	0.3612	0.3581
DenseNet121	LSTM	ต้นฉบับ	95.44	95.96	0.2337	0.1804
		Linear contrast stretch	95.31	95.89	0.2416	0.2050
		Gamma correction	95.77	96.17	0.1932	0.1895
		CLAHE	91.94	91.24	0.3150	0.3049
	GRU	ต้นฉบับ	95.28	95.53	0.2070	0.2336
		Linear contrast stretch	95.50	95.93	0.2111	0.1812
		Gamma correction	94.61	95.07	0.2344	0.2194
		CLAHE	89.58	90.59	0.3565	0.3424
VGG19	LSTM	ต้นฉบับ	94.76	95.01	0.2820	0.1830
		Linear contrast stretch	93.87	94.33	0.2948	0.2876
		Gamma correction	94.45	95.37	0.2532	0.2441
		CLAHE	86.58	86.34	0.3773	0.3829
	GRU	ต้นฉบับ	95.56	95.74	0.2271	0.2252
		Linear contrast stretch	89.25	90.01	0.3669	0.3433
		Gamma correction	94.06	94.52	0.2582	0.2313
		CLAHE	87.81	88.82	0.3725	0.3631

จากข้อมูลค่าผลลัพธ์การเรียนรู้ และการตรวจสอบระหว่างกระบวนการเรียนรู้ของแต่ละโครงข่ายในตารางที่ 4.2 และ 4.3 จะเห็นได้ว่าการใช้แบบจำลอง ResNet152V2 มีค่าประสิทธิภาพผลลัพธ์ความแม่นยำในการเรียนรู้ (Final train accuracy) และค่าความคลาดเคลื่อน ในการเรียนรู้ (Final train loss) ดีที่สุด แต่ไม่ได้มีผลลัพธ์ค่าประสิทธิภาพความแม่นยำในการตรวจสอบ (Final validate accuracy) สูงที่สุด และหากพิจารณาจากรูปที่ 4.7 จะเห็นได้ว่ากราฟประวัติค่าผลลัพธ์การเรียนรู้ของแบบจำลอง ResNet152V2 มีความแปรปรวนสูงกว่าแบบจำลองอื่น ซึ่งแสดงถึงเสถียรภาพของแบบจำลองที่ต่ำกว่า

4.4 ผลการจำแนกหมวดหมู่โรคทรวงอก และการประเมินค่าประสิทธิภาพของแบบจำลอง

การทำนายผลการจำแนกหมวดหมู่ภาวะโรคทรวงอกจากภาพถ่ายรังสีจำนวน 3 หมวดหมู่ แบ่งเป็น ภาวะโรคโควิด-19 ภาวะโรคปอดอักเสบ และภาวะปกติ ทั้งนี้ใช้ชุดข้อมูลภาพต้นฉบับ ชุดข้อมูลภาพที่ผ่านการปรับปรุงด้วยวิธี Linear contrast stretch ชุดข้อมูลภาพที่ผ่านการปรับปรุงด้วยวิธี Gamma correction และชุดข้อมูลภาพที่ผ่านการปรับปรุงด้วยวิธี CLAHE ด้วยแบบจำลองที่ผ่านกระบวนการเรียนรู้ในงานวิจัยนี้ โดยแบ่งตามชุดข้อมูลที่ใช้ในการเรียนรู้ ซึ่งในการทำนายผลการจำแนกหมวดหมู่ภาวะโรคทรวงอกใช้ภาพถ่ายรังสีทั้งภาพเอกซเรย์ และภาพ CT รวมทั้งมด 3,242 ภาพ ที่ผ่านการนำไปปรับขนาดเป็น $224 \times 224 \times 3$ เพื่อให้สามารถนำไปใช้กับแบบจำลองได้ตามมาตรฐานที่กำหนดไว้ และทำการปรับช่วงค่าความเข้มสีในแต่ละพิกเซลของภาพให้อยู่ในช่วงระหว่าง 0 ถึง 1 ด้วยการนำความเข้มสีในแต่ละพิกเซลของภาพมาหารด้วย 255 แล้ว ด้วยขนาดของกลุ่มข้อมูล (Batch size) 32 ภาพ โดยในการทดสอบการจำแนกโรคทรวงอกของโครงข่าย CNN-RNN ใช้ข้อมูลทั้งชุดข้อมูลภาพต้นฉบับ และชุดข้อมูลที่ผ่านกระบวนการปรับปรุงคุณภาพสีแบบต่าง ๆ แต่ในการทดสอบโครงข่าย CNN ใช้เพียงชุดข้อมูลภาพต้นฉบับเท่านั้น

การจำแนกหมวดหมู่โรคทรวงอกจากภาพถ่ายรังสี มีผลลัพธ์ที่ได้จากแบบจำลองในรูปความน่าจะเป็นของภาพที่ถูกจัดอยู่ในหมวดหมู่นั้น ๆ โดยใช้หมวดหมู่ที่มีค่าความน่าจะเป็นสูงสุดมาเป็นผลลัพธ์การจำแนกภาวะโรคทรวงอกจากภาพถ่ายรังสี และถูกนำมาประเมินประสิทธิภาพผลลัพธ์การจำแนกด้วยตัวชี้วัดทั้งหมด 4 ตัว คือ ค่าความแม่นยำ (Accuracy) ค่าทำนายเชิงบวก (Precision) อัตราบวกแท้จริง (Recall) และค่าคะแนน F1 (F1-score) ซึ่งคำนวณได้จากค่าสถิติการทำนายในเมทริกซ์วัดประสิทธิภาพ

เมทริกซ์วัดประสิทธิภาพการทำนายผลของแบบจำลอง CNN แสดงดังรูปที่ 4.10 และค่าประสิทธิภาพการทำนายผลการจำแนกหมวดหมู่โรคทรวงอกของแบบจำลอง CNN แสดงดังตารางที่ 4.4

		ResNet152v2			DenseNet121			VGG19					
		ค่าที่ทำนาย			ค่าที่ทำนาย			ค่าที่ทำนาย					
		โควิด-19	ปอดอักเสบ	ปกติ	โควิด-19	ปอดอักเสบ	ปกติ	โควิด-19	ปอดอักเสบ	ปกติ			
ชุดข้อมูลภาพต้นฉบับ	ค่าที่แท้จริง	โควิด-19	1212 (37.38%)	7 (0.22%)	31 (0.96%)	โควิด-19	863 (26.62%)	0 (0.00%)	387 (11.94%)	โควิด-19	1158 (35.72%)	25 (0.77%)	67 (2.07%)
	ปอดอักเสบ	2 (0.06%)	916 (28.25%)	30 (0.93%)	3 (0.09%)	765 (23.60%)	180 (5.55%)	74 (2.28%)	824 (25.42%)	50 (1.54%)			
	ปกติ	366 (11.29%)	13 (0.40%)	665 (20.51%)	12 (0.37%)	0 (0.00%)	1032 (31.83%)	643 (19.83%)	21 (0.65%)	380 (11.72%)			

รูปที่ 4.10 เมทริกซ์วัดประสิทธิภาพการทำนายผลของแบบจำลอง CNN

ตารางที่ 4.4 ค่าประสิทธิภาพการจำแนกหมวดหมู่โรคแทรกของโครงข่ายแบบจำลอง CNN

แบบจำลอง CNN	แบบจำลอง RNN	ชุดข้อมูล	ภาวะโรค แทรก	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
ResNet152V2	-	ต้นฉบับ	โควิด-19	87.48	76.71	96.96	85.65
			ปอดอักเสบ	98.40	97.86	96.62	97.24
			ปกติ	86.43	91.60	63.70	75.14
			โดยรวม	86.15	88.72	85.76	86.01
DenseNet121	-	ต้นฉบับ	โควิด-19	87.60	98.29	69.04	81.11
			ปอดอักเสบ	94.36	100.0	80.7	89.32
			ปกติ	82.14	64.54	98.85	78.09
			โดยรวม	82.05	87.61	82.86	82.84
VGG19	-	ต้นฉบับ	โควิด-19	75.05	61.76	92.64	74.11
			ปอดอักเสบ	94.76	94.71	86.92	90.65
			ปกติ	75.91	76.46	36.40	49.32
			โดยรวม	72.86	77.64	71.99	71.36

เมทริกซ์วัดประสิทธิภาพการทำนายผล และค่าประสิทธิภาพการทำนายผลการจำแนกหมวดหมู่โรคแทรกของโครงข่าย CNN-RNN ที่ใช้แบบจำลอง CNN เป็น ResNet121V2 ในแต่ละชุดข้อมูลภาพแสดงดังรูปที่ 4.11 และตารางที่ 4.5 ตามลำดับ

		ResNet152v2-LSTM				ResNet152v2-GRU				
		ค่าที่ทำนาย			ค่าที่ทำนาย			ค่าที่ทำนาย		
		โควิด-19	ปอดอักเสบ	ปกติ	โควิด-19	ปอดอักเสบ	ปกติ	โควิด-19	ปอดอักเสบ	ปกติ
ชุดข้อมูลภาพ ต้นฉบับ	ค่าที่เป็นจริง โควิด-19	1183 (36.49%)	14 (0.43%)	53 (1.63%)	1183 (36.49%)	59 (1.82%)	8 (0.25%)	1183 (36.49%)	59 (1.82%)	8 (0.25%)
	ค่าที่เป็นจริง ปอดอักเสบ	1 (0.03%)	935 (28.84%)	12 (0.37%)	122 (3.76%)	912 (28.13%)	10 (0.31%)	1 (0.03%)	15 (0.46%)	932 (28.75%)
	ค่าที่เป็นจริง ปกติ	183 (5.64%)	10 (0.31%)	851 (26.25%)	1 (0.03%)	15 (0.46%)	932 (28.75%)	185 (5.71%)	9 (0.28%)	850 (26.22%)
ชุดข้อมูลภาพ ที่ปรับปรุงด้วย Linear contrast stretch	ค่าที่เป็นจริง โควิด-19	1198 (36.95%)	6 (0.19%)	46 (1.42%)	1168 (36.03%)	16 (0.49%)	66 (2.04%)	1198 (36.95%)	6 (0.19%)	46 (1.42%)
	ค่าที่เป็นจริง ปอดอักเสบ	1 (0.03%)	930 (28.69%)	17 (0.52%)	1 (0.03%)	899 (27.73%)	48 (1.48%)	1 (0.03%)	15 (0.46%)	827 (25.51%)
	ค่าที่เป็นจริง ปกติ	185 (5.71%)	9 (0.28%)	850 (26.22%)	202 (6.23%)	15 (0.46%)	827 (25.51%)	185 (5.71%)	9 (0.28%)	850 (26.22%)

		ค่าที่ทำนาย			ค่าที่ทำนาย			
		โควิด-19	ปอดอักเสบ	ปกติ	โควิด-19	ปอดอักเสบ	ปกติ	
ชุดข้อมูลภาพ ที่ปรับปรุงด้วย Gamma correction	ค่าที่เป็นจริง	โควิด-19	1175 (36.24%)	15 (0.46%)	60 (1.85%)	1194 (36.83%)	4 (0.12%)	52 (1.60%)
		ปอดอักเสบ	4 (0.12%)	939 (28.96%)	5 (0.15%)	8 (0.25%)	926 (28.56%)	14 (0.43%)
		ปกติ	156 (4.81%)	12 (0.37%)	876 (27.02%)	191 (5.89%)	8 (0.25%)	845 (26.06%)
	ค่าที่เป็นจริง	โควิด-19	1190 (36.71%)	3 (0.09%)	57 (1.76%)	1223 (37.72%)	4 (0.12%)	23 (0.71%)
		ปอดอักเสบ	1 (0.03%)	941 (29.03%)	6 (0.19%)	1 (0.03%)	914 (28.19%)	33 (1.02%)
		ปกติ	156 (4.81%)	7 (0.22%)	881 (27.17%)	296 (9.13%)	9 (0.28%)	739 (22.79%)

รูปที่ 4.11 เมตริกชีวิตประสิทธิภาพการทำนายผลของโครงข่าย CNN-RNN
ที่ใช้แบบจำลอง CNN เป็น ResNet152V2

ตารางที่ 4.5 ค่าประสิทธิภาพการทำนายผลการจำแนกหมวดหมู่โรคทรวงอกของโครงข่าย CNN-RNN
ที่ใช้แบบจำลอง CNN เป็น ResNet152V2

แบบจำลอง CNN	แบบจำลอง RNN	ชุดข้อมูล	ภาวะโรค ทรวงอก	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
ResNet152V2	LSTM	ต้นฉบับ	โควิด-19	92.26	86.54	94.64	90.41
			ปอดอักเสบ	98.86	97.50	98.63	98.06
			ปกติ	92.04	92.90	81.51	86.83
			โดยรวม	91.58	92.31	91.59	91.77
		Linear contrast stretch	โควิด-19	92.66	86.56	95.84	90.96
			ปอดอักเสบ	98.98	98.41	98.10	98.25
			ปกติ	92.07	93.10	81.42	86.87
			โดยรวม	91.86	92.69	91.79	92.03
		Gamma correction	โควิด-19	92.75	88.01	94.00	90.91
			ปอดอักเสบ	98.89	97.20	99.05	98.12
			ปกติ	92.81	93.09	83.91	88.26
			โดยรวม	92.23	92.77	92.32	92.43
		CLAHE	โควิด-19	93.31	88.34	95.20	91.64
			ปอดอักเสบ	99.48	98.95	99.26	99.1
			ปกติ	93.03	93.33	84.39	88.64
			โดยรวม	92.91	93.54	92.95	93.13

ResNet152V2	GRU	ต้นฉบับ	โควิด-19	94.14	90.58	94.64	92.57
			ปอดอักเสบ	98.95	98.11	98.31	98.21
			ปกติ	93.65	92.49	87.36	89.85
			โดยรวม	93.37	93.73	93.44	93.54
		Linear contrast stretch	โควิด-19	91.21	85.19	93.44	89.12
			ปอดอักเสบ	97.53	96.67	94.83	95.74
			ปกติ	89.79	87.89	79.21	83.32
			โดยรวม	89.27	89.92	89.16	89.39
		Gamma correction	โควิด-19	92.13	85.71	95.52	90.35
			ปอดอักเสบ	98.95	98.72	97.68	98.20
			ปกติ	91.83	92.76	80.94	86.45
			โดยรวม	91.46	92.40	91.38	91.67
		CLAHE	โควิด-19	90.01	80.46	97.84	88.30
			ปอดอักเสบ	98.55	98.60	96.41	97.49
			ปกติ	88.86	92.96	70.79	80.37
			โดยรวม	88.71	90.67	88.35	88.72

เมทริกซ์วัดประสิทธิภาพการทำนายผล และค่าประสิทธิภาพการทำนายผลการจำแนกหมวดหมู่โรคทรวงอกของโครงข่าย CNN-RNN ที่ใช้แบบจำลอง CNN เป็น DenseNet121 ในแต่ละชุดข้อมูลภาพแสดงดังรูปที่ 4.12 และตารางที่ 4.6 ตามลำดับ

		DenseNet121-LSTM				DenseNet121-GRU				
		ค่าที่ทำนาย			ค่าที่ทำนาย			ค่าที่ทำนาย		
		โควิด-19	ปอดอักเสบ	ปกติ	โควิด-19	ปอดอักเสบ	ปกติ	โควิด-19	ปอดอักเสบ	ปกติ
ชุดข้อมูลภาพต้นฉบับ	ค่าที่เป็นจริง	โควิด-19	1222 (37.69%)	5 (0.15%)	23 (0.71%)	โควิด-19	1195 (36.86%)	4 (0.12%)	51 (1.57%)	
	ปอดอักเสบ	0 (0.00%)	936 (28.87%)	12 (0.37%)	ปอดอักเสบ	3 (0.09%)	910 (28.07%)	35 (1.08%)		
	ปกติ	277 (8.54%)	10 (0.31%)	757 (23.35%)	ปกติ	242 (7.46%)	12 (0.37%)	790 (24.37%)		
ชุดข้อมูลภาพที่ปรับปรุงด้วย Linear contrast stretch	ค่าที่เป็นจริง	โควิด-19	1193 (36.80%)	2 (0.06%)	55 (1.70%)	โควิด-19	1203 (37.11%)	4 (0.12%)	43 (1.33%)	
	ปอดอักเสบ	1 (0.03%)	940 (28.99%)	7 (0.22%)	ปอดอักเสบ	1 (0.03%)	907 (27.98%)	40 (1.23%)		
	ปกติ	219 (6.76%)	15 (0.46%)	810 (24.98%)	ปกติ	288 (8.88%)	1 (0.03%)	755 (23.29%)		

		ค่าที่ทำนาย			ค่าที่ทำนาย		
		โควิด-19	ปอดอักเสบ	ปกติ	โควิด-19	ปอดอักเสบ	ปกติ
ชุดข้อมูลภาพ ที่ปรับปรุงด้วย Gamma correction	โควิด-19	1227 (37.85%)	1 (0.03%)	22 (0.68%)	1215 (37.48%)	12 (0.37%)	23 (0.71%)
	ปอดอักเสบ	1 (0.03%)	919 (28.35%)	28 (0.86%)	2 (0.06%)	908 (28.01%)	38 (1.17%)
	ปกติ	293 (9.04%)	5 (0.15%)	746 (23.01%)	366 (11.29%)	21 (0.65%)	657 (20.27%)
		ค่าที่เป็นจริง			ค่าที่เป็นจริง		
ชุดข้อมูลภาพ ที่ปรับปรุงด้วย CLAHE	โควิด-19	1221 (37.66%)	1 (0.03%)	28 (0.86%)	1206 (37.20%)	3 (0.09%)	41 (1.26%)
	ปอดอักเสบ	2 (0.06%)	935 (28.84%)	11 (0.34%)	2 (0.06%)	938 (28.93%)	8 (0.25%)
	ปกติ	310 (9.56%)	20 (0.62%)	714 (22.02%)	308 (9.50%)	12 (0.37%)	724 (22.33%)
		ค่าที่เป็นจริง			ค่าที่เป็นจริง		

รูปที่ 4.12 เมตริกชีวิตประสิทธิภาพการทำนายผลของโครงข่าย CNN-RNN
ที่ใช้แบบจำลอง CNN เป็น DenseNet121

ตารางที่ 4.6 ค่าประสิทธิภาพการทำนายผลการจำแนกหมวดหมู่โรคทรวงอกของโครงข่าย CNN-RNN
ที่ใช้แบบจำลอง CNN เป็น DenseNet121

แบบจำลอง CNN	แบบจำลอง RNN	ชุดข้อมูล	ภาวะโรค ทรวงอก	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
DenseNet121	LSTM	ต้นฉบับ	โควิด-19	90.59	81.52	97.76	88.90
			ปอดอักเสบ	99.17	98.42	98.73	98.57
			ปกติ	90.07	95.58	72.51	82.46
			โดยรวม	89.91	91.84	89.67	89.98
		Linear contrast stretch	โควิด-19	91.46	84.43	95.44	89.60
			ปอดอักเสบ	99.23	98.22	99.16	98.69
			ปกติ	90.87	92.89	77.59	84.55
			โดยรวม	90.78	91.85	90.73	90.95
		Gamma correction	โควิด-19	90.22	80.67	98.16	88.56
			ปอดอักเสบ	98.92	99.35	96.94	98.13
			ปกติ	89.27	93.72	71.46	81.09
			โดยรวม	89.20	91.25	88.85	89.26
CLAHE	โควิด-19	89.48	79.65	97.68	87.75		
	ปอดอักเสบ	98.95	97.80	98.63	98.21		
	ปกติ	88.62	94.82	68.39	79.46		
	โดยรวม	88.53	90.76	88.23	88.47		

DenseNet121	GRU	ต้นฉบับ	โควิด-19	90.75	82.99	95.60	88.85
			ปอดอักเสบ	98.33	98.27	95.99	97.12
			ปกติ	89.51	90.18	75.67	82.29
			โดยรวม	89.30	90.48	89.09	89.42
		Linear contrast stretch	โควิด-19	89.64	80.63	96.24	87.75
			ปอดอักเสบ	98.58	99.45	95.68	97.53
			ปกติ	88.53	90.10	72.32	80.24
			โดยรวม	88.37	90.06	88.08	88.51
		Gamma correction	โควิด-19	87.57	76.75	97.20	85.77
			ปอดอักเสบ	97.75	96.49	95.78	96.13
			ปกติ	86.18	91.50	62.93	74.57
			โดยรวม	85.75	88.25	85.30	85.49
		CLAHE	โควิด-19	89.08	79.55	96.48	87.2
			ปอดอักเสบ	99.23	98.43	98.95	98.69
			ปกติ	88.62	93.66	69.35	79.69
			โดยรวม	88.46	90.55	88.26	88.53

เมทริกซ์วัดประสิทธิภาพการทำนายผล และค่าประสิทธิภาพการทำนายผลการจำแนกหมวดหมู่โรคตรวจออกของโครงข่าย CNN-RNN ที่ใช้แบบจำลอง CNN เป็น VGG19 ในแต่ละชุดข้อมูลภาพแสดงดังรูปที่ 4.13 และตารางที่ 4.7 ตามลำดับ

VGG19-LSTM				VGG19-GRU					
		ค่าที่ทำนาย					ค่าที่ทำนาย		
		โควิด-19	ปอดอักเสบ	ปกติ			โควิด-19	ปอดอักเสบ	ปกติ
ชุดข้อมูลภาพต้นฉบับ	โควิด-19	1183 (36.49%)	6 (0.19%)	61 (1.88%)	โควิด-19	1184 (36.52%)	17 (0.52%)	49 (1.51%)	
	ปอดอักเสบ	2 (0.06%)	935 (28.84%)	11 (0.34%)	ปอดอักเสบ	4 (0.12%)	943 (29.09%)	1 (0.03%)	
	ปกติ	196 (6.05%)	8 (0.25%)	840 (25.91%)	ปกติ	220 (6.79%)	24 (0.74%)	800 (24.68%)	
ชุดข้อมูลภาพที่ปรับปรุงด้วย Linear contrast stretch	โควิด-19	1147 (35.38%)	3 (0.09%)	100 (3.08%)	โควิด-19	1193 (36.80%)	0 (0.00%)	57 (1.76%)	
	ปอดอักเสบ	3 (0.09%)	934 (28.81%)	11 (0.34%)	ปอดอักเสบ	8 (0.25%)	867 (26.74%)	73 (2.25%)	
	ปกติ	129 (3.98%)	4 (0.12%)	911 (28.10%)	ปกติ	202 (6.23%)	2 (0.06%)	840 (25.91%)	

		โควิด-19	ปอดอักเสบ	ปกติ			โควิด-19	ปอดอักเสบ	ปกติ	
ชุดข้อมูลภาพ ที่ปรับปรุงด้วย Gamma correction	ค่าที่เป็นจริง	โควิด-19	1234 (38.06%)	4 (0.12%)	12 (0.37%)	ค่าที่เป็นจริง	โควิด-19	1209 (37.29%)	9 (0.28%)	32 (0.99%)
		ปอดอักเสบ	2 (0.06%)	945 (29.15%)	1 (0.03%)		4 (0.12%)	923 (28.47%)	21 (0.65%)	
		ปกติ	379 (11.69%)	9 (0.28%)	656 (20.23%)		291 (8.98%)	15 (0.46%)	738 (22.76%)	
ชุดข้อมูลภาพ ที่ปรับปรุงด้วย CLAHE	ค่าที่เป็นจริง	โควิด-19	1231 (37.97%)	11 (0.34%)	8 (0.25%)	ค่าที่เป็นจริง	โควิด-19	1183 (36.49%)	59 (1.81%)	8 (0.25%)
		ปอดอักเสบ	0 (0.00%)	944 (29.12%)	4 (0.12%)		122 (3.76%)	912 (28.13%)	10 (0.31%)	
		ปกติ	454 (14.00%)	20 (0.62%)	570 (17.58%)		1 (0.03%)	15 (0.46%)	932 (28.75%)	

รูปที่ 4.13 เมตริกชีวิตประสิทธิภาพการทำนายผลของโครงข่าย CNN-RNN
ที่ใช้แบบจำลอง CNN เป็น VGG19

ตารางที่ 4.7 ค่าประสิทธิภาพการทำนายผลการจำแนกหมวดหมู่โรคทรวงอกของโครงข่าย CNN-RNN
ที่ใช้แบบจำลอง CNN เป็น VGG19

แบบจำลอง CNN	แบบจำลอง RNN	ชุดข้อมูล	ภาวะโรค ทรวงอก	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
VGG19	LSTM	ต้นฉบับ	โควิด-19	91.83	85.66	94.64	89.93
			ปอดอักเสบ	99.17	98.52	98.63	98.57
			ปกติ	91.49	92.11	80.46	85.89
			โดยรวม	91.24	92.10	91.24	91.46
		Linear contrast stretch	โควิด-19	92.75	89.68	91.76	90.71
			ปอดอักเสบ	99.35	99.26	98.52	98.89
			ปกติ	92.47	89.14	87.26	88.19
			โดยรวม	92.29	92.69	92.51	92.60
		Gamma correction	โควิด-19	87.75	76.41	98.72	86.14
			ปอดอักเสบ	99.51	98.64	99.68	99.16
			ปกติ	87.63	98.06	62.84	76.60
			โดยรวม	87.45	91.04	87.08	87.30
		CLAHE	โควิด-19	85.41	73.06	98.48	83.89
			ปอดอักเสบ	98.92	96.82	99.58	98.18
			ปกติ	85.01	97.94	54.60	70.11
			โดยรวม	84.67	89.27	84.22	84.06

VGG19	GRU	ต้นฉบับ	โควิด-19	91.05	84.09	94.72	89.09
			ปอดอักเสบ	98.58	95.83	99.47	97.62
			ปกติ	90.93	94.12	76.63	84.48
			โดยรวม	90.28	91.35	90.27	90.40
		Linear contrast stretch	โควิด-19	91.76	85.03	95.44	89.93
			ปอดอักเสบ	97.44	99.77	91.46	95.43
			ปกติ	89.70	86.60	80.46	83.42
			โดยรวม	89.45	90.47	89.12	89.59
		Gamma correction	โควิด-19	89.64	80.39	96.72	87.80
			ปอดอักเสบ	98.49	97.47	97.36	97.41
			ปกติ	88.93	93.30	70.69	80.44
			โดยรวม	88.53	90.39	88.26	88.55
		CLAHE	โควิด-19	92.07	84.70	96.96	90.42
			ปอดอักเสบ	99.07	97.81	99.05	98.43
			ปกติ	91.21	94.59	77.11	84.96
			โดยรวม	91.18	92.37	91.04	91.27

จากผลลัพธ์การทำนายผลการจำแนกโรคที่ตรวจออกจากภาพถ่ายรังสีทั้งแบบจำลอง CNN และแบบจำลอง CNN-RNN ที่แสดงในตารางที่ 4.4 ถึง 4.7 สามารถแสดงโครงข่ายที่ให้ผลลัพธ์การจำแนกโดยรวมดีที่สุดจากทุกโครงข่ายที่พัฒนาขึ้นโดยแบ่งตามแบบจำลอง CNN ที่ใช้ ดังแสดงในตารางที่ 4.8

ตารางที่ 4.8 ผลลัพธ์ประสิทธิภาพที่ดีที่สุดของแต่ละโครงข่ายตามแบบจำลอง CNN ที่ใช้

แบบจำลอง CNN	แบบจำลอง RNN	ชุดข้อมูล	ภาวะโรคที่ตรวจออก	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
ResNet152V2	GRU	ต้นฉบับ	โควิด-19	94.14	90.58	94.64	92.57
			ปอดอักเสบ	98.95	98.11	98.31	98.21
			ปกติ	93.65	92.49	87.36	89.85
			โดยรวม	93.37	93.73	93.44	93.54
DenseNet121	LSTM	Linear contrast stretch	โควิด-19	91.46	84.43	95.44	89.60
			ปอดอักเสบ	99.23	98.22	99.16	98.69
			ปกติ	90.87	92.89	77.59	84.55
			โดยรวม	90.78	91.85	90.73	90.95
VGG19	LSTM	Linear contrast stretch	โควิด-19	92.75	89.68	91.76	90.71
			ปอดอักเสบ	99.35	99.26	98.52	98.89
			ปกติ	92.47	89.14	87.26	88.19
			โดยรวม	92.29	92.69	92.51	92.60

จากตารางที่ 4.8 แสดงให้เห็นว่าการใช้แบบจำลอง CNN-RNN ให้ผลลัพธ์ประสิทธิภาพความแม่นยำโดยรวมสูงกว่าการใช้เฉพาะแบบจำลอง CNN โดยแบบจำลอง ResNet152V2-GRU ที่ใช้ภาพต้นฉบับในการประมวลผลให้ประสิทธิภาพโดยรวมดีที่สุดในการจำแนกภาวะโรคทรวงอก ด้วยค่าความแม่นยำเท่ากับ 93.37% ค่า precision 93.73% ค่า recall 93.44% และ ค่า F1-score 93.54% ซึ่งหากแบ่งค่าประสิทธิภาพในการจำแนกตามกรณีต่าง ๆ ของภาวะโรคทรวงอก ResNet152V2-GRU ที่ใช้ภาพต้นฉบับในการประมวลผลนั้นให้ประสิทธิภาพความแม่นยำสูงที่สุดในการจำแนกภาวะโรคโควิด-19 และภาวะปกติ โดยมีค่าความแม่นยำอยู่ที่ 94.14% และ 93.65% ตามลำดับ ในส่วนของภาวะโรคปอดอักเสบ VGG19-LSTM ที่ใช้เทคนิค Linear contrast stretch ในการปรับปรุงคุณภาพสีของข้อมูลภาพ ขาเข้าก่อนการประมวลผลให้ประสิทธิภาพความแม่นยำในการจำแนกสูงที่สุดอยู่ที่ 99.35% นอกจากนี้ผู้จัดทำได้เปรียบเทียบผลลัพธ์ประสิทธิภาพโดยรวมสูงสุดที่ได้จากวิธีการที่นำเสนอกับวิธีการในงานวิจัยก่อนหน้านี้ ดังแสดงในตารางที่ 4.9

ตารางที่ 4.9 การเปรียบเทียบประสิทธิภาพของวิธีการที่นำเสนอกับงานวิจัยอื่น

ผู้เขียน/งานวิจัย	จำนวนข้อมูล	ประเภทหมวดหมู่	วิธีการ	ตัวบ่งชี้ค่าประสิทธิภาพ (%)			
				ACC	Precision	Recall	F1-score
Ozturk และคณะ [56]	625 X-rays	Multiclass	DCNN	87.02	89.96	85.35	-
Asnaoui และคณะ [57]	6,087 X-rays	Multiclass	Inception-ResNetV2	92.18	92.38	92.11	92.07
Rahimzadeh และคณะ [58]	15,805 X-rays	Multiclass	Xception-ResNet50V2	91.40	72.83	87.31	-
Joshi และคณะ [59]	746 CT	Binary	LiMS-Net	92.11	-	88.77	92.59
Wu และคณะ [40]	495 CT	Binary	ResNet50	76.00	-	81.10	-
Xu และคณะ [41]	618 CT	Multiclass	ResNet+ LocationAttention	86.70	81.30	86.70	83.90
Perumal และคณะ [42]	205 X-rays and 202 CT	Multiclass	VGG16	93.00	91.00	90.00	-
งานวิจัยที่นำเสนอ	9271 X-rays and 6939 CT	Multiclass	ResNet152V2+ GRU	93.37	93.72	93.44	93.54

4.5 เวลาในการประมวลผลของกระบวนการเรียนรู้ และทำนายผลการจำแนกภาวะโรค ตรวจออกจากภาพถ่ายรังสี

เวลาในการประมวลผลของกระบวนการเรียนรู้ (Training time) และเวลาในการทำนายผลการจำแนก (Predict time) ภาวะโรคตรวจออกจากภาพถ่ายรังสี ของแต่ละโครงข่ายแสดงดังตารางที่ 4.10 โดยในส่วนกระบวนการเรียนรู้เป็นเวลาทั้งหมดที่โครงข่ายใช้เรียนรู้ข้อมูลภาพรังสีตรวจออก และส่วนการทำนายผล เป็นเวลาต่อหนึ่งภาพที่ได้จากการหาค่าเฉลี่ยเวลาในการทำนายผลภาพชุดข้อมูลทดสอบทั้งหมด

ตารางที่ 4.10 เวลาที่ใช้ในการเรียนรู้ และทำนายผลการจำแนกภาวะโรคตรวจออกจากภาพถ่ายรังสี

เวลา	ชุดข้อมูล	ResNet152V2			DenseNet121			VGG19		
		-	LSTM	GRU	-	LSTM	GRU	-	LSTM	GRU
การเรียนรู้ (นาที:วินาที)	ต้นฉบับ	156:17	107:29	99:57	163:42	105:32	89:14	122:45	121:39	100:56
	Linear contrast stretch	-	104:41	101:35	-	103:55	92:49	-	115:1	104:28
	Gamma correction	-	108:44	103:49	-	103:26	90:31	-	120:08	104:14
	CLAHE	-	110:05	98:33	-	108:48	94:22	-	114:55	105:49
การทำนาย (วินาที/ภาพ)	ต้นฉบับ	0.49	0.28	0.21	0.27	0.04	0.04	0.8	0.13	0.12
	Linear contrast stretch	-	0.28	0.28	-	0.04	0.04	-	0.18	0.16
	Gamma correction	-	0.29	0.28	-	0.04	0.03	-	0.18	0.16
	CLAHE	-	0.29	0.26	-	0.05	0.05	-	0.17	0.16

จากข้อมูลในตารางที่ 4.10 หากเปรียบเทียบโครงข่ายที่ใช้เฉพาะ CNN พบว่าแบบจำลอง VGG19 ใช้เวลาในการเรียนรู้ที่น้อยที่สุดโดยใช้เวลาอยู่ที่ 122 นาที 45 วินาที และ ResNet152V2 ใช้เวลาในการเรียนรู้มากที่สุดเป็นเวลา 156 นาที 17 วินาที แต่หากเปรียบเทียบเฉพาะโครงข่าย CNN-RNN พบว่าแบบจำลอง DenseNet121-GRU ใช้เวลาในการเรียนรู้ที่น้อยที่สุดอยู่ที่ 89 นาที 14 วินาที ในทางกลับกัน VGG19-LSTM ใช้เวลาในการเรียนรู้มากที่สุดเป็นเวลา 121 นาที 39 วินาที

ในส่วนการทำนายผลหากเปรียบเทียบโครงข่ายที่ใช้เฉพาะ CNN แบบจำลอง DenseNet152 ใช้เวลาในการทำนายผลน้อยที่สุด และแบบจำลอง VGG19 ใช้เวลาในการทำนายผลมากที่สุด โดยมีเวลาที่ใช้อยู่ที่ 0.27 วินาทีต่อภาพ และ 0.8 วินาทีต่อภาพ ตามลำดับ และเมื่อเปรียบเทียบเฉพาะโครงข่าย CNN-RNN พบว่าแบบจำลอง DenseNet152-GRU ใช้เวลาในการทำนายผลน้อยที่สุด และแบบจำลอง ResNet152V2 ใช้เวลาในการทำนายผลมากที่สุด

จากข้อมูลข้างต้นแสดงให้เห็นว่าการใช้ชุดข้อมูลภาพรังสีทรวงอกที่มีการปรับปรุงด้วยวิธีการที่ต่างกันเป็นข้อมูลขาเข้าของโครงข่ายต่าง ๆ นั้นมีเวลาที่ใกล้เคียงกันไม่ได้แตกต่างกันอย่างมีนัยสำคัญ และ การใช้แบบจำลอง CNN-RNN ในการจำแนกภาวะโรคทรวงอกจากภาพถ่ายรังสี ใช้เวลาน้อยกว่าการใช้ เฉพาะแบบจำลอง CNN ทั้งในส่วนของเรียนรู้ เนื่องจากพารามิเตอร์ของโครงข่าย CNN-RNN มีค่าน้อย กว่าโครงข่าย CNN ด้วยเหตุผลว่ามีชั้น Fully connected layer ที่มีการเชื่อมโยงกันทุกโหนดมากกว่า

บทที่ 5

สรุปผลการวิจัยและการต่อยอดในอนาคต

5.1 สรุปผลการวิจัย

ในงานวิจัยนี้เสนอโครงข่ายแบบผสม CNN-RNN เพื่อจำแนกภาวะโรคทรวงอก 3 ประเภท ได้แก่ โรคโควิด-19 โรคปอดอักเสบ และภาวะปกติ จากทั้งภาพเอกซเรย์ และภาพ CT ทรวงอก ที่มีจำนวนทั้งหมด 16,210 ภาพจากแหล่งข้อมูลสาธารณะ โดยแบ่ง 80% ของภาพทั้งหมดใช้สำหรับการเรียนรู้ของแบบจำลอง และอีก 20% ที่เหลือใช้สำหรับการทดสอบ ในส่วนของข้อมูลสำหรับการตรวจสอบระหว่างการเรียนรู้ ได้จากการนำเอา 20% ของภาพในชุดสำหรับการเรียนรู้มาใช้ อีกทั้งมีการนำเทคนิคการปรับปรุงคุณภาพสีทั้งหมด 3 แบบ อันประกอบไปด้วย Linear contrast stretch Gamma correction และ CLAHE มาประยุกต์ใช้เพื่อเพิ่มความเด่นชัดของรายละเอียดภายในภาพ รวมถึงลดสัญญาณรบกวนภายในภาพก่อนนำไปใช้ในการเรียนรู้ และทำนายผลการจำแนกภาวะโรคทรวงอกของโครงข่ายเพื่อใช้ร่วมเปรียบเทียบค่าประสิทธิภาพการจำแนก โดยในงานวิจัยนี้เลือกนำแบบจำลอง CNN ทั้งหมด 3 แบบจำลอง ได้แก่ ResNet152V2 DenseNet121 และ VGG19 ที่ผ่านการเรียนรู้ด้วยชุดข้อมูล ImageNet แล้ว และโครงข่ายแบบจำลองประเภท RNN ได้แก่ LSTM และ GRU มาใช้ในการศึกษา นอกจากนี้ยังได้นำเฉพาะแบบจำลอง CNN ทั้ง 3 โครงข่าย ที่ไม่ถูกรวมเข้ากับแบบจำลอง RNN มาใช้ในการทดลองเพื่อเปรียบเทียบประสิทธิภาพอีกด้วย

จากผลลัพธ์ที่ได้แสดงให้เห็นว่าการใช้โครงข่ายแบบจำลอง CNN-RNN ให้ผลลัพธ์การจำแนกภาวะโรคทรวงอกที่มีประสิทธิภาพความแม่นยำโดยรวมสูงกว่าการใช้เฉพาะแบบจำลอง CNN อีกทั้งยังใช้เวลาในการประมวลผลที่น้อยกว่า โดยโครงข่าย ResNet152V2-GRU ที่ใช้ภาพต้นฉบับในการประมวลผลนั้นให้ประสิทธิภาพการจำแนกโดยรวมสูงที่สุดเมื่อเทียบกับโครงข่ายอื่น ๆ โดยมีค่าความแม่นยำ อยู่ที่ 93.37% ค่า precision 93.73% ค่า recall 93.44% และ ค่า F1-score เท่ากับ 93.54% นอกจากนี้ยังให้ผลลัพธ์ค่าความแม่นยำสูงที่สุดในการจำแนกภาวะโรค COVID-19 และภาวะปกติอีกด้วย โดยมีค่าความแม่นยำ คือ 94.14% และ 93.65% ตามลำดับ โดยใช้เวลาในการเรียนรู้ข้อมูล 99 นาที 57 วินาที และทำนายผล 0.21 วินาทีต่อภาพ แต่ในแง่การจำแนกภาวะโรคปอดอักเสบ VGG19-LSTM ที่ใช้ภาพที่ผ่านกระบวนการปรับปรุงคุณภาพสีด้วยเทคนิค Linear contrast stretch ให้ผลลัพธ์ประสิทธิภาพสูงที่สุด โดยมีค่าความแม่นยำเท่ากับ 99.35%

โครงข่ายแบบจำลอง CNN-RNN ที่นำเสนอสามารถนำไปใช้งานได้ไม่เพียงแต่ในการตรวจหาโรคโควิด-19 จากภาพถ่ายรังสีเท่านั้น แต่ยังสามารถใช้ในการวิเคราะห์ และวินิจฉัยโรคปอดอักเสบจาก

ภาพถ่ายรังสี 2 ชนิด คือ ภาพเอกซเรย์ และภาพ CT ได้อีกด้วย ซึ่งแสดงให้เห็นว่าวิธีการที่นำเสนอมีประสิทธิภาพสูงในการจำแนกโรคทรวงอก 3 ประเภท อีกทั้งยังมีความยืดหยุ่นต่อข้อมูลที่นำเข้ามาเนื่องจากสามารถประมวลผลได้ในภาพถ่ายรังสีทั้ง 2 ประเภท

5.2 การพัฒนาต่อยอดในอนาคต

ในงานวิจัยนี้ เราได้มุ่งเน้นการจำแนกภาวะโรคทรวงอกเป็นสามหมวดหมู่ ได้แก่ ภาวะโรคโควิด-19, ภาวะโรคปอดอักเสบ และภาวะปกติ จากภาพเอกซเรย์ และภาพ CT ด้วยโครงข่าย CNN-RNN โดยในโครงข่ายประเภท RNN ได้หยิบยืม LSTM และ GRU มาประยุกต์ใช้งาน แม้ผลลัพธ์ที่ได้จะให้ประสิทธิภาพสูง แต่ยังมีแนวทางอื่นๆ ที่สามารถวิจัยต่อยอดเพื่อพัฒนาประสิทธิภาพ และการประยุกต์ใช้งานของแบบจำลองให้ดียิ่งขึ้นในอนาคต ดังนี้

1.) พัฒนาโครงข่ายให้สามารถจำแนกหมวดหมู่ภาวะของโรคปอดจากภาพเอกซเรย์ และภาพ CT ได้มากขึ้นจากปัจจุบันสามารถจำแนกได้สามหมวดหมู่ เช่น วัณโรคปอด หลอดลมอักเสบ และมะเร็งปอด และต่อยอดให้จำแนกแบบหลายหมวดหมู่ในภาพเดียวได้ เพื่อใช้จัดการกับกรณีผู้ป่วยอาจมีมากกว่าหนึ่งภาวะพร้อมกัน

2.) เลื่อนำสถาปัตยกรรม RNN ขั้นสูง ได้แก่ การใช้ Bidirectional Gated Recurrent Unit (Bi-GRU) และ Bidirectional Long Short-Term Memory (Bi-LSTM) มาประยุกต์ใช้ในโครงข่าย CNN-RNN เพื่อหาความสัมพันธ์ของคุณลักษณะ และจำแนกหมวดหมู่ภาวะโรคทรวงอก พร้อมเปรียบเทียบประสิทธิภาพผลลัพธ์การจำแนกกับโครงข่ายเดิมที่พัฒนาขึ้นในงานวิจัยนี้ เพื่อหาโครงข่ายที่ให้ผลลัพธ์การจำแนกหมวดหมู่ภาวะโรคทรวงอกดีที่สุด

3.) เลื่อนำเทคนิคฟัซซี่ลอจิก (Fuzzy Logic) มาประยุกต์ใช้เป็นฟังก์ชันการจำแนกหมวดหมู่ภาวะโรคทรวงอก (Classifier) เพื่อเปรียบเทียบ และวิเคราะห์ว่าส่งผลกระทบต่อประสิทธิภาพในการจำแนกของโครงข่าย CNN-RNN หรือไม่ และช่วยเพิ่มประสิทธิภาพการจำแนกของโครงข่ายที่พัฒนาในงานวิจัยนี้ที่ใช้ฟังก์ชันการจำแนกเป็น Softmax

บรรณานุกรม

- [1] WHO. Director-General’s Opening Remarks at the Media Briefing on COVID-19–11 March 2020. Available online: <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>.
- [2] Clemente-Suárez VJ, Navarro-Jiménez E, Moreno-Luna L, Saavedra-Serrano MC, Jimenez M, Simón JA, Tornero-Aguilera JF. The Impact of the COVID-19 Pandemic on Social, Health, and Economy. *Sustainability*. 2021; 13(11):6314. <https://doi.org/10.3390/su13116314>.
- [3] Centers for Disease Control and Prevention. “**Human Coronavirus Types.**” [Online]. Available : <https://www.cdc.gov/coronavirus/types.html>. 2024.
- [4] กฤษฏี ประภาสวัตต์ วลัยลักษณ์ ชัยสูตร อภิญญา เจริญศักดิ์ และ คณะบรรณาธิการ 2546. **รังสีวินิจฉัย Diagnostic radiology**. กรุงเทพฯ : มหาวิทยาลัยมหิดล.
- [5] พรรณี แสงแก้ว มนัสวี เลาะวิธี และ ทวีป แสงแห่งธรรม. “การถ่ายภาพเอกซเรย์คอมพิวเตอร์เพื่อการวินิจฉัยโรคระบาดโควิด-19” *Thai Journal of Physics*, ฉบับที่ 37(2), 2563. หน้า 45-59
- [6] Zeng F, Cai Y, Guo Y, et al. The diagnostic value of chest X-ray in coronavirus disease 2019: A comparative study of X-ray and CT. *Science Progress*. 2021;104(3).
- [7] ศรสุมภา ลี้มเจริญ (พ.บ.) และ จิตราพร อินทรารักษ์ (พ.บ.) “วินิจฉัยทางรังสีวิทยาในโควิด-19” *BJM*, ฉบับที่ 7(1), 2020
- [8] Gonzalez, R. C. 2002. **Digital Image Processing (2 nd Edition)**. Prentice Hall
- [9] รศ. ดร. สนั่น ศรีสุข 2564. **การประมวลผลภาพขั้นสูง (Advanced Image Processing)**. มหาวิทยาลัยนครพนม
- [10] The MathWorks, Inc. “**Binary Images.**” [Online]. Available : <http://matlab.izmiran.ru/help/toolbox/images/intro4.html>. 2005.
- [11] Sharma M. “**Histograms in Image Processing with skimage-Python.**” [Online]. Available : <https://towardsdatascience.com/histograms-in-image-processing-withskimage-python-be5938962935>. 2019.
- [12] DaSilva V. “**Computer Vision for Busy Developers.**” [Online]. Available : https://medium.com/@vad7_1_0/cv-for-busy-developers-image-processing-d93ed4836880. 2019.

บรรณานุกรม (ต่อ)

- [13] ไอคูร์ย กายจนสุรัตน์. “การระบุบุคคลโดยใช้สามเหลี่ยมสีลอเรนซ์และรูปแบบเส้นเลือด ณ ตำแหน่งขั้วประสาทของจอประสาทตา.” วิทยานิพนธ์วิศวกรรมศาสตรดุษฎีบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้า บัณฑิตวิทยาลัย, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง. 2563.
- [14] Robert F., Simon P., Ashley W., Erik W. “**Image Arithmetic**” [Online].
Available : <https://homepages.inf.ed.ac.uk/rbf/HIPR2/arthops.htm>. 2004.
- [15] Patterson, J. and Gibson, A. 2017. Deep learning: A practitioner’s approach. O’Reilly Media.
- [16] ดวงเพ็ญ เจตน์พิพัฒน์พงษ์ และ รัชกฤษ ธนพัฒน์ดล. “การศึกษาเปรียบเทียบผลการเปลี่ยนแปลงจำนวนคำตอบของโครงข่ายประสาทเทียมสำหรับการพยากรณ์ระดับแม่น้ำยม จังหวัดแพร่” **วารสารการจัดการสิ่งแวดล้อม**, ปีที่ 6, ฉบับที่ 2, 2553. หน้า 35-53
- [17] ธนาวุฒิ ประกอบผล. “โครงข่ายประสาทเทียม” **วารสาร มฉก.วิชาการ**, ปีที่ 12, ฉบับที่ 24, 2552. หน้า 73-87
- [18] Luthfi, R. 2019 **Neural Network: The Dead Neuron**. [Online].
Available:<https://towardsdatascience.com/neural-network-the-dead-neuron-eaa92e575748>
- [19] Patterson, J. and Gibson, A. 2017. Deep learning: A practitioner’s approach. O’Reilly Media.
- [20] Goodfellow, I. Bengio, Y. and Courville, A. 2016. Deep Learning. London, England: MIT Press.
- [21] Alzubaidi, L. Al-Shamma, O. Fadhel, M. A. Farhan, L. Zhang, J. and Duan, Y. “Optimizing the performance of breast cancer classification by employing the same domain transfer learning from hybrid deep convolutional neural network model.” *Electronics*, 2020, 9(3), 445.
- [22] ศุภณัฐ จินตวัฒน์สกุล “การจำแนกภาพขวดแบบเซตเปิดด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชัน” วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิชา สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ, 2562
- [23] Szegedy, C. Ioffe, S. Vanhoucke, V. and Alemi, A. “Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning”, in AAAI, 2017, pp. 4278–4284
- [24] Sandler, M. Howard, A. Zhu, M. Zhmoginov, A. and Chen, L.-C. “MobileNetV2: Inverted residuals and linear bottlenecks,” in 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018.

บรรณานุกรม (ต่อ)

- [25] He, K. Zhang, X. Ren, S. and Sun, J. “Deep residual learning for image recognition,” in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016. [9] Simonyan, K. and Zisserman, A. “Very deep convolutional networks for large-scale image recognition,” arXiv [cs.CV], 2014.
- [26] Simonyan, K. and Zisserman, A. (2015) Very Deep Convolutional Networks for Large-Scale Image Recognition. The 3rd International Conference on Learning Representations (ICLR2015).
- [27] Khattar, Anuradha, and S. M. K. Quadri. “Generalization of Convolutional Network to Domain Adaptation Network for Classification of Disaster Images on Twitter.” Multimedia Tools and Applications, vol. 81, no. 21, 2022, pp. 30437–30464
- [28] He, K., Zhang, X., Ren, S., & Sun, J. 2016. Identity Mappings in Deep Residual Networks. ArXiv, abs/1603.05027.
- [29] Malik, H., Naeem, A., Sadeghi-Niaraki, A. et al. Multi-classification deep learning models for detection of ulcerative colitis, polyps, and dyed-lifted polyps using wireless capsule endoscopy images. *Complex Intell. Syst.* 10, 2024, 2477–2497.
- [30] Huang, G., Liu, Z., Van Der Maaten, L. and Weinberger, K.Q. 2017. Densely Connected Convolutional Networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, 21-26 July 2017, 4700-4708.
- [31] Radwan, Noha. Leveraging Sparse and Dense Features for Reliable State Estimation in Urban Environments, 2019.
- [32] Sepp Hochreiter; Jürgen Schmidhuber. 1997. "Long short-term memory". *Neural Computation.* 9 (8): 1735–1780. doi:10.1162/neco.1997.9.8.1735. PMID 9377276. S2CID 1915014. [20] L. C. Jain and L. R. Medsker. 1999. *Recurrent Neural Networks: Design and Applications* (1st. ed.). CRC Press, Inc., USA.
- [33] Awad, Asmaa Ahmed, et al. “An Improved Long Short Term Memory Network for Intrusion Detection.” *PloS One*, vol. 18, no. 8, 2023, p. e0284795.

บรรณานุกรม (ต่อ)

- [34] Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H. & Bengio, Y. 2014. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation (cite arxiv:1406.1078Comment: EMNLP 2014)
- [35] Subrmanian, M., Shanmugavadivel, K., Nandhini, P.S., Sowmya, R. “Evaluating the Performance of LSTM and GRU in Detection of Distributed Denial of Service Attacks Using CICDDoS2019 Dataset.” Proceedings of 7th International Conference on Harmony Search, Soft Computing and Applications, Springer Nature Singapore, 2022, pp. 395–406.
- [36] Zhang, Jianpeng & Xie, Yutong & Li, Yi & Shen, Chunhua & Xia, Yong. 2020. COVID-19 Screening on Chest X-ray Images Using Deep Learning based Anomaly Detection. arXiv preprint arXiv:2003.12338.
- [37] Ezz El-Din Hemdan, Marwa A. Shouman, and Mohamed Esmail Karar. 2020. Covidx-net: a framework of deep learning classifiers to diagnose covid-19 in x-ray images. arXiv preprint arXiv:2003.11055.
- [38] Islam, M. Z., Islam, M. M., & Asraf, A. 2020. A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images. Informatics in medicine unlocked, 20, 100412.
- [39] Rahman, T., Khandakar, A., Qiblawey, Y., Tahir, A., Kiranyaz, S., Abul Kashem, S. B., Islam, M. T., Al Maadeed, S., Zughair, S. M., Khan, M. S., & Chowdhury, M. 2021. Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images. Computers in biology and medicine, 132, 10.
- [40] Wu, X., Hui, H., Niu, M., Li, L., Wang, L., He, B., Yang, X., Li, L., Li, H., Tian, J., & Zha, Y. 2020. Deep learning-based multi-view fusion model for screening 2019 novel coronavirus pneumonia: A multicentre study. European journal of radiology, 128, 109041.
- [41] Xu X, Jiang X, Ma C, Du P, Li X, Lv S, Yu L, Ni Q, Chen Y, Su J, Lang G, Li Y, Zhao H, Liu J, Xu K, Ruan L, Sheng J, Qiu Y, Wu W, Liang T, Li L). 2020. A Deep Learning System to Screen Novel Coronavirus Disease 2019 Pneumonia. Engineering (Beijing). Oct;6(10):1122-1129. doi: 10.1016/j.eng.2020.04.010.

บรรณานุกรม (ต่อ)

- [42] Perumal, V., Narayanan, V. & Rajasekar, S.J.S. 2021. Detection of COVID-19 using CXR and CT images using Transfer Learning and Haralick features. *Appl Intell* 51, 341–358.
- [43] Cohen, J.P.; Morrison, P.; Dao, L.; Roth, K.; Duong, T.Q; Ghassemi, M. Covid-19 image data collection: Prospective predictions are 287 the future. *arXiv preprint arXiv:2006.2020*, 11988, <https://github.com/ieee8023/covid-chestxray-dataset>.
- [44] Chowdhury, M.E.; Rahman, T.; Khandakar, A.; Mazhar, R.; Kadir, M.A.; Mahbub, Z.B.; Islam, K.R.; Khan, M.S.; Iqbal, A.; 289 Al-Emadi, N.A.; Reaz, M.B. Can AI Help in Screening Viral and COVID-19 Pneumonia?. *IEEE Access* 2020, 8, 132665-132676. 290
- [45] Kang, Z.; Xiaohong, L.; Jun, S.; Zhihuan, L.; Ye, S.; Xingwang, W.; Yunfei, Z.; Wenhua, L.; Chengdi, W.; Ke, W.; Linsen, Y.; 292 Ming, G.; Zhongguo, Z.; Liang, L.; Jin, W.; Zehong, Y.; Huimin, C.; Jie, X.; Lei, Y.; Wenjia, C.; Wenqin, X.; Shaoxu, W.; Wei, Z.; 293 Shanping, J.; Lianghong, Z.; Xuan, Z.; Li, W.; Liu, L.; Jiaming, L.; Haiping, Y.; Winston, W.; Oulan, L.; Charlotte, Z.; Liang, L.; 294 Tao, W.; Ruiyun, D.; Kang, W.; Yong, Z.; Ting, C.; Johnson, Y.N.L.; Manson, F.; Jianxing, H.; Tianxin, L.; Weimin, L.; Guangyu, W. 295 Clinically Applicable AI System for Accurate Diagnosis, Quantitative Measurements, and Prognosis of COVID-19 Pneumonia 296 Using Computed Tomography. *Cell* 2020, 181(6), 1423-1433.e11.
- [46] Kermamy, D.S.; Goldbaum, M.; Cai, W.; Valentim, C.; Liang, H.; Baxter, S. L.; McKeown, A.; Yang, G.; Wu, X.; Yan, F.; Dong, J.; 298 Prasadha, M. K.; Pei, J.; Ting, M.; Zhu, J.; Li, C.; Hewett, S.; Dong, J.; Ziyar, I.; Shi, A.; Zhang, K. Identifying Medical Diagnoses and 299 Treatable Diseases by Image-Based Deep Learning. *Cell* 2018, 172(5), 1122–1131.e9.
- [47] Cleverley J., Piper J., Jones M.M. The role of chest radiography in confirming covid-19 pneumonia. *BMJ*. 2020:370. doi: 10.1136/bmj.m2426.
- [48] Islam R, Tarique M. Chest X-Ray Images to Differentiate COVID-19 from Pneumonia with Artificial Intelligence Techniques. *Int J Biomed Imaging*. 2022 Dec 22;2022:5318447. doi: 10.1155/2022/5318447. PMID: 36588667; PMCID: PMC9800093.
- [49] Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace.
- [50] F. Chollet and others, Keras. GitHub. from <https://github.com/keras-team/keras>.

บรรณานุกรม (ต่อ)

- [51] T. Shermin, S. W. Teng, M. Murshed, G. Lu, F. Sohel, and M. Paul, "Enhanced transfer learning with ImageNet trained classification layer," in *Image and Video Technology*, Cham: Springer International Publishing, 2019, pp. 142–155.
- [52] J. Wang, Y. Yang, J. Mao, Z. Huang, C. Huang and W. Xu, "CNN-RNN: A Unified Framework for Multi-label Image Classification," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, pp. 2285-2294, doi: 10.1109/CVPR.2016.251.
- [53] Shorten, C., Khoshgoftaar, T.M. A survey on Image Data Augmentation for Deep Learning. *J Big Data* 6, 60 (2019).
- [54] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin et al. "TensorFlow: LargeScale Machine Learning on Heterogeneous Distributed Systems," 2015, arXiv preprint arXiv:1603.04467.
- [55] E. Bisong, *Building Machine Learning and Deep Learning Models on Google Cloud Platform*, Springer, 2019.
- [56] Ozturk, T., Talo, M., Yildirim, E. A., Baloglu, U. B., Yildirim, O., & Rajendra Acharya, U. 2020. Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Computers in biology and medicine*, 121, 103792.
- [57] K. El Asnaoui, Y. Chawki. Using X-ray images and deep learning for automated detection of coronavirus disease, *J. Biomol. Struct. Dyn.*, 2020, 0 1–12.
- [58] M. Rahimzadeh and A. Attar, "A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2," *Inform. Med. Unlocked*, vol. 19, no. 100360, p. 100360, 2020.
- [59] Joshi, A.M.; Nayak, D.R.; Das, D.; Zhang, Y.D. LiMS-Net: A Lightweight Multi-Scale CNN for COVID-19 Detection from Chest CT Scans. *ACM Trans. Manag. Inf. Syst.* 2022, 14, 1–17.