



การเรียนรู้เชิงลึกด้วยโครงข่ายเส้นประสาทเทียมในการจำแนกการเป็นโรคมะเร็งตับโดยใช้
ภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า

DEEP LEARNING WITH CONVOLUTIONAL NEURAL NETWORK FOR CLASSIFICATION
OF HEPATOCELLULAR CARCINOMA ON MAGNETIC RESONANCE IMAGING

กานต์พิชชา พรหมมา

บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ

2566

การเรียนรู้เชิงลึกด้วยโครงข่ายเส้นประสาทเทียมในการจำแนกการเป็นโรคมะเร็งตับโดยใช้
ภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า



กานต์พิชชา พรหมมา

สารนิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
วิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการข้อมูล
คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
ปีการศึกษา 2566
ลิขสิทธิ์ของมหาวิทยาลัยศรีนครินทรวิโรฒ

DEEP LEARNING WITH CONVOLUTIONAL NEURAL NETWORK FOR CLASSIFICATION
OF HEPATOCELLULAR CARCINOMA ON MAGNETIC RESONANCE IMAGING



A Master's Project Submitted in Partial Fulfillment of the Requirements
for the Degree of MASTER OF SCIENCE
(Data Science)

Faculty of Science, Srinakharinwirot University

2023

Copyright of Srinakharinwirot University

สารนิพนธ์
เรื่อง
การเรียนรู้เชิงลึกด้วยโครงข่ายเส้นประสาทเทียมในการจำแนกการเป็นโรคมะเร็งตับโดยใช้ภาพถ่ายการตรวจ
คลื่นแม่เหล็กไฟฟ้า
ของ
กานต์พิชชา พรหมมา

ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการข้อมูล
ของมหาวิทยาลัยศรีนครินทรวิโรฒ

(รองศาสตราจารย์ นายแพทย์ฉัตรชัย เอกปัญญาสกุล)
คณบดีบัณฑิตวิทยาลัย

คณะกรรมการสอบปากเปล่าสารนิพนธ์

..... ที่ปรึกษาหลัก (อาจารย์ ดร.โสภณ มงคลลักษณ์)	 ประธาน (รองศาสตราจารย์ ดร.อรรณพ หมั่นสกุล)
..... ที่ปรึกษาร่วม (รองศาสตราจารย์ ดร.วรพันธ์ คู่สกุลนิรันดร์)	 กรรมการ (ผู้ช่วยศาสตราจารย์ ดร.ศิริสรรพ เหล่าหะเกียรติ)

ชื่อเรื่อง	การเรียนรู้เชิงลึกด้วยโครงข่ายประสาทเทียมในการจำแนกการเป็นโรคมะเร็ง ตับโดยใช้ภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า
ผู้วิจัย	กานต์พิชชา พรหมมา
ปริญญา	วิทยาศาสตร์มหาบัณฑิต
ปีการศึกษา	2566
อาจารย์ที่ปรึกษา	อาจารย์ ดร. โสภณ มงคลลักษณ์
อาจารย์ที่ปรึกษาร่วม	รองศาสตราจารย์ ดร. วรพันธ์ คู่สกุลนิรันดร์

โรคมะเร็งตับเป็นหนึ่งในสาเหตุการเสียชีวิตอันดับต้น ๆ ของคนทั่วโลก และเป็นมะเร็งเพียงชนิดเดียวที่สามารถวินิจฉัยโดยใช้ภาพถ่ายทางรังสีวินิจฉัยได้ ไม่จำเป็นต้องใช้การตรวจแบบตัดชิ้นเนื้อ (Biopsy) ซึ่งเทคนิคการตรวจที่ดีที่สุด คือ การตรวจคลื่นแม่เหล็กไฟฟ้า (Magnetic Resonance Imaging) เพราะสามารถเห็นความแตกต่างของเนื้อเยื่อได้ดี แต่ทั้งนี้ก็ต้องอาศัยรังสีแพทย์ในการวินิจฉัยโรค เนื่องจากแพทย์มีจำนวนน้อยจึงส่งผลให้เกิดความล่าช้าในการรักษา งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อสร้างแบบจำลองการจำแนกการเป็นโรคมะเร็งตับและไม่เป็นโรคมะเร็งตับ (Binary Classification) จากภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้าโดยใช้แบบจำลองโครงข่ายประสาทเทียมคอนโวลูชัน VGG16 และ ResNet-50 โดยมีกระบวนการปรับโครงสร้างแบบจำลองและค่าพารามิเตอร์ต่าง ๆ เป็นทั้งหมด 4 ชุดการทดลอง (Version) ซึ่งผู้วิจัยใช้ข้อมูลจากภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้าของฐานข้อมูล The Cancer Imaging Archive (TCIA) จำนวน 2,251 รูป โดยแบบจำลอง ResNet-50 มีประสิทธิภาพมากกว่าแบบ VGG16 ในทุกผลการทดลองทั้งหมด 4 ชุด และแบบจำลอง ResNet-50 Version 2 มีค่า Accuracy มากที่สุด โดยมีค่าเท่ากับ 0.69 หากวิเคราะห์ด้วยเมตริกซ์ False Negatives (FN) แบบจำลอง VGG16 Version 2 มีค่า False Negatives น้อยกว่า ResNet-50 Version 2 โดยมีรูปที่ถูกทายเป็น False Negatives อยู่ที่ 32 รูป และ 54 รูป ตามลำดับ จากผลลัพธ์นี้แสดงให้เห็นว่าการเรียนรู้เชิงลึกด้วยโครงข่ายประสาทเทียมคอนโวลูชันสามารถจำแนกการเป็นโรคมะเร็งตับได้ และการเลือกประเภทของแบบจำลองส่งผลกระทบอย่างมีนัยสำคัญต่อประสิทธิภาพของแบบจำลอง ในการศึกษาครั้งนี้ โครงสร้างโครงข่ายประสาทเทียมคอนโวลูชันแบบ ResNet-50 มีประสิทธิภาพมากกว่าแบบ VGG16 เมื่อวัดด้วยเมตริกซ์ Accuracy แต่หากเปรียบเทียบด้วยเมตริกซ์ False Negatives โครงสร้างโครงข่ายประสาทเทียมคอนโวลูชันแบบ VGG16 มีประสิทธิภาพดีกว่าแบบ ResNet-50

คำสำคัญ : โรคมะเร็งตับ, การตรวจคลื่นแม่เหล็กไฟฟ้า, การจำแนกประเภท, การเรียนรู้เชิงลึก, โครงข่ายประสาทเทียมคอนโวลูชัน

Title	DEEP LEARNING WITH CONVOLUTIONAL NEURAL NETWORK FOR CLASSIFICATION OF HEPATOCELLULAR CARCINOMA ON MAGNETIC RESONANCE IMAGING
Author	KANPHITCHA PROMMA
Degree	MASTER OF SCIENCE
Academic Year	2023
Thesis Advisor	Dr. Sophon Mongkolluksamee
Co Advisor	Associate Professor Dr. Worapan Kusakunniran

Hepatocellular carcinoma is one of the leading causes of death globally. It is the only type of cancer that can be diagnosed using radiological imaging, without the need for a biopsy. The best diagnostic technique is magnetic resonance imaging (MRI), as it can visualize tissue differences effectively. However, it still relies on radiologists for diagnosis, which can cause delays in treatment due to their limited numbers. This research aims to develop a model for binary classification of hepatocellular carcinoma and non-hepatocellular carcinoma from magnetic resonance imaging (MRI) using deep learning techniques with the convolutional neural network (CNN) architectures VGG16 and ResNet-50. The four experimental setups were conducted with variations in model architectures and parameters. The researchers utilized data from MRI images from The Cancer Imaging Archive (TCIA) database, comprising 2,251 images. The ResNet-50 model outperformed VGG16 in all four experimental setups, with ResNet-50 Version 2 achieving the highest accuracy of 0.69. Analyzing False Negatives (FN) matrices revealed that VGG16 Version 2 had fewer false negatives than ResNet-50 Version 2, with 32 and 54 misclassified images, respectively. These results indicate that deep learning with convolutional neural networks can effectively classify hepatocellular carcinoma. Furthermore, the choice of model type significantly impacts model performance. In this study, the ResNet-50 CNN architecture outperformed VGG16 in terms of accuracy, but when considering False Negatives, VGG16 demonstrated a better performance than ResNet-50.

Keyword : Hepatocellular carcinoma, Magnetic resonance imaging, Binary Classification, Deep learning, Convolutional neural network

กิตติกรรมประกาศ

สารนิพนธ์นี้สามารถสำเร็จลุล่วงได้ด้วยดี ด้วยการสนับสนุนอย่างดียิ่งในทุก ๆ ด้าน ทั้งในด้านการให้ความรู้ แนวทางการดำเนินการต่างๆ ในการจัดทำสารนิพนธ์ ข้อคำแนะนำการปรับปรุงข้อบกพร่องต่าง ๆ ที่เป็นประโยชน์จากความคิดเห็นของคณะกรรมการผู้ควบคุมสารนิพนธ์ ด้วยความเมตตากรุณาช่วยเหลือและเอาใจใส่อย่างดี ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างยิ่ง

ขอกราบขอบพระคุณ อาจารย์ ดร.โสภณ มงคลลักษณ์ และรองศาสตราจารย์ ดร.วรพันธ์ คูสฤกษ์นิรันดร์ ที่ได้ให้ความเมตตากรุณาเป็นที่ปรึกษา และให้ความช่วยเหลือชี้แนะแนวทางการดำเนินงานต่างๆ ที่เป็นประโยชน์ทั้งในด้านการศึกษาและการจัดทำสารนิพนธ์นี้ด้วยความเอาใจใส่ตลอดมา ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอกราบขอบพระคุณอาจารย์และกรรมการบริหารหลักสูตรสาขาวิชาวิทยาการข้อมูล คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒทุกท่าน ที่ได้กรุณาประสิทธิ์ประสาทความรู้ต่างๆ ให้กับผู้วิจัย ตลอดจนความช่วยเหลือในการจัดทำสารนิพนธ์ครั้งนี้

ขอขอบคุณ นพ.กฤตภาส ชัยศุจยากร และดร.วัชรภรณ์ ต้นโชติศรีนนท์ ที่ให้คำแนะนำและความช่วยเหลืออย่างดียิ่งสำหรับการเตรียมข้อมูลตัวอย่างที่ใช้ในการจัดทำสารนิพนธ์ครั้งนี้

ขอขอบคุณพี่ ๆ และเพื่อน ๆ สาขาวิชาวิทยาการข้อมูล รวมถึงบุคคลอื่น ๆ อีกหลายท่านที่ไม่ได้กล่าวนามไว้ ณ ที่นี้ ที่ได้ให้ความช่วยเหลือและให้กำลังใจกับผู้วิจัยตลอดมา

สุดท้ายนี้ ผู้วิจัยขอโน้มรำลึกถึงคุณบิดามารดา และครูบาอาจารย์ที่อบรมสั่งสอนให้ความรู้ รวมถึงเป็นอีกหนึ่งกำลังใจให้การสนับสนุนกับผู้วิจัยด้วยดีตลอดมา

กานต์พิชชา พรพมา

สารบัญ

หน้า

บทคัดย่อภาษาไทย	ง
บทคัดย่อภาษาอังกฤษ	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ	ช
สารบัญตาราง.....	ณ
สารบัญรูปภาพ	ญ
บทที่ 1 บทนำ.....	1
1.1 ความสำคัญและความเป็นมาของงานวิจัย.....	1
1.2 วัตถุประสงค์ของการวิจัย.....	3
1.3 ความสำคัญของการวิจัย	3
1.4 ขอบเขตของการวิจัย	4
1.5 กรอบแนวคิดในการวิจัย	5
1.6 ประโยชน์ที่คาดว่าจะได้รับจากงานวิจัย	6
บทที่ 2 วรรณกรรมและงานวิจัยที่เกี่ยวข้อง.....	7
2.1 การตรวจคลื่นแม่เหล็กไฟฟ้า (Magnetic Resonance Imaging: MRI)	7
2.2 ทฤษฎีเกี่ยวกับกระบวนการก่อนการประมวลผลภาพ (Image Pre-processing).....	9
2.3 ทฤษฎีเกี่ยวกับการเพิ่มปริมาณข้อมูลภาพ (Image Augmentation).....	11
2.4 ทฤษฎีเกี่ยวกับการจัดการความไม่สมดุลกันของข้อมูล (Imbalance Dataset)	11
2.5 ทฤษฎีเกี่ยวกับอัลกอริทึมการเรียนรู้เชิงลึกด้วยโครงข่ายประสาทเทียมคอนโวลูชัน (Deep Learning with Convolutional Neural Network)	13
2.6 การวัดประสิทธิภาพของแบบจำลอง (Model Evaluation)	20

2.7 งานวิจัยที่เกี่ยวข้อง (Related work).....	24
บทที่ 3 วิธีการดำเนินการวิจัย.....	31
3.1 กระบวนการทำงานของแบบจำลอง (Workflows Process of Model)	31
3.2 การเก็บรวบรวมข้อมูล (Data Acquisition).....	33
3.3 การระบุประเภทข้อมูล (Data Labeling)	34
3.4 การทำความสะอาดข้อมูล (Data Cleansing).....	34
3.5 การเตรียมข้อมูล (Data Preprocessing)	35
3.6 การสำรวจข้อมูล (Exploratory Data Analysis: EDA).....	35
3.7 การแบ่งข้อมูลสำหรับชุดฝึกสอนและชุดทดสอบ (Train-Test Split)	37
3.8 การสร้างแบบจำลองการจำแนก (Classification Model)	38
3.9 การวัดประสิทธิภาพของแบบจำลอง (Model Evaluation)	45
บทที่ 4 ผลการดำเนินงานวิจัย.....	46
4.1 ผลการประเมินประสิทธิภาพแบบจำลองด้วยอัลกอริทึม VGG16	46
4.2 ผลการประเมินประสิทธิภาพแบบจำลองด้วยอัลกอริทึม ResNet-50	68
บทที่ 5 สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ	93
5.1 สรุปผลการวิจัย.....	93
5.2 อภิปรายผล	95
5.3 ข้อเสนอแนะ.....	98
บรรณานุกรม	99
ประวัติผู้เขียน.....	102

สารบัญตาราง

หน้า

ตาราง 1 ตารางการปรับข้อมูลรูปภาพด้วยเทคนิคต่าง ๆ ก่อนนำไปใช้ในการสร้างแบบจำลอง ...	38
ตาราง 2 ตารางแสดงค่าพารามิเตอร์ต่าง ๆ ที่ใช้ในการสร้างแบบจำลอง VGG16 Version 1	39
ตาราง 3 ตารางแสดงค่าพารามิเตอร์ต่าง ๆ ที่ใช้ในการสร้างแบบจำลอง VGG16 Version 2	40
ตาราง 4 ตารางแสดงค่าพารามิเตอร์ต่าง ๆ ที่ใช้ในการสร้างแบบจำลอง VGG16 Version 3	41
ตาราง 5 ตารางแสดงค่าพารามิเตอร์ต่าง ๆ ที่ใช้ในการสร้างแบบจำลอง VGG16 Version 4	41
ตาราง 6 ตารางแสดงค่าพารามิเตอร์ต่าง ๆ ที่ใช้ในการสร้างแบบจำลอง ResNet50 Version 1 .	42
ตาราง 7 ตารางแสดงค่าพารามิเตอร์ต่าง ๆ ที่ใช้ในการสร้างแบบจำลอง ResNet50 Version 2 .	43
ตาราง 8 ตารางแสดงค่าพารามิเตอร์ต่าง ๆ ที่ใช้ในการสร้างแบบจำลอง ResNet50 Version 3 .	44
ตาราง 9 ตารางแสดงค่าพารามิเตอร์ต่าง ๆ ที่ใช้ในการสร้างแบบจำลอง ResNet50 Version 4 .	44
ตาราง 10 ตารางแสดงประสิทธิภาพการจำแนกของแบบจำลอง VGG16 แต่ละ Version	46
ตาราง 11 ตารางเปรียบเทียบค่า Accuracy เฉลี่ยแต่ละ fold ข้อมูลของแบบจำลอง VGG16	67
ตาราง 12 ตารางแสดงประสิทธิภาพการจำแนกของแบบจำลอง ResNet-50 แต่ละ Version	68
ตาราง 13 ตารางเปรียบเทียบค่า Accuracy เฉลี่ยแต่ละ fold ข้อมูลของแบบจำลอง ResNet-50	89

สารบัญรูปภาพ

หน้า

ภาพประกอบ 1 ภาพตัดในแนวขวาง (Axial slide)	8
ภาพประกอบ 2 ในระยะ T1 มีภาพที่เกิดขึ้นเป็น 6 ระยะย่อย (six time points: TPs)	8
ภาพประกอบ 3 เทคนิค Largest Contour Segmentation (LCS) Procedure	9
ภาพประกอบ 4 (ก) ค่า Histogram ภาพต้นฉบับ, (ข) ค่า Histogram ที่ถูกตัด	10
ภาพประกอบ 5 เทคนิคการเพิ่มปริมาณข้อมูลภาพ (Image Augmentation)	11
ภาพประกอบ 6 เทคนิคการจัดการความไม่สมดุลกันของข้อมูล (Imbalanced Datasets)	13
ภาพประกอบ 7 ความแตกต่างระหว่างโครงข่ายประสาทเทียมและการเรียนรู้เชิงลึก	14
ภาพประกอบ 8 แผนภูมิสถาปัตยกรรม VGGNet	16
ภาพประกอบ 9 สถาปัตยกรรม VGGNet	16
ภาพประกอบ 10 สถาปัตยกรรม ResNet	18
ภาพประกอบ 11 การเลือกโครงข่ายประสาทเทียมคอนโวลูชันที่เหมาะสมกับชุดข้อมูล (Input data)	19
ภาพประกอบ 12 ตัวอย่างตาราง Confusion Matrix ขนาด 2x2	20
ภาพประกอบ 13 ตัวอย่างตาราง Confusion Matrix ขนาด 2x2	23
ภาพประกอบ 14 ROC curve แสดงความสัมพันธ์ระหว่างแกน X (false positive rate หรือ 1-specificity) และแกน Y (true positive rate หรือ sensitivity)	23
ภาพประกอบ 15 รูปภาพขั้นตอนการทำ K-FOLD CROSS –VALIDATION	24
ภาพประกอบ 16 แสดงขั้นตอนในการสร้างแบบจำลองของงานวิจัย Deep Learning with Convolutional Neural Network for Differentiation of Liver Masses at Dynamic Contrast-enhanced CT: A Preliminary Study	26
ภาพประกอบ 17 ข้อมูลแสดงประสิทธิภาพของแบบจำลอง (Model Evaluation) จากการใช้ข้อมูลภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI)	27

ภาพประกอบ 18 โครงสร้างเป็น Alexnet และใช้ Weight เป็น “Imagenet”	28
ภาพประกอบ 19 ผลลัพธ์การวัดประสิทธิภาพของแบบจำลอง โดยใช้ โครงสร้างเป็น Alexnet และใช้ Weight เป็น “Imagenet”	29
ภาพประกอบ 20 การวัดประสิทธิภาพของแบบจำลองด้วยโครงข่ายประสาทเทียม (CNN) ด้วย โครงสร้างเป็น VGG16 และใช้ Weight เป็น “Imagenet”	30
ภาพประกอบ 21 แสดง Flowchart กระบวนการสร้างแบบจำลอง	31
ภาพประกอบ 22 จำนวนชุดข้อมูลผู้ป่วยที่เป็นโรคมะเร็งตับ (Hepatocellular carcinoma : HCC) และผู้ป่วยที่ไม่เป็นโรคมะเร็งตับ (Non Hepatocellular carcinoma : Non HCC)	36
ภาพประกอบ 23 รายละเอียดผู้ป่วยที่เป็นโรคมะเร็งตับ (Hepatocellular carcinoma : HCC) และผู้ป่วยที่ไม่เป็นโรคมะเร็งตับ (Non Hepatocellular carcinoma : Non HCC)	36
ภาพประกอบ 24 การจัดกลุ่มผู้ป่วยแต่ละรายใหม่ในกลุ่มผู้ป่วยที่เป็นโรคมะเร็งตับ (Hepatocellular carcinoma : HCC)	37
ภาพประกอบ 25 การจัดกลุ่มผู้ป่วยแต่ละรายใหม่ในกลุ่มผู้ป่วยที่ไม่เป็นโรคมะเร็งตับ (Non Hepatocellular carcinoma : Non HCC)	38
ภาพประกอบ 26 Confusion Matrix แบบจำลอง VGG16 Version 1 fold 1	47
ภาพประกอบ 27 Learning Curve แบบจำลอง VGG16 Version 1 fold 1	47
ภาพประกอบ 28 Confusion Matrix แบบจำลอง VGG16 Version 1 fold 2	48
ภาพประกอบ 29 Learning Curve แบบจำลอง VGG16 Version 1 fold 2	48
ภาพประกอบ 30 Confusion Matrix แบบจำลอง VGG16 Version 1 fold 3	49
ภาพประกอบ 31 Learning Curve แบบจำลอง VGG16 Version 1 fold 3	49
ภาพประกอบ 32 Confusion Matrix แบบจำลอง VGG16 Version 1 fold 4	50
ภาพประกอบ 33 Learning Curve แบบจำลอง VGG16 Version 1 fold 4	50
ภาพประกอบ 34 Confusion Matrix แบบจำลอง VGG16 Version 1 fold 5	51
ภาพประกอบ 35 Learning Curve แบบจำลอง VGG16 Version 1 fold 5	51

ภาพประกอบ 36 Confusion Matrix แบบจำลอง VGG16 Version 2 fold 1	52
ภาพประกอบ 37 Learning Curve แบบจำลอง VGG16 Version 2 fold 1	52
ภาพประกอบ 38 Confusion Matrix แบบจำลอง VGG16 Version 2 fold 2	53
ภาพประกอบ 39 Learning Curve แบบจำลอง VGG16 Version 2 fold 2	53
ภาพประกอบ 40 Confusion Matrix แบบจำลอง VGG16 Version 2 fold 3	54
ภาพประกอบ 41 Learning Curve แบบจำลอง VGG16 Version 2 fold 3	54
ภาพประกอบ 42 Confusion Matrix แบบจำลอง VGG16 Version 2 fold 4	55
ภาพประกอบ 43 Learning Curve แบบจำลอง VGG16 Version 2 fold 4	55
ภาพประกอบ 44 Confusion Matrix แบบจำลอง VGG16 Version 2 fold 5	56
ภาพประกอบ 45 Learning Curve แบบจำลอง VGG16 Version 2 fold 5	56
ภาพประกอบ 46 Confusion Matrix แบบจำลอง VGG16 Version 3 fold 1	57
ภาพประกอบ 47 Learning Curve แบบจำลอง VGG16 Version 3 fold 1	57
ภาพประกอบ 48 Confusion Matrix แบบจำลอง VGG16 Version 3 fold 2	58
ภาพประกอบ 49 Learning Curve แบบจำลอง VGG16 Version 3 fold 2	58
ภาพประกอบ 50 Confusion Matrix แบบจำลอง VGG16 Version 3 fold 3	59
ภาพประกอบ 51 Learning Curve แบบจำลอง VGG16 Version 3 fold 3	59
ภาพประกอบ 52 Confusion Matrix แบบจำลอง VGG16 Version 3 fold 4	60
ภาพประกอบ 53 Learning Curve แบบจำลอง VGG16 Version 3 fold 4	60
ภาพประกอบ 54 Confusion Matrix แบบจำลอง VGG16 Version 3 fold 5	61
ภาพประกอบ 55 Learning Curve แบบจำลอง VGG16 Version 3 fold 5	61
ภาพประกอบ 56 Confusion Matrix แบบจำลอง VGG16 Version 4 fold 1	62
ภาพประกอบ 57 Learning Curve แบบจำลอง VGG16 Version 4 fold 1	62
ภาพประกอบ 58 Confusion Matrix แบบจำลอง VGG16 Version 4 fold 2	63

ภาพประกอบ 59 Learning Curve แบบจำลอง VGG16 Version 4 fold 2	63
ภาพประกอบ 60 Confusion Matrix แบบจำลอง VGG16 Version 4 fold 3	64
ภาพประกอบ 61 Learning Curve แบบจำลอง VGG16 Version 4 fold 3	64
ภาพประกอบ 62 Confusion Matrix แบบจำลอง VGG16 Version 4 fold 4	65
ภาพประกอบ 63 Learning Curve แบบจำลอง VGG16 Version 4 fold 4	65
ภาพประกอบ 64 Confusion Matrix แบบจำลอง VGG16 Version 4 fold 5	66
ภาพประกอบ 65 Learning Curve แบบจำลอง VGG16 Version 4 fold 5	66
ภาพประกอบ 66 สรุปผลการประเมินประสิทธิภาพแบบจำลองด้วยอัลกอริทึม VGG16	67
ภาพประกอบ 67 Confusion Matrix แบบจำลอง ResNet-50 Version 1 fold 1	69
ภาพประกอบ 68 Learning Curve แบบจำลอง ResNet-50 Version 1 fold 1	69
ภาพประกอบ 69 Confusion Matrix แบบจำลอง ResNet-50 Version 1 fold 2.....	70
ภาพประกอบ 70 Learning Curve แบบจำลอง ResNet-50 Version 1 fold 2	70
ภาพประกอบ 71 Confusion Matrix แบบจำลอง ResNet-50 Version 1 fold 3.....	71
ภาพประกอบ 72 Learning Curve แบบจำลอง ResNet-50 Version 1 fold 3	71
ภาพประกอบ 73 Confusion Matrix แบบจำลอง ResNet-50 Version 1 fold 4.....	72
ภาพประกอบ 74 Learning Curve แบบจำลอง ResNet-50 Version 1 fold 4	72
ภาพประกอบ 75 Confusion Matrix แบบจำลอง ResNet-50 Version 1 fold 5.....	73
ภาพประกอบ 76 Learning Curve แบบจำลอง ResNet-50 Version 1 fold 5	73
ภาพประกอบ 77 Confusion Matrix แบบจำลอง ResNet-50 Version 2 fold 1	74
ภาพประกอบ 78 Learning Curve แบบจำลอง ResNet-50 Version 2 fold 1	74
ภาพประกอบ 79 Confusion Matrix แบบจำลอง ResNet-50 Version 2 fold 2.....	75
ภาพประกอบ 80 Learning Curve แบบจำลอง ResNet-50 Version 2 fold 2	75
ภาพประกอบ 81 Confusion Matrix แบบจำลอง ResNet-50 Version 2 fold 3.....	76

ภาพประกอบ 82 Learning Curve แบบจำลอง ResNet-50 Version 2 fold 3	76
ภาพประกอบ 83 Confusion Matrix แบบจำลอง ResNet-50 Version 2 fold 4	77
ภาพประกอบ 84 Learning Curve แบบจำลอง ResNet-50 Version 2 fold 4	77
ภาพประกอบ 85 Confusion Matrix แบบจำลอง ResNet-50 Version 2 fold 5	78
ภาพประกอบ 86 Learning Curve แบบจำลอง ResNet-50 Version 2 fold 5	78
ภาพประกอบ 87 Confusion Matrix แบบจำลอง ResNet-50 Version 3 fold 1	79
ภาพประกอบ 88 Learning Curve แบบจำลอง ResNet-50 Version 3 fold 1	79
ภาพประกอบ 89 Confusion Matrix แบบจำลอง ResNet-50 Version 3 fold 2	80
ภาพประกอบ 90 Learning Curve แบบจำลอง ResNet-50 Version 3 fold 2	80
ภาพประกอบ 91 Confusion Matrix แบบจำลอง ResNet-50 Version 3 fold 3	81
ภาพประกอบ 92 Learning Curve แบบจำลอง ResNet-50 Version 3 fold 3	81
ภาพประกอบ 93 Confusion Matrix แบบจำลอง ResNet-50 Version 3 fold 4	82
ภาพประกอบ 94 Learning Curve แบบจำลอง ResNet-50 Version 3 fold 4	82
ภาพประกอบ 95 Confusion Matrix แบบจำลอง ResNet-50 Version 3 fold 5	83
ภาพประกอบ 96 Learning Curve แบบจำลอง ResNet-50 Version 3 fold 5	83
ภาพประกอบ 97 Confusion Matrix แบบจำลอง ResNet-50 Version 4 fold 1	84
ภาพประกอบ 98 Learning Curve แบบจำลอง ResNet-50 Version 4 fold 1	84
ภาพประกอบ 99 Confusion Matrix แบบจำลอง ResNet-50 Version 4 fold 2	85
ภาพประกอบ 100 Learning Curve แบบจำลอง ResNet-50 Version 4 fold 2	85
ภาพประกอบ 101 Confusion Matrix แบบจำลอง ResNet-50 Version 4 fold 3	86
ภาพประกอบ 102 Learning Curve แบบจำลอง ResNet-50 Version 4 fold 3	86
ภาพประกอบ 103 Confusion Matrix แบบจำลอง ResNet-50 Version 4 fold 4	87
ภาพประกอบ 104 Learning Curve แบบจำลอง ResNet-50 Version 4 fold 4	87

ภาพประกอบ 105 Confusion Matrix แบบจำลอง ResNet-50 Version 4 fold 5.....	88
ภาพประกอบ 106 Learning Curve แบบจำลอง ResNet-50 Version 4 fold 5.....	88
ภาพประกอบ 107 สรุปผลการประเมินประสิทธิภาพแบบจำลองด้วยอัลกอริทึม ResNet-50	89



บทที่ 1

บทนำ

1.1 ความสำคัญและความเป็นมาของงานวิจัย

ปัจจุบันโรคมะเร็งตับ นับเป็นสาเหตุการเสียชีวิตอันดับ 3 ของคนทั่วโลกในกลุ่มของโรคมะเร็ง อ้างอิงจากสถิติการเสียชีวิตของ WHO (World Health Organization) ในปี 2563 (WHO, 2021) โดยส่วนใหญ่แล้วมักเป็นผู้ป่วยมะเร็งตับในกลุ่มที่เกิดจากเซลล์ตับเอง (Hepatocellular carcinoma) ซึ่งจัดเป็นมะเร็งชนิดปฐมภูมิ (Primary Cancer) ปกติแล้วในมะเร็งของอวัยวะอื่น ๆ ต้องอาศัยผลการตรวจทางพยาธิวิทยา (Pathology) ที่เกี่ยวข้องกับการศึกษาและวินิจฉัยโรคจากการตรวจชิ้นเนื้ออวัยวะต่าง ๆ เท่านั้น แต่การตรวจโดยใช้ชิ้นเนื้อ (Biopsy) นั้นอาจนำไปสู่ผลข้างเคียงที่ไม่พึงประสงค์อย่างหนึ่ง คือ ภาวะ Tumor seeding หรือ Needle seeding เป็นภาวะที่เกิดจากชิ้นเนื้อในการตรวจหลุดลอยไปอยู่ในเซลล์อวัยวะอื่น ๆ ข้างเคียงตามทางเดินเข็ม (Needle Biopsy) ที่ทำการตัดชิ้นเนื้อนั้น ๆ หากเซลล์ที่ตัดมาเป็นเซลล์มะเร็งก็จะสามารถแพร่กระจายเซลล์ไปสู่อวัยวะข้างเคียงเหล่านั้นได้ทันที เปรียบเสมือนการงอกของเมล็ดพันธุ์พืชที่ไปเจริญในบริเวณอื่น ๆ

ในมะเร็งตับชนิด Hepatocellular carcinoma นี้ เป็นมะเร็งชนิดเดียวที่สามารถวินิจฉัยโดยใช้ภาพถ่ายทางรังสีวินิจฉัย (Diagnostic imaging) ได้ โดยไม่จำเป็นต้องใช้ผลการตรวจทางพยาธิวิทยาในการยืนยันคำวินิจฉัยทุกครั้ง และนอกจากนี้ยังมีแนวทางการรักษามุ่งเน้นในเรื่องหลักการเฝ้าระวังโรค (Surveillance) ทุก 6 เดือน ในกลุ่มผู้ป่วยที่มีความเสี่ยงการเป็นโรคมะเร็งตับ ตัวอย่างเช่น กลุ่มผู้ป่วยที่เป็นโรคไวรัสตับอักเสบบี (Hepatitis B infection) เพื่อให้สามารถตรวจหารอยโรค (Lesion) ของมะเร็งตับได้ พร้อมทั้งสามารถกำหนดแนวทางการรักษาได้อย่างเหมาะสม รวดเร็วทันต่อการรักษา และเพื่อเป็นการเพิ่มอัตราการรอดชีวิตของผู้ป่วยให้มากที่สุด การใช้ภาพถ่ายทางรังสีวินิจฉัยในการยืนยันคำวินิจฉัยโรคของรังสีแพทย์ได้ นั้นจะสามารถลดภาวะความเสี่ยงต่าง ๆ ตามที่กล่าวมาข้างต้นได้ และยังสามารถลดความเสี่ยงและความเจ็บปวดที่อาจเกิดจากการตรวจชิ้นเนื้อดังกล่าวนี้ได้อีกด้วย

การตรวจทางรังสีวินิจฉัยนั้น มีมากมายหลากหลายเทคนิค อาทิ การตรวจอัลตราซาวด์ช่องท้อง (Ultrasound) ที่ใช้ในการคัดกรองโรค (Screening) และติดตามโรค (Follow up) การตรวจวินิจฉัยโรคด้วยเครื่องเอกซเรย์คอมพิวเตอร์ (Computerized Tomography Scan: CT scan) แพทย์จะฉายรังสีเอกซเรย์ตามบริเวณที่ต้องการตรวจ แล้วใช้คอมพิวเตอร์สร้างเป็นภาพฉายอวัยวะภายในร่างกาย สามารถใช้เพื่อประกอบการวินิจฉัยหาความผิดปกติของ

ร่างกายได้ แต่เทคนิคที่ได้รับการยอมรับมากที่สุดในการใช้ยืนยันคำวินิจฉัยโรคมะเร็งตับจากข้อมูลภาพถ่ายทางรังสีวินิจฉัย คือ การตรวจคลื่นแม่เหล็กไฟฟ้า (Magnetic Resonance Imaging: MRI) เนื่องจากสามารถตรวจหาคุณลักษณะของมะเร็งตับชนิด Hepatocellular carcinoma ในกรณีของผู้ป่วยที่มีภาวะตับแข็งได้อย่างชัดเจน มีความแม่นยำสูงในการวินิจฉัยก่อนมะเร็งตับแข็งขนาดใหญ่ (Taouli, Losada, Holland, และ Krinsky, 2004) ภาพที่ได้มีความคมชัดสูงสามารถบอกขอบเขตของรอยโรคได้ถูกต้องและแม่นยำ เนื่องจากสามารถให้ข้อมูลความแตกต่างของเนื้อเยื่ออวัยวะตับได้ดี อีกทั้งการตรวจชนิดนี้เป็นการตรวจด้วยคลื่นแม่เหล็กไฟฟ้าจึงลดความเสี่ยงในการสัมผัสรังสีวินิจฉัยทางการแพทย์ ภาพถ่ายทางรังสีวินิจฉัยนี้จะถูกนำไปใช้ประกอบการวินิจฉัยการเป็นโรคมะเร็งตับของรังสีแพทย์ เพื่อใช้ในการกำหนดวางแผนทางการรักษาผู้ป่วยได้อย่างเหมาะสมและเป็นไปตามหลักมาตรฐานสากลในปัจจุบัน อ้างอิงจาก AASLD 2018 Guidance on HCC Diagnosis แต่ด้วยรอยโรคของมะเร็งตับชนิด Hepatocellular carcinoma ที่มีขนาดค่อนข้างเล็กต้องอาศัยรังสีแพทย์ผู้เชี่ยวชาญและมีประสบการณ์ในการอ่านผลตรวจจากภาพถ่ายทางรังสีวินิจฉัย ประกอบกับภาระงานที่หนักจากจำนวนผู้ป่วยที่เพิ่มมากขึ้นในแต่ละวัน ทำให้ไม่สามารถตรวจวินิจฉัยได้อย่างทันทั่วถึง จนนำไปสู่อัตราการเสียชีวิตของผู้ป่วยที่เป็นมะเร็งตับชนิด Hepatocellular carcinoma ที่เพิ่มมากขึ้นในแต่ละปี เพราะมะเร็งชนิดนี้มีระยะเวลาในการดำเนินของโรคที่ค่อนข้างเร็ว ผู้ป่วยมักเสียชีวิตภายใน 6 เดือนหรือ 1 ปี เพราะกว่าจะได้รับ การตรวจวินิจฉัยก็มักพบเป็นระยะสุดท้ายของโรคแล้ว

จากความก้าวหน้าทางเทคโนโลยีในปัจจุบัน พบว่า การสร้างแบบจำลองด้วยเทคนิคการเรียนรู้เชิงลึก (Deep Learning) เป็นวิธีการเรียนรู้แบบอัตโนมัติที่เหมาะสมกับข้อมูลที่ไม่มีโครงสร้าง (Unstructured Data) ซึ่งจะรวมถึงข้อมูลในลักษณะของรูปภาพด้วย โดยเกิดจากการเลียนแบบการทำงานของโครงข่ายระบบประสาทของมนุษย์ (Neurons) โดยนำระบบโครงข่ายเส้นประสาทเทียม (Neural Network) เหล่านี้ มาวางซ้อนทับกันหลายๆ ชั้น (Layer) จากนั้น นำไปใช้ในการเรียนรู้ข้อมูลตัวอย่างที่ต้องการศึกษา (Training) และถูกนำไปใช้ในการหารูปแบบหรือจัดหมวดหมู่ข้อมูล โครงข่ายเส้นประสาทเทียมที่มีความสามารถในการจดจำรูปแบบของภาพได้ (Pattern recognition of images) และเป็นที่นิยมใช้ในปัจจุบันนั้น เป็นโครงข่ายประสาทเทียมคอนโวลูชัน (Convolutional neural network: CNN) จัดเป็นโครงข่ายประสาทเทียมหนึ่งในกลุ่ม Bio-inspired ที่มีการใช้หลักการเดียวกันกับ Spatial Convolution โดยจะจำลองการมองเห็นของมนุษย์ที่มองพื้นที่ออกเป็นที่ย่อย ๆ และนำกลุ่มของพื้นที่ย่อย ๆ เหล่านั้น

มารวมกันอีกครั้งหนึ่ง เพื่อตรวจจับว่าในภาพนั้น ๆ เป็นสิ่งใด และการเรียนรู้เชิงลึกนี้ยังถือว่าเป็นศาสตร์อีกแขนงหนึ่งของการเรียนรู้ของเครื่องจักร (Machine Learning) อีกด้วย

ในงานวิจัยนี้ จึงมีวัตถุประสงค์ในการสร้างแบบจำลองจำแนกประเภท (Binary Classification) การเป็นโรคมะเร็งตับและไม่เป็นโรคมะเร็งตับจากภาพทางรังสีวินิจฉัย จากข้อมูลภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (Magnetic Resonance Imaging: MRI) ในบริเวณช่องท้อง เพื่อให้สามารถใช้เป็นข้อมูลประกอบการรักษาได้อย่างแม่นยำ รวดเร็วทันต่อการรักษา และสามารถเพิ่มอัตราการรอดชีวิตของผู้ป่วยได้

1.2 วัตถุประสงค์ของการวิจัย

ในการวิจัยครั้งนี้ผู้วิจัยได้ตั้งความมุ่งหมาย ไว้ดังนี้

1.2.1 เพื่อสร้างแบบจำลองการจำแนกประเภท (Binary Classification) การเป็นโรคมะเร็งตับและไม่เป็นโรคมะเร็งตับจากภาพทางรังสีวินิจฉัย โดยใช้เทคนิคการเรียนรู้เชิงลึกด้วยโครงข่ายประสาทเทียมคอนโวลูชัน (Deep Learning with Convolutional Neural Network)

1.2.2 เพื่อให้สามารถเลือกโครงสร้างโครงข่ายประสาทเทียมคอนโวลูชันแบบจำลองแบบ VGG16 และแบบ ResNet-50 ที่มีประสิทธิภาพในการจำแนกประเภท (Binary Classification) การเป็นโรคมะเร็งตับและไม่เป็นโรคมะเร็งตับจากภาพทางรังสีวินิจฉัยได้

1.3 ความสำคัญของการวิจัย

งานวิจัยนี้จะนำภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) มาศึกษาและจำแนกประเภท (Binary Classification) การเป็นโรคมะเร็งตับและไม่เป็นโรคมะเร็งตับจากภาพทางรังสีวินิจฉัยด้วยเทคนิคการเรียนรู้เชิงลึกด้วยโครงข่ายประสาทเทียมคอนโวลูชัน (Deep Learning with Convolutional Neural Network) ซึ่งเป็นเครื่องมือสำหรับการสร้างแบบจำลองด้วยเทคนิคการเรียนรู้แบบมีผู้สอน (Supervised Learning) ประเภทการจำแนกประเภทแบบ 2 ประเภท (Binary Classification) เพื่อเป็นข้อมูลสนับสนุนการตัดสินใจสำหรับรังสีแพทย์ให้สามารถวินิจฉัยโรคได้อย่างแม่นยำมากยิ่งขึ้น สามารถวางแผนทางการรักษาผู้ป่วยได้อย่างเหมาะสม และเพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองแบบต่าง ๆ ซึ่งทดลองกับข้อมูลภาพถ่ายทางรังสีวินิจฉัยจากฐานข้อมูลเปิด (Open database) ของ The Cancer Imaging Archive (TCIA) สำหรับผู้ป่วยที่เป็นโรคมะเร็งตับชนิด Hepatocellular carcinoma ใช้ชุดข้อมูล The Cancer Genome Atlas Liver Hepatocellular Carcinoma Collection (TCGA-LIHC) และสำหรับผู้ป่วยที่ไม่เป็นโรคมะเร็งตับชนิด Hepatocellular carcinoma จะใช้ชุดข้อมูลทั้งหมด 2 ชุด คือ

The Cancer Genome Atlas Cervical Kidney Renal Papillary Cell Carcinoma Collection (TCGA-KIRP) ซึ่งเป็นข้อมูลผู้ป่วยที่เป็นโรคมะเร็งไต และอีกชุดข้อมูลหนึ่งจะเป็นข้อมูลผู้ป่วยที่เป็นโรคมะเร็งตับอ่อน The Clinical Proteomic Tumor Analysis Consortium Pancreatic Ductal Adenocarcinoma Collection (CPTAC-PDA) ซึ่งข้อมูลทั้งหมดจะต้องผ่านการทำความสะอาดข้อมูล (Preprocessing) ก่อน ซึ่งจำนวนข้อมูลผู้ป่วยทั้งหมดเริ่มแรกมีทั้งหมด 229 ราย เป็นชุดข้อมูลผู้ป่วยที่เป็นโรคมะเร็งตับชนิด Hepatocellular carcinoma มีจำนวนข้อมูลผู้ป่วยทั้งหมด 98 ราย ใช้ข้อมูลชุด TCGA-LIHC สำหรับผู้ป่วยที่ไม่เป็นโรคมะเร็งตับชนิด Hepatocellular carcinoma ใช้ชุดข้อมูล TCGA-KIRP และ CPTAC-PDA มีจำนวนข้อมูลผู้ป่วยทั้งหมด 33 ราย และ 98 ราย ตามลำดับ หลังจากที่ได้ทำความสะอาดข้อมูลแล้ว และเลือกเฉพาะในระยະ T1 (Pulse sequence) เท่านั้นตามวัตถุประสงค์ของงานวิจัย

ดังนั้น ชุดข้อมูลที่จะนำไปใช้การทำการวิจัยครั้งนี้ คือ มีจำนวนข้อมูลผู้ป่วยทั้งหมด 39 ราย โดยแบ่งเป็นชุดข้อมูลผู้ป่วยที่เป็นโรคมะเร็งตับชนิด Hepatocellular carcinoma คือ TCGA-LIHC จำนวนทั้งหมด 20 ราย และเป็นชุดข้อมูลสำหรับผู้ป่วยที่ไม่เป็นโรคมะเร็งตับชนิด Hepatocellular carcinoma คือ TCGA-KIRP และ CPTAC-PDA จำนวนทั้งหมด 13 ราย และ 6 ราย ตามลำดับ

1.4 ขอบเขตของการวิจัย

1.4.1 ข้อมูลที่ใช้ในการวิจัย

ใช้ภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) ในระยະ T1 ที่เกิดขึ้นเป็น 6 ระยะเวลาย่อย (six time points: TPs) คือ ระยะเวลาก่อนฉีดสารทึบแสง (Pre-contrast phase: TP1) และ ระยะเวลาหลังฉีดสารทึบแสง 5 ระยะเวลาย่อย (Post-contrast: TP2 - TP6)

1.4.2 ประเภทงานวิจัย

สร้างแบบจำลองการจำแนกระบบรายงานคะแนนภาพรังสีด้วยระดับที่ใช้ในการบอกการเป็นโรคมะเร็งจากภาพถ่ายทางรังสีวินิจฉัย 2 ประเภทออกจากกัน (Binary Classification) ในการจำแนกการเป็นโรคมะเร็งตับและไม่เป็นโรคมะเร็งตับ

1.4.3 คุณสมบัติที่ใช้ศึกษา

ต้องการเปรียบเทียบประสิทธิภาพของแบบจำลองที่มีโครงสร้างโครงข่ายประสาทเทียมคอนโวลูชันแบบจำลองแบบ “VGG16” และแบบ “ResNet-50” ในการจำแนกการเป็นโรคมะเร็งตับและไม่เป็นโรคมะเร็งตับจากภาพถ่ายทางรังสีวินิจฉัย 2 ประเภทออกจากกัน (Binary Classification)

1.5 กรอบแนวคิดในการวิจัย

การพัฒนาแบบจำลองเพื่อจำแนกการเป็นโรคมะเร็งตับและไม่เป็นโรคมะเร็งตับจากภาพถ่ายทางรังสีวินิจฉัย สามารถแบ่งออกเป็น 5 ขั้นตอน ดังนี้

ขั้นตอนที่หนึ่ง รวบรวมข้อมูลผู้ป่วยที่มีการสั่งการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) พร้อมทั้งทำความสะอาดข้อมูลให้พร้อมใช้งาน พบว่า ข้อมูลหลังจากทำความสะอาดข้อมูล และเลือกใช้เฉพาะภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) ในระยะ T1 เท่านั้น มีจำนวนข้อมูลผู้ป่วยทั้งหมด 39 ราย โดยเป็นชุดข้อมูลผู้ป่วยที่เป็นโรคมะเร็งตับชนิด Hepatocellular carcinoma คือ TCGA-LIHC จำนวนทั้งหมด 20 ราย และเป็นชุดข้อมูลผู้ป่วยที่ไม่เป็นโรคมะเร็งตับชนิด Hepatocellular carcinoma คือ TCGA-KIRP และ CPTAC-PDA จำนวนทั้งหมด 13 ราย และ 6 ราย ตามลำดับ หลังจากนั้นทำการวิเคราะห์และสำรวจข้อมูลในมุมมองต่างๆ (Exploratory Data Analysis : EDA)

ขั้นตอนที่สอง ได้ให้แพทย์เป็นผู้คัดเลือกภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) ในระยะ T1 เพื่อในการสร้างแบบจำลองการจำแนกการเป็นโรคมะเร็งตับและไม่เป็นโรคมะเร็งตับ โดยต้องมีพื้นที่อวัยวะตับมากกว่าร้อยละ 50 ในภาพถ่ายการตรวจนั้น เนื่องจากหากมีพื้นที่อวัยวะตับน้อยกว่าร้อยละ 50 ในภาพถ่ายการตรวจ รอยโรคที่ปรากฏในอวัยวะตับนั้นจะมีขนาดเล็กค่อนข้างมาก ทำให้วินิจฉัยโรคได้ยาก ขอบเขตงานวิจัยนี้จึงได้กำหนดเกณฑ์ในการคัดเลือกภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) ที่มีพื้นที่ตับมากกว่าร้อยละ 50 เท่านั้น

ถัดมาเป็นการเตรียมข้อมูลรูปภาพ โดยการปรับขนาดของรูปภาพให้มีขนาดเท่ากับ 224 x 224 pixels เพื่อให้ตรงกับขนาดของรูปในขั้นตอนการสร้างแบบจำลอง (Input shape) จากนั้นจะเป็นการใช้เทคนิคการลดปริมาณข้อมูล (Undersampling) และเพิ่มปริมาณข้อมูล (Oversampling) เพื่อให้ข้อมูลรูปภาพที่ใช้สร้างแบบจำลองนั้นสมดุลกันในทุก 2 ประเภทข้อมูล

หลังจากนั้น จึงทำการประมวลผลภาพ (Image Preprocessing) เพื่อปรับแต่งภาพก่อนจะนำไปใช้วิเคราะห์หรือการประมวลผลอื่น ๆ ของข้อมูลภาพการสอน (Training data) จากข้อมูลภาพเดิม โดยการปรับภาพพื้นหลังทั้งหมดให้เป็นสีดำสนิท และใช้เทคนิคค้นหา Contour ที่ใหญ่ที่สุด (Find the largest contour) เพื่อเลือกบริเวณในรูปภาพที่เราสนใจ จากนั้นจึงทำการตัดเลือกรูปภาพนั้นมาจัดให้อยู่ตรงกลางในพื้นที่รูปสี่เหลี่ยมแล้วเติมขอบของภาพเพิ่มเติม (Padding) สุดท้ายใช้การปรับภาพด้วยเทคนิค CLAHE (Contrast Limited Adaptive Histogram Equalization) เพื่อเพิ่มเติมรายละเอียดของภาพให้ชัดเจนมากขึ้น

ขั้นตอนที่สาม มุ่งเน้นไปที่การสร้างแบบจำลองจำแนกการเป็นโรคมะเร็งตับและไม่ใช่โรคมะเร็งตับจากภาพทางรังสีวินิจฉัยการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) ซึ่งเป็นปัญหาแบบการจำแนกแบบ 2 ประเภท (Binary Classification) ของผู้ป่วยที่เป็นโรคมะเร็งตับและไม่ใช่โรคมะเร็งตับ โดยใช้เทคนิคการเรียนรู้เชิงลึกด้วยโครงข่ายประสาทเทียมคอนโวลูชัน (Deep Learning with Convolutional Neural Network) ด้วยโครงสร้างโครงข่ายประสาทเทียมคอนโวลูชันแบบจำลองแบบ VGG16 และแบบ ResNet-50 โดยจะแบ่งข้อมูลออกเป็นข้อมูลสอน (Training Data) ประมาณร้อยละ 80 และชุดข้อมูลทดสอบ (Testing Data) ประมาณร้อยละ 20 ซึ่งตรงนี้ได้ใช้การทดสอบแบบไขว้พับ (k - Fold Cross Validation) แบบ 5 ส่วน (5-Fold Cross Validation) เข้ามาร่วมด้วย เนื่องจากชุดข้อมูลรูปภาพที่ใช้ในการวิจัยครั้งนี้ มีความไม่สมดุลของข้อมูลอยู่ ดังนั้น จึงเลือกใช้การประเมินประสิทธิภาพของแบบจำลองนี้ มาใช้ในขั้นตอนการแบ่งข้อมูล (Split data)

ขั้นตอนที่สี่ มุ่งเน้นเรื่องการเปรียบเทียบประสิทธิภาพของโครงสร้างโครงข่ายประสาทเทียมคอนโวลูชัน แบบจำลองแบบ VGG16 และแบบ ResNet-50

1.6 ประโยชน์ที่คาดว่าจะได้รับจากงานวิจัย

1.6.1 สามารถนำแบบจำลองไปใช้ในการจำแนกการเป็นโรคมะเร็งตับและไม่ใช่โรคมะเร็งตับจากภาพทางรังสีวินิจฉัยได้ เพื่อเป็นข้อมูลประกอบการตัดสินใจสำหรับรังสีแพทย์ให้สามารถวินิจฉัยได้อย่างแม่นยำ และสามารถวางแผนแนวทางการรักษาผู้ป่วยได้อย่างเหมาะสม

1.6.2 สามารถเลือกโครงสร้างโครงข่ายประสาทเทียมคอนโวลูชันแบบจำลองแบบ VGG16 และแบบ ResNet-50 ที่มีประสิทธิภาพในการจำแนกการเป็นโรคมะเร็งตับและไม่ใช่โรคมะเร็งตับจากภาพทางรังสีวินิจฉัยได้

1.6.3 สามารถใช้แบบจำลองในการจำแนกการเป็นโรคมะเร็งตับและไม่ใช่โรคมะเร็งตับจากภาพทางรังสีวินิจฉัยได้อย่างแม่นยำ รวดเร็วทันต่อการรักษา และสามารถเพิ่มอัตราการรอดชีวิตของผู้ป่วยได้

1.6.4 สามารถนำแบบจำลองไปเป็นต้นแบบในการพัฒนา ปรับปรุง ต่อยอดให้มีประสิทธิภาพมากยิ่งขึ้นในอนาคตได้

บทที่ 2

วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

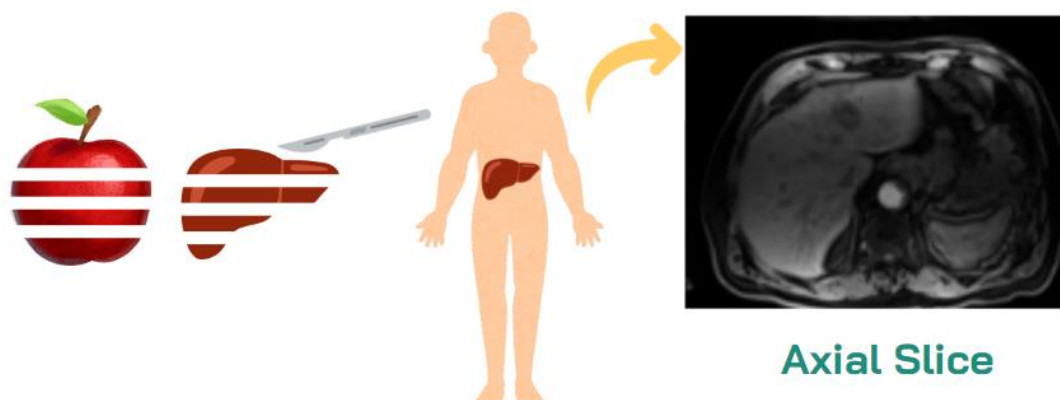
ในการวิจัยครั้งนี้ ผู้วิจัยได้ศึกษาเอกสารทฤษฎีและงานวิจัยที่เกี่ยวข้องกับการพัฒนาแบบจำลองในการจำแนกการเป็นโรคมะเร็งตับและไม่ใช่โรคมะเร็งตับจากภาพทางรังสีวินิจฉัย และได้นำเสนอตามหัวข้อดังต่อไปนี้

- 2.1 การตรวจคลื่นแม่เหล็กไฟฟ้า (Magnetic Resonance Imaging: MRI)
- 2.2 ทฤษฎีเกี่ยวกับกระบวนการก่อนการประมวลผลภาพ (Image Pre-processing)
- 2.3 ทฤษฎีเกี่ยวกับการเพิ่มปริมาณข้อมูลภาพ (Image Augmentation)
- 2.4 ทฤษฎีเกี่ยวกับการจัดการความไม่สมดุลกันของข้อมูล (Imbalance Dataset)
- 2.5 ทฤษฎีเกี่ยวกับอัลกอริทึมการเรียนรู้เชิงลึกด้วยโครงข่ายประสาทเทียมคอนโวลูชัน (Deep Learning with Convolutional Neural Network)
- 2.6 การวัดประสิทธิภาพของแบบจำลอง (Model Evaluation)
- 2.7 งานวิจัยที่เกี่ยวข้อง (Related work)

2.1 การตรวจคลื่นแม่เหล็กไฟฟ้า (Magnetic Resonance Imaging: MRI)

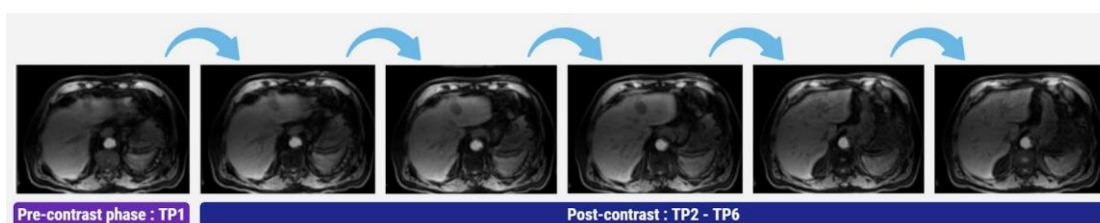
การตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) เป็นการตรวจร่างกายโดยใช้เครื่องตรวจที่มีคลื่นสนามแม่เหล็กความเข้มสูงของไฮโดรเจนอะตอม (Hydrogen, H) ซึ่งเป็นส่วนประกอบหลักที่อยู่ภายในร่างกายมนุษย์ เช่น โมเลกุลของน้ำ (H_2O) เมื่อผู้ป่วยเข้าไปอยู่ภายใต้สนามแม่เหล็กไฟฟ้า เครื่องจะส่งสัญญาณคลื่นวิทยุที่มีความถี่จำเพาะออกมา (Radiofrequency) เข้าไปกระตุ้นระบบอวัยวะที่จะตรวจ จากนั้น อวัยวะที่ถูกกระตุ้นจะมีการเปลี่ยนแปลงระดับพลังงานตามขบวนการทางฟิสิกส์ ที่เรียกว่า “การกำทอน (Resonance)” หลังจากหยุดกระตุ้นไฮโดรเจนอะตอมภายในร่างกาย แล้วมีการคายพลังงานออกมา ทั้งนี้จะมีอุปกรณ์รับสัญญาณที่ได้ออกมาเปลี่ยนแปลงเป็นสัญญาณภาพแสดงผ่านทางจอภาพ ได้ภาพเสมือนจริงของอวัยวะภายในร่างกายต่าง ๆ รวมไปถึงในส่วนที่เป็นก้อนเนื้อต่าง ๆ (Mass) หรือมะเร็ง (Cancer) อีกด้วย ภาพที่ได้จะเป็นภาพตามระนาบได้ทั้ง แนวขวาง แนวยาว และแนวเฉียง โดยอยู่ในรูปแบบของภาพถ่าย 3 มิติ (มหาวิทยาลัยมหิดล, 2018) การตรวจนี้เป็นเทคนิคที่ได้รับการยอมรับมากที่สุดในการใช้ยืนยันคำวินิจฉัยโรคมะเร็งตับจากข้อมูลภาพถ่ายทางรังสีวินิจฉัย เนื่องจากสามารถตรวจหาคุณลักษณะของมะเร็งตับชนิด Hepatocellular carcinoma ในกรณีของผู้ป่วยที่มีภาวะตับแข็งได้อย่างชัดเจน มีความแม่นยำสูงในการวินิจฉัยก้อนมะเร็งตับแข็งขนาดใหญ่ ภาพที่ได้มีความคมชัดสูง สามารถ

บอกขอบเขตของโรคได้ถูกต้องและแม่นยำ เนื่องจากให้ความแตกต่างของเนื้อเยื่อได้ดี และใช้ในการตรวจวินิจฉัยรอยโรคของผู้ป่วย



ภาพประกอบ 1 ภาพตัดในแนวขวาง (Axial slice)

การตรวจคลื่นแม่เหล็กไฟฟ้า แบ่งออกเป็น T1 และ T2 โดยที่ T1-weighted เปรียบเสมือน “Anatomy scan” เพราะสามารถแยกขอบเขตของเนื้อเยื่อ อวัยวะต่างๆ ได้อย่างชัดเจน เห็นพื้นที่ (fluid) สีดำ เป็นอวัยวะ (Solid organ) และกล้ามเนื้อ (Muscle) เป็นสีเทา (Gray) ส่วนไขมัน (Fat) สีสว่าง (Bright) ในขณะที่ T2-weighted เปรียบเสมือน “Pathology scan” เพราะบริเวณที่มีพื้นที่ (Fluid collection) หรือการบวมของเนื้อเยื่อ (Edematous tissue) จะสว่างเด่นชัดขึ้นมาจากเนื้อเยื่อข้างเคียง โดยในระยะ T1 มีภาพที่เกิดขึ้นเป็น 6 ระยะเวลาย่อย (six time points: TPs) คือ ระยะเวลาก่อนฉีดสารทึบแสง (Pre-contrast phase: TP1) และระยะหลังฉีดสารทึบแสง 5 ระยะเวลาย่อย (Post-contrast: TP2 - TP6)



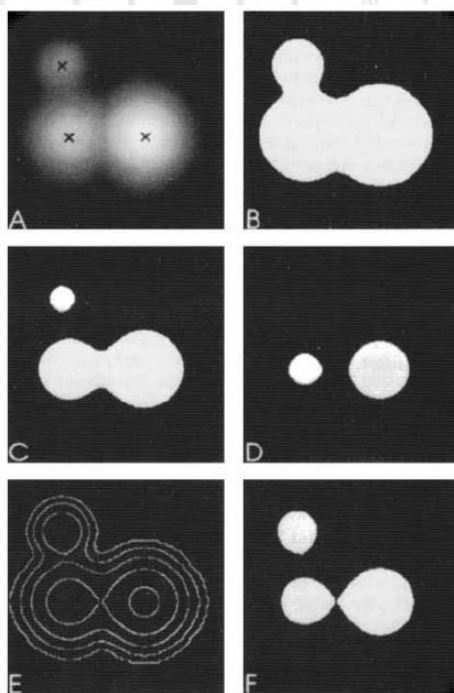
ภาพประกอบ 2 ในระยะ T1 มีภาพที่เกิดขึ้นเป็น 6 ระยะเวลาย่อย (six time points: TPs)

2.2 ทฤษฎีเกี่ยวกับกระบวนการก่อนการประมวลผลภาพ (Image Pre-processing)

2.2.1 เทคนิคหา Contour ที่ใหญ่ที่สุด (The largest contour)

เทคนิค Contour image คือ การแสดงขอบของวัตถุในภาพ โดยการวาดเส้นเชื่อมต่อความสูงหรือความเข้มของพิกเซล (Pixel) ที่มีค่าเท่ากันในภาพ ดังนั้น เส้น contour จะสามารถแสดงความเข้มของวัตถุและลักษณะที่สำคัญของวัตถุนั้นๆ ในภาพได้ เทคนิคนี้มักถูกใช้ในการประมวลผลภาพ เช่น ในการตรวจจับวัตถุ การติดตามวัตถุ หรือในงานทางการแพทย์เพื่อการวินิจฉัยภาพทางการแพทย์ เป็นต้น

ดังนั้น เทคนิคการหา The largest Contour image จึงเป็นการนำเทคนิคการทำ Contour image มาหา contour ที่ใหญ่ที่สุด หรือ contour ที่มีพื้นที่ของพื้นที่ที่ครอบคลุม (area) มากที่สุดในภาพ เทคนิคนี้มักถูกใช้ในการตรวจจับวัตถุหรือสิ่งที่ต้องการภายในภาพ การหา contour ที่ใหญ่ที่สุด จะช่วยให้สามารถแยกแยะวัตถุหรือลักษณะที่สำคัญต่าง ๆ ของภาพได้ดีขึ้น โดยการหา contour ที่ใหญ่ที่สุดนั้น สามารถทำได้โดยการใช้วิธีการวนรอบ (loop) ตั้งแต่การหา contour ทั้งหมดในภาพและหาพื้นที่ของแต่ละ contour แล้วเลือก contour ที่มีพื้นที่มากที่สุดนั้น ๆ ออกมา



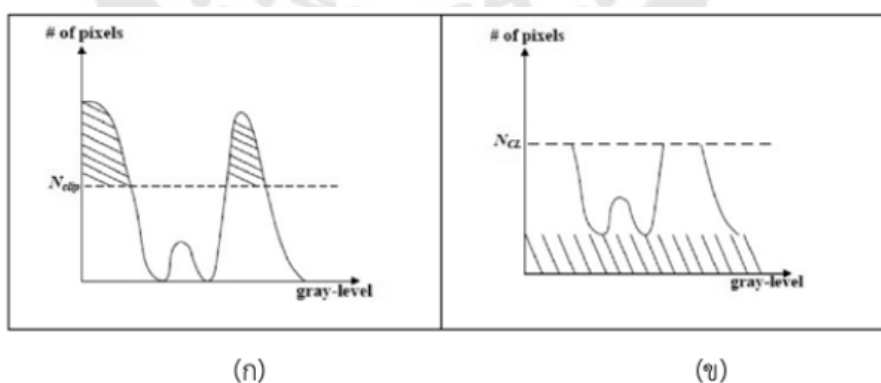
ภาพประกอบ 3 เทคนิค Largest Contour Segmentation (LCS) Procedure

ที่มา: (Manders, Hoebe, Strackee, Vossepoel, และ Aten, 1996)

2.2.2 Applying Contrast Limited Adaptive Histogram Equalization (CLAHE)

เป็นเทคนิคการเพิ่มคุณภาพรูปภาพที่พัฒนามาจากเทคนิค Histogram Equalization (HE) โดย S.M.Pizer วิธีการนี้จะพิจารณารายละเอียดข้อมูลจาก Histogram Equalization (HE) ในแต่ละค่าพิกเซล (Pixel) บนบริเวณพื้นส่วนกลางของภาพต้นฉบับ โดยค่า Histogram ที่มีระดับสูงกว่าค่าเฉลี่ยพิกเซลในระดับ gray จะถูกนำมากระจายให้กับทุก พิกเซลในรูปภาพ gray scale โดยค่าที่นำมากระจายจะเป็นค่า Histogram ที่มีระดับสูงกว่าค่าเฉลี่ยพิกเซลในระดับ gray (พิเชษ, พัฒนพงษ์, พนิดา, พัฒนพงษ์, และ นัฐริยา, 2013) เทคนิคนี้จึงเป็นการเพิ่มคุณภาพของรูปภาพด้วยการปรับแต่งความคมชัดและความเข้มของภาพที่มีปัญหาเรื่องความสว่างและความมืดไม่คงที่ในพื้นที่ต่าง ๆ ภายในรูปภาพนั้น ๆ เป็นการปรับแต่งค่าสีในภาพอ้างอิงจากการแบ่งภาพเป็นส่วน ๆ และปรับค่าสีแต่ละส่วนอย่างอัตโนมัติ การใช้เทคนิค CLAHE จึงช่วยลดการสูญเสียข้อมูลที่อาจเกิดจากการเพิ่มความเข้มของรูปภาพได้ โดยในขั้นตอนสุดท้ายเทคนิค CLAHE จะใช้การแบ่งสี เพื่อปรับสีให้เข้ากับเส้นขอบโดยอัตโนมัติ

เทคนิคการเพิ่มคุณภาพของรูปภาพประเภท CLAHE จึงมีความเหมาะสมสำหรับภาพที่มีความแตกต่างในความสว่างของพื้นที่ต่าง ๆ ของภาพ เช่น ภาพที่ถ่ายในเงาหรือภาพที่มีการกระจายความสว่างไม่สม่ำเสมอ เทคนิคนี้มักถูกใช้ในการปรับปรุงภาพถ่ายต่าง ๆ ทางทางการแพทย์หรือการปรับปรุงรูปภาพที่ถ่ายในสภาพแวดล้อมที่มีแสงน้อยหรือมีการกระจายความสว่างไม่สม่ำเสมอ เป็นต้น

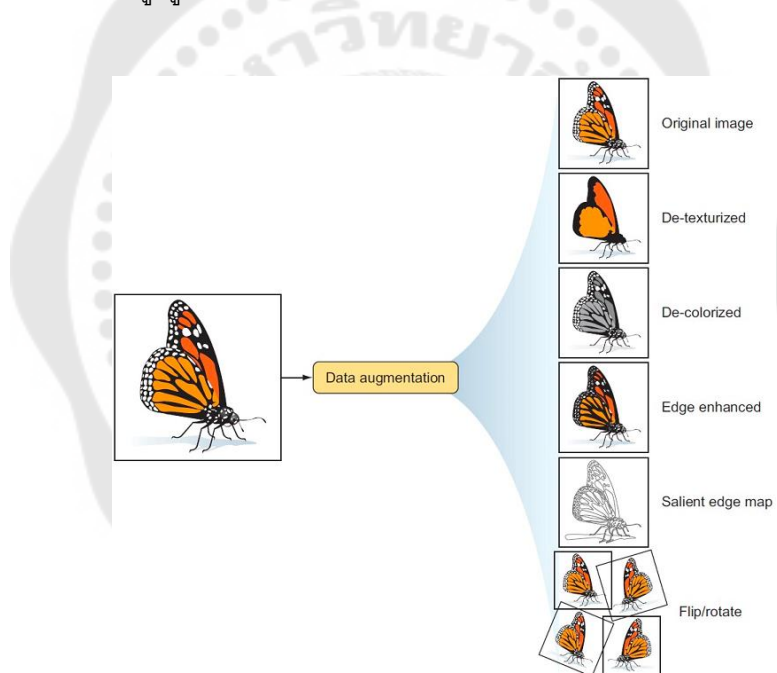


ภาพประกอบ 4 (ก) ค่า Histogram ภาพต้นฉบับ, (ข) ค่า Histogram ที่ถูกตัด

ที่มา: (พิเชษ และคนอื่น ๆ, 2013)

2.3 ทฤษฎีเกี่ยวกับการเพิ่มปริมาณข้อมูลภาพ (Image Augmentation)

ในการสร้างแบบจำลอง อาจพบปัญหาเรื่องความไม่สมดุลกันของข้อมูล เช่น รูปภาพถ่ายทางรังสีวินิจฉัยของผู้ป่วยที่เป็นโรคมะเร็งและไม่เป็นโรคมะเร็งมีสัดส่วนที่ต่างกันมากเกินไป จนนำไปสู่ปัญหาเรื่องแบบจำลองเรียนรู้ได้ดีกับชุดข้อมูลสอน (Training Data) มากๆ แต่มีประสิทธิภาพต่ำในการจำแนกข้อมูลที่ไม่เคยมาก่อน (Overfitting) ปัญหานี้สามารถแก้ได้ด้วยการเพิ่มจำนวนข้อมูลในการสอนเป็นการสังเคราะห์ข้อมูลขึ้นมาเอง ในปัจจุบันมีการใช้เทคนิค เช่น การหมุนภาพ (Rotate) การเลื่อนภาพ (Shift) การกลับภาพ (Flip) เป็นต้น มาใช้ในการเพิ่มจำนวนข้อมูล ซึ่งนอกจากเป็นการเพิ่มจำนวนของข้อมูลแล้ว เทคนิคนี้ยังช่วยเพิ่มความหลากหลายของชุดข้อมูลภาพอีกด้วย โดยในงานวิจัยนี้จะใช้เทคนิคการหมุนภาพ (Rotate) และการกลับภาพ (Flip) มาใช้ในการเตรียมข้อมูลรูปภาพก่อนการนำไปใช้ในการสร้างแบบจำลองต่อไป



ภาพประกอบ 5 เทคนิคการเพิ่มปริมาณข้อมูลภาพ (Image Augmentation)

ที่มา: <https://freecontent.manning.com/the-computer-vision-pipeline-part-3-image-preprocessing/>

2.4 ทฤษฎีเกี่ยวกับการจัดการความไม่สมดุลกันของข้อมูล (Imbalance Dataset)

ข้อมูลไม่สมดุล (Imbalanced Datasets) หมายถึง จำนวนข้อมูลในกลุ่มหนึ่งมากกว่าจำนวนข้อมูลของอีกกลุ่มหนึ่งเป็นจำนวนมาก ซึ่งอาจเกิดขึ้นได้เป็นปกติเนื่องจากลักษณะ

ธรรมชาติของข้อมูลที่ล้วนมีความแตกต่างของจำนวนในแต่ละกลุ่มอยู่แล้ว หรืออาจเกิดจากข้อจำกัดในการเก็บข้อมูลอันเนื่องมาจากค่าใช้จ่ายในการเก็บข้อมูลที่สูง เช่น ข้อมูลรูปภาพถ่ายทางรังสีวินิจฉัยของผู้ป่วยที่เป็นโรคมะเร็งและไม่เป็นโรคมะเร็ง หรือปัญหาเรื่องรูปแบบรายงานการอ่านผลภาพถ่ายทางรังสีวินิจฉัยต่างๆ ที่ไม่มีมาตรฐานที่ชัดเจน (Gold Standard) ต้องอาศัยบุคลากรในการวิเคราะห์และจัดหมวดหมู่ข้อมูล ทั้งนี้ก็นำมาซึ่งค่าใช้จ่ายในการจัดการข้อมูลเหล่านี้ เป็นต้น จากทั้งหมดนี้ส่งผลให้เกิดความเอนเอียงในการจำแนกข้อมูล ทำให้ข้อมูลในประเภทที่มีสัดส่วนน้อย (Minority) ก่อให้เกิดความผิดพลาดในการจำแนกประเภทข้อมูลมากกว่าประเภทข้อมูลที่มีสัดส่วนมากกว่า (Majority) ซึ่งจะนำไปสู่ปัญหา “ปัญหาการแบ่งกลุ่มข้อมูลผิดกลุ่ม (Misclassification)” ในปัจจุบันวิธีการที่นิยมใช้แก้ปัญหาคือการแบ่งกลุ่มข้อมูลที่ไม่สมดุลกัน คือ การจัดการปัญหาในระดับข้อมูล (Data Level) โดยอาจใช้วิธีการสุ่มตัวอย่างซ้ำ (Resampling Techniques) ซึ่งมีวิธีการเลือกใช้ที่หลากหลาย เช่น

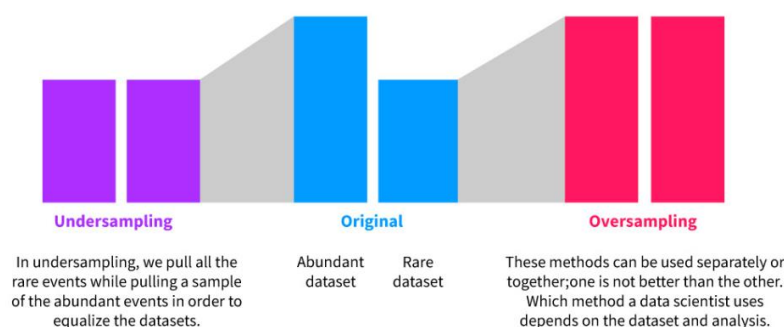
2.4.1 เทคนิคการทำ Oversampling

เป็นวิธีการที่ใช้ในการสุ่มเพิ่มจำนวนข้อมูลในกลุ่มรองให้มีจำนวนใกล้เคียงกับข้อมูลในกลุ่มหลัก มักจะใช้วิธีการเพิ่มจำนวนข้อมูลในชนิดข้อมูล (Class) ที่มีจำนวนน้อยลง โดยมักจะทำการสุ่มคัดลอกข้อมูลที่มีอยู่ในชนิดข้อมูล (Class) นั้น ๆ เป็นการช่วยปรับความสมดุลของข้อมูลและช่วยเพิ่มประสิทธิภาพให้กับแบบจำลอง

2.4.2 เทคนิคการทำ Undersampling

เป็นการสุ่มเลือกข้อมูลจากกลุ่มหลักให้ได้จำนวนที่ใกล้เคียงกับกลุ่มรอง ทำโดยการสุ่มลบจำนวนตัวอย่างของชนิดข้อมูล (Class) มากลง เพื่อให้มีจำนวนตัวอย่างใกล้เคียงกับชนิดข้อมูล (Class) ที่มีจำนวนน้อย เทคนิคการทำ Undersampling นี้อาจทำให้สูญเสียข้อมูลไปบางส่วน แต่ในบางกรณีอาจช่วยให้แบบจำลองมีประสิทธิภาพในการจำแนกข้อมูลได้ดีขึ้น โดยลดปัญหาที่เกิดจากความไม่สมดุลของข้อมูลได้ ทั้งนี้ ยังมีขั้นตอนวิธีการอื่น ๆ ที่เป็นที่ยอมรับเรื่องการสุ่มตัวอย่างซ้ำ ได้แก่ SMOTE, CE-SMOTE และ Borderline-SMOTE เป็นต้น โดยการจัดการข้อมูลที่ไม่สมดุลกันนี้ ต้องใช้ความรู้พื้นฐานทางสถิติโดยเฉพาะเรื่องการเทคนิคการสุ่มข้อมูลเข้ามาประยุกต์ใช้ร่วมด้วย

โดยในงานวิจัยนี้ จะเลือกใช้วิธีในการสุ่มเพิ่มจำนวนข้อมูล (Oversampling) และสุ่มเลือกข้อมูลจากกลุ่มหลักให้ได้จำนวนที่ใกล้เคียงกับกลุ่มรอง (Undersampling) ในการเตรียมข้อมูลก่อนนำไปใช้ต่อการสร้างแบบจำลองการเป็นโรคมะเร็งตับและไม่เป็นโรคมะเร็งตับจากภาพทางรังสีวินิจฉัย โดยใช้เทคนิคการเรียนรู้เชิงลึกด้วยโครงข่ายประสาทเทียมคอนโวลูชัน (Deep Learning with Convolutional Neural Network) (Yan, Chen, Shyu, และ Chen, 2015)



ภาพประกอบ 6 เทคนิคการจัดการความไม่สมดุลกันของข้อมูล (Imbalanced Datasets)

ที่มา: <https://www.mastersindatascience.org/learning/statistics-data-science/undersampling/>

2.4.3 เทคนิคการทำ Class weight

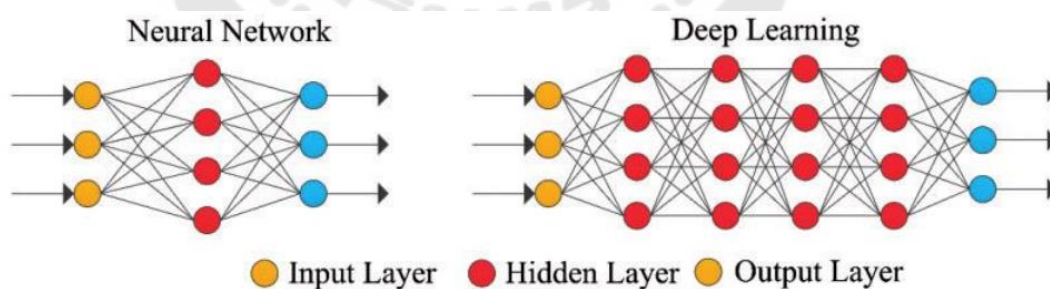
เป็นค่าที่ใช้ในการกำหนดความสำคัญของแต่ละชนิดของข้อมูลในงานการจำแนกประเภท (classification) โดยเฉพาะในกรณีที่ข้อมูลมีความไม่สมดุลกัน (Imbalance data) ชนิดข้อมูลที่มีจำนวนตัวอย่างน้อยมาก อาจมีความสำคัญมากขึ้นเมื่อเปรียบเทียบกับอีกชนิดข้อมูลที่มีจำนวนตัวอย่างมากกว่า ดังนั้น การกำหนดค่า class weight ช่วยให้แบบจำลองสามารถเรียนรู้และปรับค่าน้ำหนักของแต่ละชนิดข้อมูลให้มีการพิจารณาความสำคัญของชนิดข้อมูลในการคำนวณค่าสูญเสีย (loss) หรือการประเมินประสิทธิภาพของแบบจำลองได้ การกำหนดค่า class weight สามารถทำได้ในหลายวิธี แต่วิธีที่นิยม คือ กำหนดโดยคำนึงถึงอัตราส่วนของจำนวนตัวอย่างในแต่ละชนิดข้อมูล ซึ่งค่านี้สามารถคำนวณได้จากส่วนสูงสุดของชนิดข้อมูลที่มีจำนวนตัวอย่างมากที่สุด หรือจากการกำหนดค่ามาโดยตรงตามความเหมาะสมของงานในแต่ละกรณี

2.5 ทฤษฎีเกี่ยวกับอัลกอริทึมการเรียนรู้เชิงลึกด้วยโครงข่ายประสาทเทียมคอนโวลูชัน (Deep Learning with Convolutional Neural Network)

2.5.1 การเรียนรู้เชิงลึก (Deep Learning)

เป็นสาขาหนึ่งของการประมวลผลข้อมูลและปัญญาประดิษฐ์ที่ใช้โครงข่ายประสาทเทียม (Artificial Neural Networks) เลียนแบบการทำงานของเส้นประสาทของมนุษย์ (Neurons) มาประยุกต์ใช้ในการเรียนรู้และทำนายข้อมูลซับซ้อนแบบอัตโนมัติโดยใช้ข้อมูลขนาดใหญ่ในการฝึกสอน เพื่อให้มีระดับการเรียนรู้ที่ลึกมาก ๆ และให้สามารถจำแนกข้อมูลได้อย่างแม่นยำ โครงข่ายประสาทเทียมมักถูกออกแบบให้มีชั้นข้อมูลมากมาย (Deep layers) โดยที่แต่ละชั้นจะ

ประมวลผลข้อมูลและการเรียนรู้อย่างอิสระต่อกันไม่มีการเชื่อมโยงต่อกัน ซึ่งในขั้นตอนการเรียนรู้นี้จะให้โครงข่ายประสาทเทียมเรียนรู้จากข้อมูลของตัวอย่าง ไม่ต้องการสร้าง ตั้งกฎ หรือกำหนดลักษณะพิเศษของปัญหานั้นล่วงหน้า ผลลัพธ์ที่ได้จากโครงข่ายประสาทเทียมนี้สามารถตรวจจับรูปแบบ (Pattern) หรือจำแนกข้อมูลได้ แต่โดยทั่วไปแล้วโครงข่ายประสาทเทียมจะมีความสามารถในการเรียนรู้ได้เพียงไม่กี่ชั้นเท่านั้น เพื่อที่จะทำให้โครงข่ายประสาทเทียมสามารถคิดวิเคราะห์และประมวลผลได้เสมือนสมองของมนุษย์ ชั้นที่ไม่ได้ทำหน้าที่นำเข้าข้อมูลหรือเป็นชั้นการจำแนกหรือทำนายชนิดข้อมูลในส่วนของโครงข่ายประสาทเทียม แต่เป็นชั้นที่อยู่แทรกระหว่างตรงนี้ (Hidden layer) จะมีหน้าที่ทำการคำนวณและเรียนรู้คุณลักษณะของข้อมูล จึงมีความจำเป็นต้องมีจำนวนหลาย ๆ ชั้น เพื่อเป็นการเรียนรู้และส่งต่อข้อมูลที่ประมวลผลแล้วไปโครงข่ายประสาทเทียมถัด ๆ ไป นี่จึงเป็นหนึ่งในเหตุผลที่โครงข่ายประสาทเทียมนี้สามารถวิเคราะห์และคำนวณข้อมูลที่มีความซับซ้อนได้ ปัจจุบันเทคโนโลยีมีความก้าวหน้ามากขึ้นคอมพิวเตอร์มีประสิทธิภาพในการประมวลผลมากขึ้น จำนวนชั้นของโครงข่ายประสาทเทียมจึงสามารถเพิ่มจำนวนชั้นได้มากขึ้น และยังมีข้อทับกันหลายชั้นเพิ่มความสามารถในการคำนวณที่มีความซับซ้อนได้มากขึ้น จึงเป็นที่มาของชื่อว่า การเรียนรู้เชิงลึก (Deep Learning) (ภาพประกอบ 2) โดยปกติแล้วรูปแบบการเรียนรู้เครื่องโดยทั่วไป เมื่อมีข้อมูลดิบ (Raw data) เข้ามาจะไม่มีขั้นตอนในการประมวลผลอย่างอัตโนมัติ เพราะต้องอาศัยความรู้เฉพาะทางจากผู้เชี่ยวชาญในการเก็บรวบรวมข้อมูลการจัดหมวดหมู่ประเภทข้อมูลต่าง ๆ ในงานนั้น (Hand-Craft Features)



ภาพประกอบ 7 ความแตกต่างระหว่างโครงข่ายประสาทเทียมและการเรียนรู้เชิงลึก

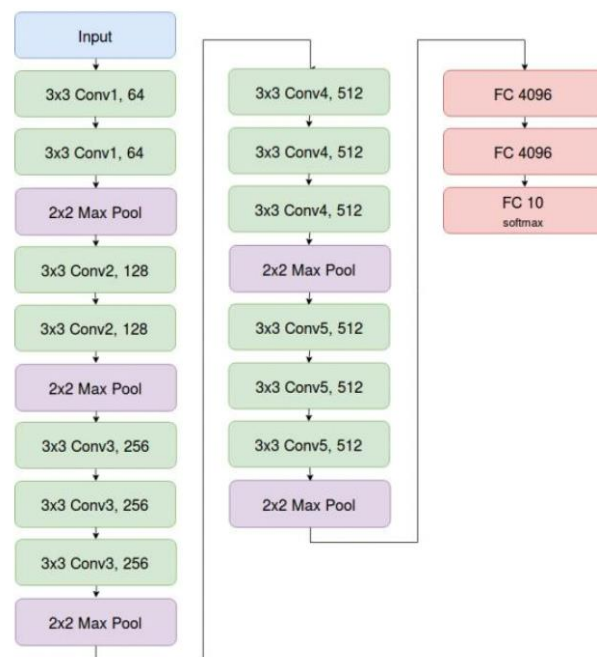
ที่มา: <https://csit.nu.ac.th>

2.5.2 โครงข่ายประสาทเทียมคอนโวลูชัน (Convolutional Neural Network: CNN)

โครงข่ายประสาทเทียมคอนโวลูชัน (Convolutional Neural Network: CNN) เป็นวิธีการที่ถูกนำมาใช้อย่างกว้างขวางทั้งในคณิตศาสตร์ สถิติ การประมวลผลสัญญาณ (Signal processing) รวมถึงการประมวลผลภาพ (Computer vision) คอนโวลูชัน ปัจจุบันถูกนำมาใช้ในรูปแบบแบบจำลองที่สามารถเรียนรู้และคัดเลือกใช้คุณลักษณะเด่น (Feature extraction) ของรูปภาพได้ ข้อดี คือ แบบจำลองเรียนรู้และคัดเลือกใช้คุณลักษณะเด่นของภาพได้เก่งกว่ามนุษย์ ดังนั้น จึงได้การประมวลผลที่แม่นยำมากกว่า ข้อแตกต่างจากโครงข่ายประสาทเทียมอัจฉริยะ (Artificial Neuron Network: ANN) คือ คำนวณน้ำหนักที่ใช้เป็นค่าชุดเดียวกันหมด จึงเป็นหนึ่งในเหตุผลที่นักวิจัยเลือกใช้ โครงข่ายประสาทเทียมคอนโวลูชัน เพราะถ้าใช้โครงข่ายประสาทเทียมอัจฉริยะปกติ นั้น แบบจำลองจะต้องเรียนรู้ค่าน้ำหนักจำนวนมาก โดยใช้การทำงานแบบโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ (Back-propagation Neural Network) โครงสร้างของโครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึก สามารถเพิ่มจำนวนของชั้นคอนโวลูชัน ได้ไม่จำกัด จึงทำให้เกิดโครงสร้างลักษณะใหม่เกิดขึ้น เช่น VGGNet และ ResNet เป็นต้น

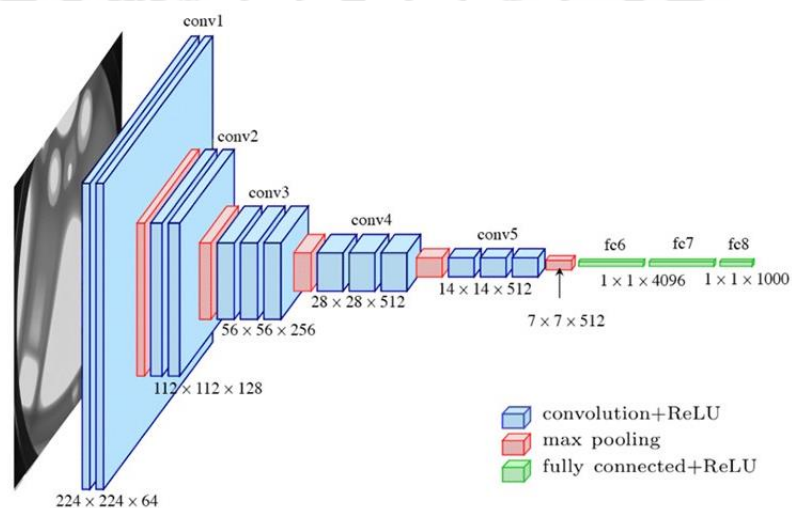
2.5.2.1 สถาปัตยกรรม VGGNet (VGGNet Architecture)

เป็นหนึ่งในแบบจำลองการเรียนรู้ของเครื่องที่ใช้ในการจำแนกประเภทภาพ ถูกพัฒนาโดยนักวิจัยที่มหาวิทยาลัย Oxford ในปี ค.ศ. 2014 และถูกนำเสนอในปี ค.ศ. 2015 ซึ่งได้รับความนิยมอย่างมากเนื่องจากความเร็วและประสิทธิภาพที่ดีในการจำแนกภาพ มีความซับซ้อนและมีจำนวนของโครงข่ายประสาทเทียมค่อนข้างมากถึง 16 หรือ 19 ชั้น โดยใช้เคอร์เนล (Kernel) ขนาด 3×3 ในการคอนโวลูชัน โครงสร้างของ VGGNet-16 ประกอบด้วย ชั้นคอนโวลูชัน จำนวน 5 กลุ่ม (Conv1-5) ในแต่ละชั้นมีจำนวนเคอร์เนล ชั้นละ 512 เคอร์เนล หลังจากชั้นคอนโวลูชัน Conv1-5 จะเป็นชั้นแมกซ์พูลลิง (Max Pooling) และถัดมาก่อนจะนำส่งข้อมูลที่ได้จากการเรียนรู้ตามโครงสร้างเหล่านี้ ไปยังชั้นเชื่อมต่อแบบสมบูรณ์ (Fully-Connected Layer) จำนวนทั้งหมด 2 ชั้น โดยมีจำนวนโนด (Node) ทั้งหมด 4,096 และ 4,096 โหนด ตามลำดับ และในชั้นสุดท้ายของชั้นเชื่อมต่อแบบสมบูรณ์ จะถูกกำหนดค่าเริ่มต้นให้มีจำนวน 1,000 โหนด (ภาพประกอบ 8, 9) โดยคำนวณหาผลลัพธ์จากการใช้วิธีซอร์ฟแมกซ์ แต่ทั้งนี้ก็สามารถปรับเปลี่ยนจำนวนโนดและการคำนวณหาผลลัพธ์จากวิธีซอร์ฟแมกซ์เป็นวิธีการอื่น ๆ ได้ ชั้นเชื่อมต่อแบบสมบูรณ์ลำดับสุดท้ายที่ใช้ในการจำแนกประเภทข้อมูลให้เหมาะสมกับงานนั้น ๆ ที่ต้องการทำ



ภาพประกอบ 8 แผนภูมิสถาปัตยกรรม VGGNet

ที่มา: <http://journal.msu.ac.th>



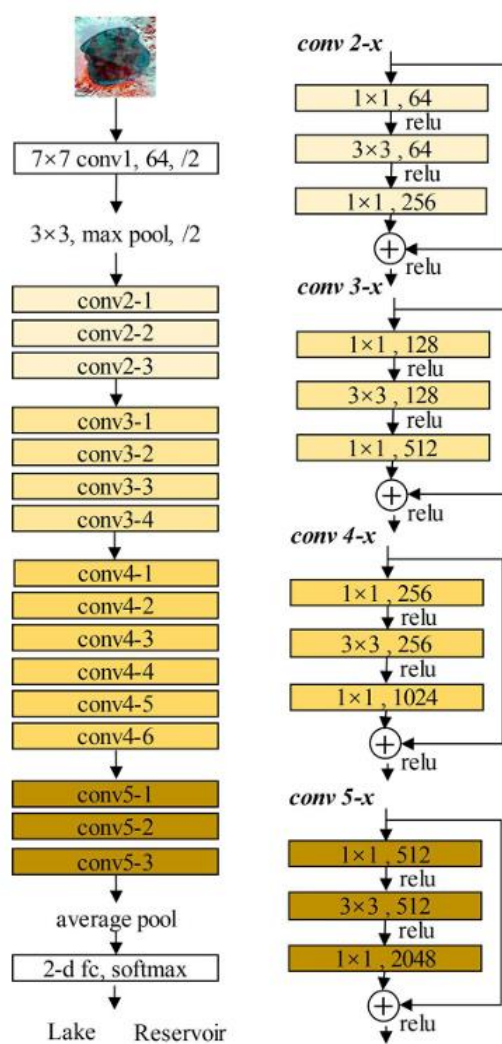
ภาพประกอบ 9 สถาปัตยกรรม VGGNet

ที่มา: <https://medium.com/analytics-vidhya/vggnet-convolutional-network-for-classification-and-detection-3543aaf61699>
<http://journal.msu.ac.th>

2.5.2.2 สถาปัตยกรรม ResNet (ResNet Architecture)

ResNet หรือชื่อเต็ม คือ Deep Residual Network ได้รับการนำเสนอครั้งแรกในงานวิจัยชื่อ Deep Residual Learning for Image Recognition ปัจจุบันได้รับความนิยมเป็นอย่างมากโดยเฉพาะงานภาพถ่ายทางการแพทย์ (Maithra และ Chiyuan, 2019) เพราะสามารถแก้ปัญหาเรื่อง Vanishing gradient ได้ ซึ่งเป็นปัญหาที่เกิดขึ้นในการฝึกสอนโครงข่ายประสาทเทียม (Neural Networks) ที่มีความลึกมาก ๆ (Deep) โดยเฉพาะในแบบจำลองที่มีจำนวนชั้น (Layers) มาก ๆ (สุริยะ, 2020) เมื่อค่าของการอนุพันธ์ของฟังก์ชันการกระตุ้น (Activation function) ในชั้นที่อยู่ก่อนหน้า (Front layers) มีค่าเข้าใกล้ศูนย์ จะทำให้การปรับค่าของพารามิเตอร์ในชั้นถัดไปกลายเป็นศูนย์ ส่งผลให้ในชั้นนั้นจะไม่มีการเรียนรู้ในชั้นตอนถัดไป และทำให้แบบจำลองไม่สามารถเรียนรู้ได้อย่างมีประสิทธิภาพ การแก้ไขปัญหา Vanishing gradient สามารถทำได้โดยการใช้ฟังก์ชันการกระตุ้นที่มีอนุพันธ์ที่มากกว่า เช่น ReLU (Rectified Linear Unit) หรือการใช้เทคนิคอื่นๆ เช่น Batch Normalization หรือการใช้เทคนิคการกำหนดค่าเริ่มต้นของน้ำหนักอย่างมีประสิทธิภาพในการฝึกสอนแบบจำลองประสาทเทียมที่มีความลึกมาก ๆ

โดยทั่วไปแล้วโครงสร้างของ ResNet-50 (Residual Network with 50 layers) ประกอบด้วยบล็อกการสร้าง (Building blocks) หลายแบบที่เรียกว่า "Residual Block" หรือ "Identity Block" ที่ช่วยให้โมเดลสามารถเรียนรู้และถอดรหัสลักษณะเด่นของภาพได้ดีขึ้น โดยลดปัญหาการสูญเสียข้อมูลในกระบวนการการฝึกสอน ทำให้สามารถสร้างแบบจำลองที่มีความลึกมากขึ้นได้โดยไม่เพิ่มความซับซ้อนของโมเดล ResNet-50 ประกอบด้วย 50 ชั้น (Layers) โดยรวมทั้งหมด ซึ่งประกอบด้วย ชั้นคอนโวลูชัน (Convolutional layers), ชั้นพูลลิง (Pooling layers), ชั้นเชื่อมต่อแบบสมบูรณ์ (Fully-Connected Layer) และฟังก์ชันการกระตุ้น (Activation function) เช่น ReLU (Rectified Linear Unit) ที่ชั้นแรก (First Layer) จะใช้ฟิลเตอร์ (Filter) ขนาด 7x7 โดยในชั้นคอนโวลูชัน (Convolutional layers) จะใช้ฟิลเตอร์ขนาด 3x3 ดังภาพประกอบ 10 โดย ResNet-50 มีลักษณะพิเศษ คือ ใช้เทคนิคของการ Skip connections หรือ Shortcut connections ที่ช่วยให้แบบจำลองสามารถผ่านข้อมูลที่ผ่านมาก่อนหน้าได้โดยตรง ช่วยลดปัญหาการสูญเสียข้อมูลในการฝึกสอนและช่วยให้สามารถสร้างโมเดลที่มีประสิทธิภาพได้ในระดับความลึกสูงได้



ภาพประกอบ 10 สถาปัตยกรรม ResNet

ที่มา: (สุริยะ, 2020)

2.5.3 ทฤษฎีการ Transfer Learning

เป็นวิธีการฝึกสอนแบบจำลองการเรียนรู้เชิงลึก (Deep Learning) เทคนิคหนึ่ง โดยการนำแบบจำลองที่ถูกฝึกสอนไว้แล้ว (Pre-trained model) จากข้อมูลหนึ่งมาใช้งานอื่น ๆ โดยไม่ต้องฝึกสอนใหม่ตั้งแต่เริ่มต้น แนวคิดหลักของ Transfer Learning คือ การถ่ายโอนความรู้ที่เกิดจากการฝึกสอนในงานหนึ่งไปใช้งานอื่น ๆ ที่มีความเกี่ยวข้อง คล้ายกัน หรือมีชนิดข้อมูล (Class) เดียวกันในชุดข้อมูลฝึกสอน ขั้นตอนหลักของ Transfer Learning ประกอบด้วย

2.5.3.1 การเลือกแบบจำลองต้นแบบ (Pre-trained Model)

เลือกแบบจำลองต้นแบบที่ถูกฝึกสอนมาแล้วจากงานที่มีความคล้ายคลึงกับงานที่เราต้องการใช้ Transfer Learning ปัจจุบันมีโมเดลที่ถูกฝึกสอนไว้แล้วในหลายสถาปัตยกรรมเช่น ResNet, VGGNet, และ MobileNet เป็นต้น

2.5.3.2 การปรับแต่ง (Fine-tuning)

เป็นการปรับปรุงแบบจำลองที่ถูกถ่ายโอนความรู้มาให้เหมาะสมกับงานที่ต้องการ โดยการปรับค่าพารามิเตอร์ (Parameters) หรือโครงสร้างของแบบจำลองต้นแบบ (Architecture model) เพื่อให้สอดคล้องกับข้อมูลและปัญหาของงานนั้น ๆ โดยเฉพาะในงานภาพตรวจต่าง ๆ ทางทางการแพทย์ที่มีวิธีการเลือกโครงข่ายประสาทเทียมคอนโวลูชัน (Deep Learning with Convolutional Neural Network) ที่แตกต่างกันตามชุดข้อมูล (Input data) (Davila, Colan, และ Hasegawa, 2024)

Table 1

Overview of fine-tuning approaches for medical image classification.

Reference	Medical domain	Backbone	Source domain	Fine-tune method	Application
Kermany et al. [45].	X-ray	Inception	ImageNet	LP	Pneumonia
Yadav et al. [46]	X-ray	VGG, Inception	ImageNet	LP	Pneumonia
Liang et al. [47]	X-ray	CNN	ChestX-ray14	FT	Pneumonia
Chouhan et al. [48]	X-ray	Ensemble	ImageNet	LP	Pneumonia
Apostolopoulos et al. [49]	X-ray	Multiple	ImageNet	Selective	Covid-19
Maghdid et al. [50]	X-ray/CT	AlexNet	ImageNet	FT	Covid-19
Ahuja et al. [51]	CT	ResNet, SqueezeNet	ImageNet	FT	Covid-19
Basaia et al. [52]	MRI	CNN	ADNI	FT	Alzheimer
Oh et al. [53]	MRI	Inception	CAE	FT	Alzheimer
Eitel et al. [54]	MRI	CNN	ADNI	FT	Multiple sclerosis
Jonsson et al. [55]	MRI	ResNet	Own	LP-FT	Brain age
Naser et al. [56]	MRI	VGG	TCIA	LP	Tumor grading
Nawaz et al. [57]	Microscopy	AlexNet	ImageNet	LP	Breast cancer
Ferreira et al. [58]	Microscopy	Inception Resnet	ImageNet	LP + Selective	Breast cancer
Bayramoglu [59]	Microscopy	Multiple	ImageNet	Dynamic	Cell Nuclei
Vesal et al. [60]	Microscopy	Inception, ResNet	ImageNet	FT	Breast cancer
Vrbancic et al. [35]	Microscopy	VGG	ImageNet	Selective	Osteosarcoma
Mahbod et al. [61]	Dermoscopy	Ensemble	ImageNet	Selective	Melanoma
Mahbod et al. [62]	Dermoscopy	Ensemble	ImageNet	FT	Melanoma
Mukhlif et al. [63]	Dermoscopy	Multiple	ImageNet/Multiple	Selective	Melanoma
Hasan et al. [64]	Dermoscopy	CNN	ImageNet	FT	Melanoma
Spolaor et al. [65]	Dermoscopy	VGG	ImageNet	Selective	Melanoma
Tajbakhsh et al. [66]	Endoscopy	AlexNet	ImageNet	LP-FT	Colonic Polyps
Kim et al. [67]	Endoscopy	AlexNet	ImageNet	LP	Colonic Polyps
Patrini et al. [68]	Endoscopy	Multiple	ImageNet	LP	Frame selection
Liu et al. [69]	Endoscopy	Multiple	ImageNet	LP	Gastric cancer
Jaafari et al. [70]	Endoscopy	Inception Resnet	ImageNet	LP + Selective	Surgical tools

ภาพประกอบ 11 การเลือกโครงข่ายประสาทเทียมคอนโวลูชันที่เหมาะสมกับชุดข้อมูล (Input data)

ที่มา: (Davila และคนอื่น ๆ, 2024)

2.5.3.3 การฝึกสอนต่อ (Training)

จากนั้น ทำการฝึกสอนแบบจำลองด้วยข้อมูลที่มีอยู่ในงานใหม่ ซึ่งแบบจำลองจะใช้ความรู้ที่เรียนรู้มาแล้วจากงานเดิม (Pre-trained) มาเรียนรู้และฝึกอบรมปรับปรุงด้วยข้อมูลใหม่ที่ได้รับ (Input data)

Transfer Learning มักถูกใช้ในกรณีที่ข้อมูลสำหรับฝึกสอนแบบจำลองใหม่มีขนาดเล็กหรือมีข้อมูลไม่เพียงพอ หรือในอีกกรณี คือ ขาดความเหมาะสมที่จะใช้แบบจำลองใหม่ โดยเริ่มจากศูนย์ เพราะการฝึกสอนแบบจำลองใหม่ อาจต้องใช้เวลาและทรัพยากรในการฝึกอบรมมากกว่าการใช้ Transfer Learning เทคนิคนี้จึงเข้ามาช่วยลดเวลาและความยุ่งยากในการพัฒนาแบบจำลองใหม่ ๆ ได้

2.6 การวัดประสิทธิภาพของแบบจำลอง (Model Evaluation)

2.6.1 Confusion Matrix

Confusion Matrix เป็นตารางข้อมูลสำคัญในการวัดความสามารถของการเรียนรู้ของเครื่อง (Machine learning) ในการแก้ปัญหาการจำแนกประเภท (Classification) มักใช้ในการประเมินผลลัพธ์ของการทำนาย (Prediction) ที่ถูกทำนายจากแบบจำลองที่สร้างขึ้นจากการเรียนรู้ของเครื่อง โดยมีแนวคิดเปรียบเทียบสัดส่วนของข้อมูลที่แบบจำลองทำนาย (Prediction) กับความจริงของข้อมูล (True data)

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

ภาพประกอบ 12 ตัวอย่างตาราง Confusion Matrix ขนาด 2x2

ที่มา: <https://medium.com/@cheng3374/วัดประสิทธิภาพ-model-จาก-confusion-matrix-69d391bcd48>

โดยค่า TP, TN, FP และ FN อธิบายได้ดังนี้

True Positives (TP): คือจำนวนตัวอย่างที่แบบจำลองทำนายว่าเป็น Positive และมีค่าจริงเป็น Positive หมายถึงการทำนายถูกต้องว่าเป็น Positive

True Negatives (TN): คือจำนวนตัวอย่างที่จำลองทำนายว่าเป็น Negative และมีค่าจริงเป็น Negative หมายถึงการทำนายถูกต้องว่าเป็น Negative

False Positives (FP): คือจำนวนตัวอย่างที่จำลองทำนายว่าเป็น Positive แต่มีค่าจริงเป็น Negative หมายถึงการทำนายผิดว่าเป็น Positive ซึ่งเรียกอีกชื่อหนึ่งว่า Type I error

False Negatives (FN): คือจำนวนตัวอย่างที่จำลองทำนายว่าเป็น Negative แต่มีค่าจริงเป็น Positive หมายถึงการทำนายผิดว่าเป็น Negative ซึ่งเรียกอีกชื่อหนึ่งว่า Type II error “จริง”

ซึ่งในทางการแพทย์ เราต้องการลดค่า False Negatives (FN) ให้น้อยที่สุดเท่าที่จะเป็นไปได้ เนื่องจาก False Negatives คือการระบุว่าผู้ป่วยไม่มีโรค ทั้งที่จริงแล้ว ผู้ป่วยเป็นโรค ซึ่งอาจทำให้พลาดการรักษาหรือไม่ได้รับการรักษาตามที่ต้องการ ส่งผลกระทบต่อสุขภาพของผู้ป่วยและสังคมโดยรวมอย่างมาก เช่น การพลิกฟื้นของโรค คุณภาพชีวิตที่ลดลง หรือการแพร่ระบาดของโรค เป็นต้น เพราะถึงแม้ว่า การลด False Positives ก็เป็นสิ่งที่ดีเช่นกัน แต่มีผลกระทบน้อยกว่า False Negatives โดยในบางกรณี False Positives อาจทำให้มีความกังวลเกินไปหรือต้องรับการตรวจสอบเพิ่มเติมเท่านั้น แต่ก็ยังมีความเป็นไปได้ที่จะไม่มีผลกระทบอื่นๆ ต่อสุขภาพของผู้ป่วย

ดังนั้น การลด False Negatives ถือเป็นความสำคัญสูงสุดในทางการแพทย์ในส่วนนี้ แม้ว่าการควบคุมทั้งสองรายการนี้จะเป็นสิ่งที่ดีที่สุดในกรณีใดก็ตาม เพราะมันจะช่วยลดความเสี่ยงทั้งสองรูปแบบของข้อผิดพลาดที่เกิดขึ้นได้ลงในการวินิจฉัยและการรักษาโรคในอนาคต

2.6.1.1 Accuracy

คือ อัตราส่วนของตัวอย่างที่แบบจำลองทำนายถูกต้องทั้งหมด เป็นค่าเปรียบเทียบระหว่างจำนวนทั้งหมดของ True Positives และ True Negatives กับจำนวนทั้งหมดของตัวอย่างที่ทำนายทั้งหมด ดังสมการด้านล่าง

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

2.6.1.2 Precision

คือ อัตราส่วนของตัวอย่างที่โมเดลทำนายว่าเป็น Positive และมีค่าจริงเป็น Positive รวมกับจำนวนทั้งหมดของ True Positives

$$\text{Precision} = \frac{TP}{TP + FP}$$

2.6.1.3 Recall (Sensitivity)

คืออัตราส่วนของตัวอย่างที่มีค่าจริงเป็น Positive และโมเดลทำนายถูกต้องว่าเป็น Positive เทียบกับจำนวนทั้งหมดของ True Positives และ False Negatives

$$\text{Recall} = \frac{TP}{TP + FN}$$

2.6.1.4 F1-Score

เป็นค่าเฉลี่ยฮาร์มอนิกค่า (harmonic mean) ระหว่าง precision และ recall และใช้สำหรับการประเมินความสามารถของโมเดลในการจำแนกประเภท

$$\text{F1 score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

2.6.2 Sensitivity (ความไว)

สัดส่วนของข้อมูลที่เป็น “จริง” หารด้วยเหตุการณ์ทั้งหมด (ทั้งที่เป็นส่วนที่แบบจำลองทำนายถูกและผิด เหมือนกับสูตรในการคำนวณของ Recall)

2.6.3 Specificity (ความจำเพาะ)

สัดส่วนของข้อมูลที่เป็น “ไม่จริง” สำหรับภาวะนั้น ๆ (เช่น สัดส่วนของการตรวจไม่พบโรคในผู้ป่วยไม่ป่วย หารด้วยเหตุการณ์ทั้งหมด (ทั้งที่เป็นส่วนที่แบบจำลองทำนายถูกและผิด))

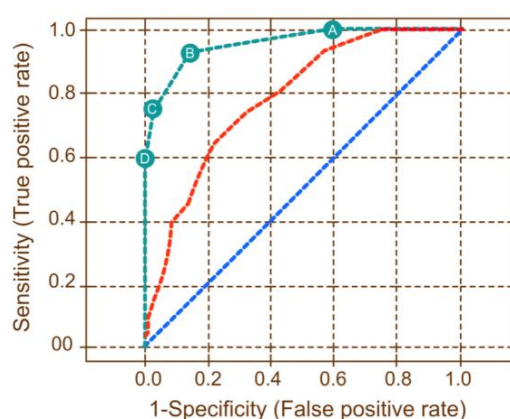
		Predicted Class		
		Positive	Negative	
Actual Class	Positive	True Positive (TP)	False Negative (FN) Type II Error	Sensitivity $\frac{TP}{(TP + FN)}$
	Negative	False Positive (FP) Type I Error	True Negative (TN)	Specificity $\frac{TN}{(TN + FP)}$
		Precision $\frac{TP}{(TP + FP)}$	Negative Predictive Value $\frac{TN}{(TN + FN)}$	Accuracy $\frac{TP + TN}{(TP + TN + FP + FN)}$

ภาพประกอบ 13 ตัวอย่างตาราง Confusion Matrix ขนาด 2x2

ที่มา: https://medium.com/@mirthful_sunset_cattle_231/confusion-matrix-48cc396b1b58

2.6.4 Receiver Operating Characteristics (ROC) Curve

เป็นกราฟที่มีโครงสร้าง แกน x เป็น False positive rate (หรือมีค่าเท่ากับ 1-Specificity) และแกน y เป็น True positive rate (หรือเท่ากับ Sensitivity) เป็นการพิจารณาคุณสมบัติทั้ง 2 ค่าของ sensitivity และ specificity ไปพร้อมๆ กัน ยกตัวอย่างเช่น การมีทางเลือก 2 ทาง คือ การเป็นโรคหรือการไม่เป็นโรค



ภาพประกอบ 14 ROC curve แสดงความสัมพันธ์ระหว่างแกน X (false positive rate หรือ 1-specificity) และแกน Y (true positive rate หรือ sensitivity)

ที่มา: https://meded.psu.ac.th/binla/class02/B7_367_261/Diagnostic_testing/index6.html

2.6.5 K-Fold Cross Validation

Cross Validation เป็นวิธีการคาดการณ์ค่าความผิดพลาดของแบบจำลอง พื้นฐานหลักการจะเป็นการสุ่มตัวอย่าง (Resampling) โดยการแบ่งข้อมูลออกเป็นกลุ่มหลาย ๆ กลุ่ม เรียกว่า folds ซึ่งเราสามารถกำหนดได้ว่า ต้องการแบ่งกลุ่มข้อมูลออกเป็นกี่กลุ่ม แต่โดยปกติแล้ว ที่นิยมใช้กัน มักกำหนดให้มีค่าเท่ากับ 5 (5 folds) และทำการหาข้อผิดพลาดจำนวน K รอบ โดยแต่ละรอบคำนวณข้อมูลชุดหนึ่งจากข้อมูล K ชุด จะถูกเลือกมาเป็นข้อมูลชุดทดสอบ (Validation set) และข้อมูลอีก K-1 ชุด จะถูกใช้เป็นข้อมูลชุดฝึกสอน (Training set)

ในแต่ละรอบของการทำ Cross Validation แบบจำลองจะถูกฝึกอบรมด้วยข้อมูลชุดฝึกสอน (Training set) และจะถูกทดสอบด้วยข้อมูลชุดทดสอบ (Validation set) เพื่อประเมินประสิทธิภาพของแบบจำลองที่ได้ และจะถูกนำมารวมกันเพื่อให้ได้ประเมินประสิทธิภาพของแบบจำลองโดยภาพรวม หลักการของ K-Fold Cross Validation สามารถช่วยลดความผิดพลาดที่เกิดจากการใช้ข้อมูลจำกัดในการทดสอบประสิทธิภาพของแบบจำลอง และช่วยให้การประเมินประสิทธิภาพนั้น มีความเชื่อถือมากยิ่งขึ้นต่อผลลัพธ์ที่ได้ในชุดข้อมูลที่ไม่เคยเห็นมาก่อน

K-FOLD CROSS – VALIDATION

Iteration 1	Test	Train	Train	Train	Train
Iteration 2	Train	Test	Train	Train	Train
Iteration 3	Train	Train	Test	Train	Train
Iteration 4	Train	Train	Train	Test	Train
Iteration 5	Train	Train	Train	Train	Test

ภาพประกอบ 15 รูปภาพขั้นตอนการทำ K-FOLD CROSS – VALIDATION

2.7 งานวิจัยที่เกี่ยวข้อง (Related work)

การทบทวนวรรณกรรมของงานวิจัยนี้ได้ทำการศึกษาค้นคว้าวิจัยที่เกี่ยวข้องกับการสร้างแบบจำลองการจำแนกประเภทแบบหลายประเภท (Multiclass Classification) ของโรคมะเร็งตับ (Hepatocellular carcinoma) โดยใช้เทคนิคการเรียนรู้เชิงลึก ด้วยโครงข่ายประสาทเทียมคอนโวลูชัน (Deep Learning with Convolutional Neural Network)

(1) บทความวิจัยเรื่อง Deep Learning with Convolutional Neural Network for Differentiation of Liver Masses at Dynamic Contrast-enhanced CT: A Preliminary Study (Yasaka, Akai, Abe, และ Kiryu, 2018)

งานวิจัยนี้ ต้องการทดสอบประสิทธิภาพการวินิจฉัยการจำแนกความแตกต่างของก้อนเนื้อตับ ว่าเป็นก้อนเนื้อดีหรือก้อนเนื้อร้าย (มะเร็ง) ออกเป็นทั้งหมด 5 ประเภท (Category) คือ

- Category A : โรคมะเร็งตับ (Hepatocellular carcinoma - HCC)
- Category B : ก้อนเนื้อมะเร็งตับ นอกเหนือจาก Category A และ HCCs ในระยะแรก
- Category C : ก้อนเนื้อตับที่คลุมเครือหรือมีรอยโรค ที่ไม่ใช่ Category D และ E
- Category D : เนื้องอกหลอดเลือดซึ่งจัดเป็นปานชนิดหนึ่ง (Hemangioma)
- Category E : ซีสต์ (Cysts)

จากการตรวจวินิจฉัยโรคด้วยเครื่องเอกซเรย์คอมพิวเตอร์ (Computerized Tomography Scan: CT scan) ภายหลังจากได้รับการฉีดสารทึบแสงทางหลอดเลือดดำในระยะเวลาที่แตกต่างกัน (Dynamic Contrast-enhanced CT) เป็นข้อมูลในช่วงมกราคม 2556 - ธันวาคม 2556 ภาพถ่ายรังสีวินิจฉัยที่ไม่ได้ใช้ในครั้งนี้ คือ กรณีที่เป็นก้อนเนื้อตับที่มีการรักษาด้วย TACE และกรณีที่ก้อนเนื้อตับอยู่ในกลุ่มผู้ป่วยที่อายุน้อยกว่า 20 ปี ภาพถ่ายจากการตรวจวินิจฉัยโรคด้วยเครื่องเอกซเรย์คอมพิวเตอร์ที่ใช้ในงานวิจัยนี้ มีทั้งหมด 3 ระยะ

- ไม่มีการใช้สารทึบแสง (Non contrast-agent enhanced phase)
- สารทึบแสงยังคงอยู่ในหลอดเลือดแดงไม่ได้อยู่ในอวัยวะและเนื้อเยื่ออ่อนอื่น ๆ (Arterial phase)
- ล้างสารทึบแสงออกจากช่องท้อง (Delayed phase)

จำนวนภาพถ่ายทางรังสีวินิจฉัยที่ใช้ในงานวิจัยนี้ของก้อนเนื้อตับแต่ละประเภท (Category A, B, C, D และ E) ประกอบด้วย 240, 121, 320, 207 และ 180 ตามลำดับ

โดยใช้เทคนิคการเรียนรู้เชิงลึกด้วยโครงข่ายเส้นประสาทเทียม (Deep Learning with Convolutional Neural Network) ที่มีโครงสร้างเป็น 6 โครงข่ายประสาทเทียมคอนโวลูชัน 3 Maximum pooling และ 3 fully connected layers เป็นงานแบบการจำแนกหลายประเภท (Multi-Classification) มีกลุ่มตัวอย่างผู้ป่วย 460 คน 1,068 รูปภาพ (.JPEG) ชุดข้อมูลรูปภาพ 1,068 รูป นำมาทำการเพิ่มปริมาณข้อมูลรูปภาพ (Image Augmentation) จาก 52 ปัจจัย (Factors) ได้ชุดรูปภาพถ่ายทั้งหมด 55,536 รูป (1,068 ชุดรูปภาพ x 52) เพื่อให้เป็นข้อมูลใช้สอน

สำหรับแบบจำลอง (Training data) และมีการเตรียมข้อมูลรูปภาพโดยการทำ Image processing (ลดขนาดให้เหลือ 70 x 70 pixels) จากนั้น นำภาพถ่ายทางรังสีสินิจัดนี้ไปสร้างเป็นแบบจำลองที่แตกต่างกัน 5 ประเภท คือ

1. Triphasic model คือ สร้างแบบจำลองโดยใช้ภาพถ่ายในระยะไม่มีการใช้สารทึบแสง (Non contrast-agent enhanced phase) สารทึบแสงยังคงอยู่ในหลอดเลือดแดงไม่ได้อยู่ในอวัยวะและเนื้อเยื่ออื่น ๆ (Arterial phase) และล้างสารทึบแสงออกจากช่องท้อง (Delayed phase)

2. Art/ Del model คือ สร้างแบบจำลองโดยใช้ภาพถ่ายในระยะสารทึบแสงยังคงอยู่ในหลอดเลือดแดงไม่ได้อยู่ในอวัยวะและเนื้อเยื่ออื่น ๆ (Arterial phase) และล้างสารทึบแสงออกจากช่องท้อง (Delayed phase)

3. Unenhanced model คือ สร้างแบบจำลองโดยใช้ภาพถ่ายในระยะไม่มีการใช้สารทึบแสง (Non contrast-agent enhanced phase)

4. Arterial mode คือ สร้างแบบจำลองโดยใช้ภาพถ่ายในระยะสารทึบแสงยังคงอยู่ในหลอดเลือดแดงไม่ได้อยู่ในอวัยวะและเนื้อเยื่ออื่น ๆ (Arterial phase)

5. Delayed model คือ สร้างแบบจำลองโดยใช้ภาพถ่ายในระยะล้างสารทึบแสงออกจากช่องท้อง (Delayed phase)

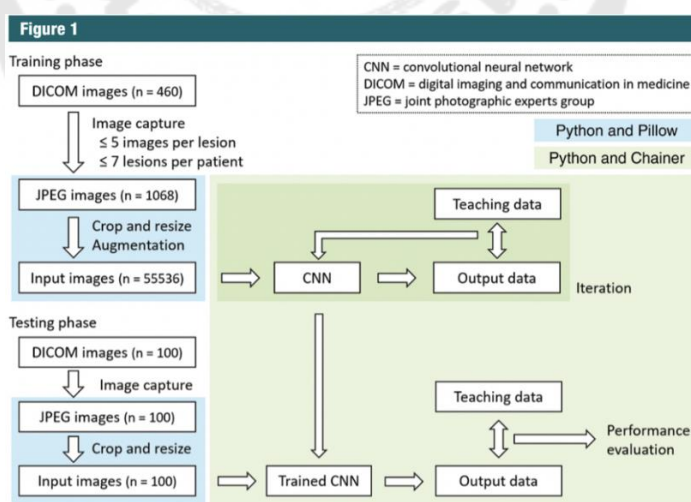


Figure 1: A flowchart of the study process from image capture to training and testing phase with CNN.

ภาพประกอบ 16 แสดงขั้นตอนในการสร้างแบบจำลองของงานวิจัย Deep Learning with Convolutional Neural Network for Differentiation of Liver Masses at Dynamic Contrast-enhanced CT: A Preliminary Study

ผลจากวิจัย พบว่า Model triphasic (Pre-contrast, Arterial, Delayed phase) มีประสิทธิภาพสูงในการจำแนกก้อนเนื้อตับออกเป็น 5 ประเภท โดยมีค่า accuracy ในการจำแนกความแตกต่างของก้อนเนื้อตับ เท่ากับ 0.84 และมีค่า receiver operating characteristic curve (ROC) เท่ากับ 0.92 ดังนั้น เทคนิคการเรียนรู้เชิงลึกด้วยโครงข่ายเส้นประสาทเทียมและการใช้ภาพถ่ายการตรวจวินิจฉัยโรคด้วยเครื่องเอกซเรย์คอมพิวเตอร์ครบทั้ง 3 ระยะตามที่ได้กล่าวมาข้างต้น แสดงให้เห็นว่า แบบจำลองมีประสิทธิภาพสูงในการจำแนกก้อนเนื้อตับโดยใช้ภาพถ่ายการตรวจวินิจฉัยโรคด้วยเครื่องเอกซเรย์คอมพิวเตอร์

(2) บทความวิจัยเรื่อง State of the Art in Artificial Intelligence and Radiomics in Hepatocellular Carcinoma (Castaldo และคนอื่น ๆ, 2021)

งานวิจัยนี้ มุ่งเน้นในการทบทวนงานวิจัยต่างๆ ที่ใช้เทคนิคการเรียนรู้ของเครื่อง (Machine Learning) และทำการศึกษาเกี่ยวข้องกับเรื่องโรคมะเร็งตับที่เกิดจากเซลล์ตับเอง (Hepatocellular carcinoma) พร้อมนำเสนอแนวทางการพัฒนางานด้านนี้ในอนาคตว่า การตรวจคลื่นแม่เหล็กไฟฟ้า (Magnetic Resonance Imaging: MRI) มีความแม่นยำมากที่สุดในปัจจุบัน ใช้การวัดประสิทธิภาพของแบบจำลอง (Model Evaluation) ด้วย Accuracy และ ROC (Receiver Operating Characteristic curve) ซึ่งการใช้ข้อมูลภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI)

มีประสิทธิภาพสูงในการจำแนกโรคมะเร็งตับ (Hepatocellular carcinoma) โดยมีค่า Accuracy สูงมากกว่า 90% ดังภาพประกอบ 13

Authors	ML Algorithm	Aim	Imaging Modality	Performance
Vivanti et al. [40]	CNN	Detection of tumor recurrence based on CT volume/tumor load	CT	Accuracy: 86%
Wenqi et al. [41]	CNN	Diagnostic accuracy of a three-phase CT protocol without PV vs. four-phase CT protocol	CT	Accuracy: 85.6% vs. 83.3%
Yamada et al. [42]	CNN	Diagnosis of primary liver cancers using transfer learning	CT	Mean DP: 44.1%, 44.2%, and 43.7% ***
Hamm et al. [43]	CNN	Classify liver lesions	MRI	Accuracy: 92%
Wu et al. [44]	CNN	CNN model for LI-RADS grading	MRI	AUC: 0.95 ****
Jansen et al. [45]	extremely randomized trees classifier	Automated classification system cataloguing liver lesions as adenoma, cyst, hemangioma, HCC and metastasis	MRI	Sensitivity/Specificity: 80/78%, 93/93%, 84/82%, 73/56% and 62/77% *****
Zhen et al. [46]	CNN	Detecting and categorizing liver tumors	MRI	AUC: 0.98, 0.99, 0.96 *****
Preis et al. [47]	ANN	Analyze 18F-FDG PET-CT liver uptake of patient at risk of developing HCC	PET	AUC: 0.89

ML: machine learning; CNN: convolutional neural network; k-NN: k-nearest neighbor; SVM: support vector machine; RF: random forest; DL: deep learning; PV: portal venous; AUC: area under the curve; DP: diagnostic performance; ANN: artificial neural network. * Results respectively for focal liver lesion detection and focal liver lesion characterization. ** Results respectively for healthy liver, cirrhosis and HCC. *** Results respectively by pixel shifts, rotations, and skew misalignments transfer learning methods. **** Result for differentiation between LR-3 and LR-4/LR-5 tumors. ***** Results respectively for adenoma, cyst, hemangioma, HCC and metastasis detection. ***** Results respectively for HCC, metastasis and other primary malignancies.

ภาพประกอบ 17 ข้อมูลแสดงประสิทธิภาพของแบบจำลอง (Model Evaluation)

จากการใช้ข้อมูลภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI)

(3) บทความวิจัยเรื่อง Deep learning LI-RADS grading system based on contrast enhanced multiphase MRI for differentiation between LR-3 and LR-4/LR-5 liver tumors (Wu และคนอื่น ๆ, 2020)

งานวิจัยนี้ ต้องการสร้างแบบจำลองการจำแนกประเภทของ LR-3 (ไม่เป็นมะเร็ง) และ LR-4/ LR-5 (เป็นมะเร็ง) ซึ่งในงานวิจัยนี้เป็นงานแก้ปัญหาแบบ Binary Classification ของโรคมะเร็งตับ (Hepatocellular carcinoma) โดยใช้ข้อมูลภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (Magnetic Resonance Imaging: MRI) จากจำนวนผู้ป่วยทั้งหมด 59 ราย มีภาพถ่าย LI-RADS ทั้งหมด 89 ภาพ ใช้โครงข่ายประสาทเทียม (Convolutional Neural Network: CNN) ที่มีโครงสร้างเป็น Alexnet และใช้ Weight เป็น “Imagenet” ดังภาพประกอบ 14

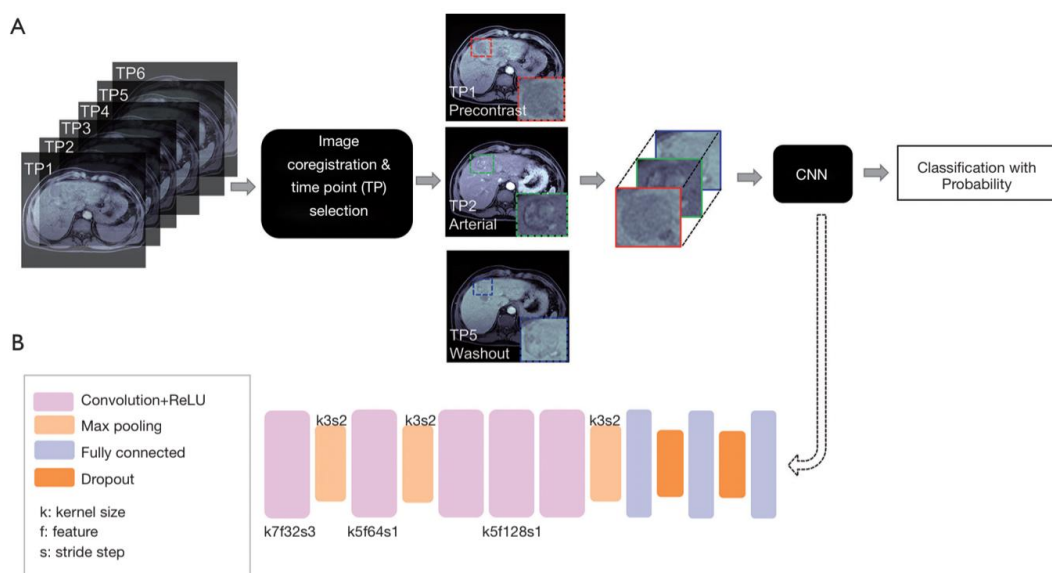


Figure 1 The design of the DL LI-RADS grading system using contrast enhanced multiphase liver MRI. (A) The workflow of the DL-driven LI-RADS grading system; (B) AlexNet model architecture used for CNN-based image feature extraction and classification. DL, deep learning; LI-RADS, Liver Imaging Reporting and Data System; CNN, convolutional neural network.

ภาพประกอบ 18 โครงสร้างเป็น Alexnet และใช้ Weight เป็น “Imagenet”

ใช้การวัดประสิทธิภาพของแบบจำลอง (Model Evaluation) ด้วย Accuracy, Sensitivity, Precision และ AUC (Area under curve) โดยจากการทดลองนี้ พบว่า การใช้ข้อมูล (Dataset) C2 นั้น ให้ประสิทธิภาพสูงสุดในการจำแนกโรคมะเร็งตับ (Hepatocellular carcinoma) โดยมีค่า Accuracy เท่ากับ 0.90, Sensitivity เท่ากับ 1.0, Precision เท่ากับ 0.84 และ AUC เท่ากับ 0.95 ซึ่งข้อแตกต่างจากการใช้ข้อมูลอื่น ๆ คือ C1 (All 6 phases), C3 (C2 + hepatobiliary phase),

C4 (Pre, Me, MW) และ C2 without transfer learning ในการสร้างแบบจำลองด้วยโครงข่ายประสาทเทียม (Convolutional Neural Network: CNN) ที่มีโครงสร้างเป็น Alexnet และใช้ Weight เป็น “Imagenet” ของชุดข้อมูล (Dataset) C2 นั้น คือ มีการใช้ข้อมูลเพียง 3 ระยะเท่านั้น ประกอบด้วย ระยะ Pre-contrast phase หมายถึง ไม่มีการใช้สารทึบแสง ระยะ Arterial phase หมายถึง สารทึบแสงยังคงอยู่ในหลอดเลือดแดงไม่ได้อยู่ในอวัยวะและเนื้อเยื่ออื่น ๆ และระยะ Washout phase หมายถึง ล้างสารทึบแสงออกจากช่องท้อง ดังภาพประกอบ 15

Table 1 The classification performance of the CNN model using datasets acquired at different combinations of image phases

Dataset	Accuracy	Precision	Sensitivity	F1 score	AUC
C1 (all 6 phases)	0.833	0.800	0.889	0.842	0.92
C2 (pre, arterial, washout)	0.900	0.835	1.0	0.909	0.95
C3 (C2 + hepatobiliary phase)	0.833	0.803	0.889	0.843	0.92
C4 (pre, me, mw)	0.789	0.814	0.756	0.780	0.91
C2 without transfer learning	0.767	0.777	0.778	0.767	0.90

pre, pre-contrast; arterial, arterial phase; washout, washout phase; me, maximum enhancement phase; mw, maximum washout phase; CNN, convolutional neural network; AUC, area under the receiver operating characteristic curve.

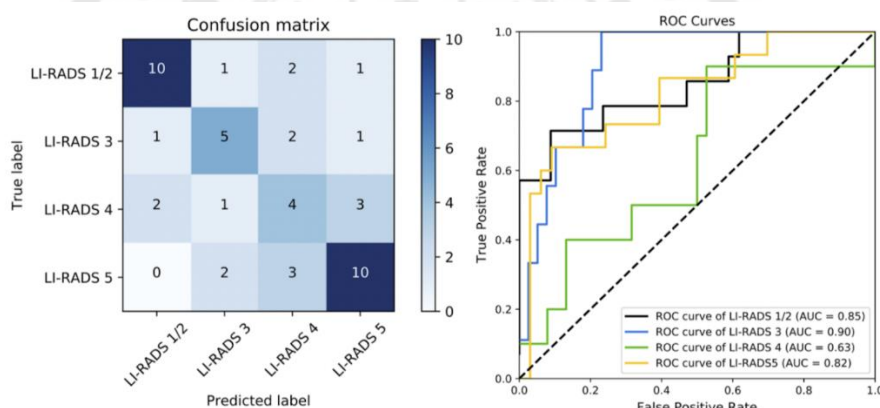
ภาพประกอบ 19 ผลลัพธ์การวัดประสิทธิภาพของแบบจำลอง โดยใช้ โครงสร้างเป็น Alexnet และใช้ Weight เป็น “Imagenet”

(4) บทความวิจัยเรื่อง Deep convolutional neural network applied to the liver imaging reporting and data system (LI-RADS) version 2014 category classification: a pilot study (Yamashita และคนอื่น ๆ, 2020)

งานวิจัยนี้ ต้องการจำแนกประเภทของ LR-1/2 (ไม่เป็นมะเร็ง), LR-3 (มีความเสี่ยงระดับปานกลางที่จะเป็นมะเร็ง), LR-4 (มีความเสี่ยงสูงที่จะเป็นมะเร็ง) และ LR-5 (เป็นมะเร็งแน่นอนจากภาพถ่ายทางรังสีวินิจฉัย) ของโรคมะเร็งตับ (Hepatocellular carcinoma) ซึ่งงานวิจัยนี้จะแตกต่างจากงานวิจัยที่ 3 ที่มุ่งเน้นไปที่การแก้ปัญหาแบบการจำแนกหลายประเภท (Multi-Classification) แทน และใช้ภาพถ่ายการตรวจวินิจฉัยโรคด้วยเครื่องเอกซเรย์คอมพิวเตอร์ (Computerized Tomography Scan: CT scan) ทั้งหมดจำนวน 163 รูป และภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI ทั้งหมดจำนวน 151 รูป ซึ่งมีการใช้ภาพถ่ายการตรวจทางรังสีวินิจฉัยแบบเดียวกับชุดข้อมูล C2 ของงานวิจัยที่ 3 ที่ได้เลือกใช้ข้อมูลภาพถ่ายเพียง 3 ระยะ (Triple-phase images) เป็นชุดข้อมูลในการสร้างแบบจำลองในครั้งนี้ (input image) ประกอบด้วย ระยะก่อนการฉีดสารทึบแสง (Pre-contrast phases), ระยะสารทึบแสงยังคงอยู่ในหลอดเลือดแดงไม่ได้อยู่ในอวัยวะและเนื้อเยื่ออื่น ๆ (Arterial phases) และระยะล้างสารทึบ

แสงออกจากช่องท้อง (Delayed phases) โดยในแต่ละ LI-RADS นั้น มีจำนวนข้อมูล ดังนี้ LR-1/2 มีจำนวนภาพถ่าย เท่ากับ 89, LR-3 มีจำนวนภาพถ่าย เท่ากับ 62, LR-4 มีจำนวนภาพถ่าย เท่ากับ 65 และ LR-5 มีจำนวนภาพถ่าย เท่ากับ 98

การสร้างแบบจำลองด้วยโครงข่ายประสาทเทียม (CNN) ใช้โครงสร้างเป็น VGG16 และใช้ Weight เป็น “Imagenet” มีการวัดประสิทธิภาพของแบบจำลอง (Model Evaluation) ด้วย Accuracy และ AUROCs (The area under the receiver operating characteristic) ได้ค่า Accuracy โดยรวมทุก LI-RADS (Overall) เท่ากับ 0.64 และ AUROCs ของ LR-1/2 (ไม่เป็นมะเร็ง), LR-3 (มีความเสี่ยงระดับปานกลางที่จะเป็นมะเร็ง), LR-4 (มีความเสี่ยงสูงที่จะเป็นมะเร็ง) และ LR-5 (เป็นมะเร็งแน่นอนจากภาพถ่ายทางรังสีวินิจฉัย) ในโรคมะเร็งตับ (Hepatocellular carcinoma) เท่ากับ 0.85, 0.90, 0.63 และ 0.82 ตามลำดับ ซึ่งพบว่า ในการจำแนก LR-4 (มีความเสี่ยงสูงที่จะเป็นมะเร็ง) มีประสิทธิภาพต่ำที่สุดจากการจำแนก LI-RADS ทั้งหมด ดังภาพประกอบ 16



ภาพประกอบ 20 การวัดประสิทธิภาพของแบบจำลองด้วยโครงข่ายประสาทเทียม (CNN)

ด้วยโครงสร้างเป็น VGG16 และใช้ Weight เป็น “Imagenet”

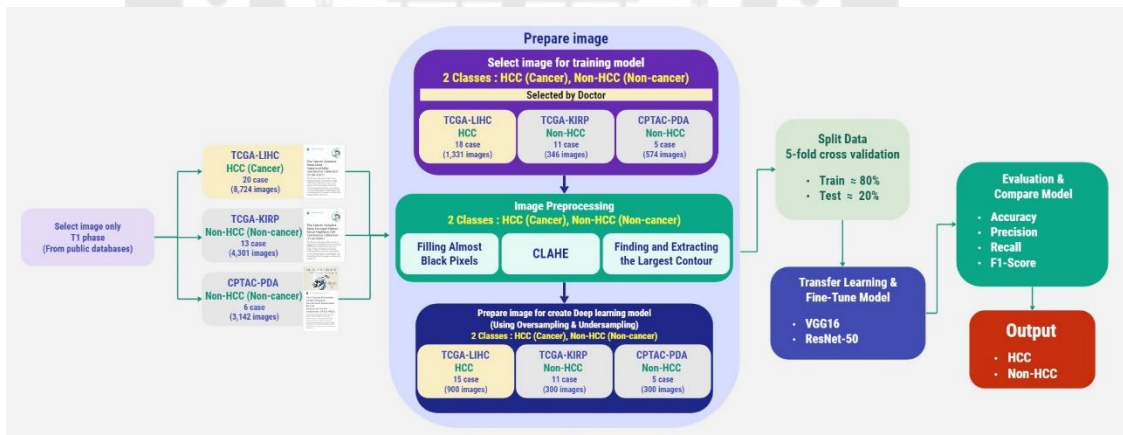
บทที่ 3

วิธีการดำเนินการวิจัย

ในการวิจัยในครั้งนี้ ผู้วิจัยได้ดำเนินการวิจัยตามขั้นตอนดังนี้

- 3.1 กระบวนการทำงานของแบบจำลอง (Workflows Process of Model)
- 3.2 การเก็บรวบรวมข้อมูล (Data Acquisition)
- 3.3 การระบุประเภทข้อมูล (Data Labeling)
- 3.4 การทำความสะอาดข้อมูล (Data Cleansing)
- 3.5 การเตรียมข้อมูล (Data Preprocessing)
- 3.6 การสำรวจข้อมูล (Exploratory Data Analysis: EDA)
- 3.7 การแบ่งข้อมูลสำหรับชุดฝึกสอนและชุดทดสอบ (Train-Test Split)
- 3.8 การสร้างแบบจำลองการจำแนก (Classification Model)
- 3.9 การวัดประสิทธิภาพของแบบจำลอง (Model Evaluation)

3.1 กระบวนการทำงานของแบบจำลอง (Workflows Process of Model)



ภาพประกอบ 21 แสดง Flowchart กระบวนการสร้างแบบจำลอง

จากรูป Flowchart กระบวนการทำงานวิจัยสร้างแบบจำลอง การพัฒนาแบบจำลอง เพื่อจำแนกการเป็นโรคมะเร็งตับและไม่ใช่โรคมะเร็งตับจากภาพถ่ายทางรังสีวินิจฉัย สามารถแบ่งออกเป็น 5 ขั้นตอน ดังนี้

ขั้นตอนที่หนึ่ง รวบรวมข้อมูลผู้ป่วยที่มีการสั่งการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) โดยเลือกใช้เฉพาะภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) ในระยะ T1 เท่านั้น พบว่า มีจำนวน

ผู้ป่วยทั้งหมด 39 ราย โดยเป็นชุดข้อมูลผู้ป่วยที่เป็นโรคมะเร็งตับชนิด Hepatocellular carcinoma คือ TCGA-LIHC จำนวนทั้งหมด 20 ราย และเป็นชุดข้อมูลผู้ป่วยที่ไม่เป็นโรคมะเร็งตับชนิด Hepatocellular carcinoma คือ TCGA-KIRP และ CPTAC-PDA จำนวนทั้งหมด 13 ราย และ 6 ราย ตามลำดับ หลังจากนั้นทำการวิเคราะห์และสำรวจข้อมูลในมุมมองต่างๆ (Exploratory Data Analysis : EDA)

ขั้นตอนที่สอง ได้ให้แพทย์เป็นผู้คัดเลือกภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) ในระยะ T1 เพื่อในการสร้างแบบจำลองการจำแนกการเป็นโรคมะเร็งตับและไม่เป็นโรคมะเร็งตับ โดยต้องมีพื้นที่รอยระดับมากกว่าร้อยละ 50 ในภาพถ่ายการตรวจนั้น เนื่องจากหากมีพื้นที่รอยระดับน้อยกว่าร้อยละ 50 ในภาพถ่ายการตรวจ รอยโรคที่ปรากฏในรอยระดับนั้นจะมีขนาดเล็กค่อนข้างมาก ทำให้วินิจฉัยโรคได้ยาก ขอบเขตงานวิจัยนี้จึงได้กำหนดเกณฑ์ในการคัดเลือกภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) ที่มีพื้นที่ระดับมากกว่าร้อยละ 50 เท่านั้น มาใช้ในการสร้างแบบจำลองการจำแนกประเภท

ถัดมาจะเป็นการมุ่งเน้นที่การเตรียมข้อมูลรูปภาพ โดยปรับขนาดของรูปภาพให้มีขนาดเท่า 224 x 224 pixels เพื่อให้เหมาะสมกับขนาดของรูปในขั้นตอนการสร้างแบบจำลอง (Input shape) เพราะในงานวิจัยนี้เราจะสร้างแบบจำลองโดยใช้โครงสร้างโครงข่ายประสาทเทียมคอนโวลูชันแบบจำลองของ VGG16 และ ResNet-50

จากนั้น แบ่งข้อมูลชุดฝึกสอน (Training set) และข้อมูลชุดทดสอบ (Validation set) โดยใช้หลักการของ K-Fold Cross Validation ซึ่งกำหนดค่า K เท่ากับ 5 (5 folds) และได้กำหนดสัดส่วนของข้อมูลชุดฝึกสอน (Training set) และข้อมูลชุดทดสอบ (Validation set) ภายในแต่ละส่วน (Fold) เป็นร้อยละ 80 และร้อยละ 20 โดยประมาณ ตามลำดับ ในขั้นตอนการแบ่งข้อมูล (Split data)

ทั้งนี้ ในกลุ่มข้อมูลผู้ป่วยที่เป็นโรคมะเร็งตับและไม่เป็นโรคมะเร็งตับนั้น มีสัดส่วนจำนวนรูปภาพที่ไม่เท่ากันในผู้ป่วยแต่ละราย จึงได้ออกแบบจัดกลุ่มของผู้ป่วยทั้ง 2 กลุ่มใหม่ โดยให้ในผู้ป่วยแต่ละรายชุดใหม่นี้ (ตามข้อมูลใน Column : data name) มีจำนวนรูปภาพเท่ากับ 60 รูปเสมอ พร้อมทั้งใช้เทคนิคการลดปริมาณข้อมูล (Undersampling) และเพิ่มปริมาณข้อมูล (Oversampling) เพื่อให้ข้อมูลรูปภาพที่นำไปใช้ในการสร้างแบบจำลองนั้นมีความสมดุลกันในทุก 2 ประเภทข้อมูล คือ กลุ่มข้อมูลผู้ป่วยที่เป็นโรคมะเร็งตับและไม่เป็นโรคมะเร็งตับ

หลังจากนั้น ทำการประมวลผลภาพ (Image Preprocessing) ของข้อมูลภาพการสอน (Training data) จากข้อมูลภาพเดิม โดยการปรับภาพพื้นหลังทั้งหมดให้เป็นสีดำสนิท และใช้

เทคนิคค้นหา Contour ที่ใหญ่ที่สุด (Find the largest contour) เพื่อทำการตัดเลือกรูปภาพนั้นมาจัดให้อยู่ตรงกลางในพื้นที่รูปสี่เหลี่ยมแล้วเติมขอบของภาพเพิ่มเติม (Padding) สุดท้ายใช้การปรับภาพด้วยเทคนิค CLAHE (Contrast Limited Adaptive Histogram Equalization) เพื่อเพิ่มเติมรายละเอียดของภาพให้ชัดเจนมากขึ้น

ขั้นตอนที่สาม มุ่งเน้นไปที่การสร้างแบบจำลองการจำแนกการเป็นโรคมะเร็งตับและไม่เป็นโรคมะเร็งตับจากภาพถ่ายทางรังสีวินิจฉัยการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) ซึ่งเป็นปัญหาแบบการจำแนกแบบ 2 ประเภท (Binary Classification) ของผู้ป่วยที่เป็นโรคมะเร็งตับและไม่เป็นโรคมะเร็งตับ โดยใช้เทคนิคการเรียนรู้เชิงลึกด้วยโครงข่ายประสาทเทียมคอนโวลูชัน (Deep Learning with Convolutional Neural Network) ด้วยโครงสร้างโครงข่ายประสาทเทียมคอนโวลูชันแบบจำลองแบบ VGG16 และแบบ ResNet-50 โดยแบ่งข้อมูลออกเป็นข้อมูลสอน (Training Data) ประมาณร้อยละ 80 และชุดข้อมูลทดสอบ (Testing Data) ประมาณร้อยละ 20 ซึ่งตรงนี้ได้ใช้การทดสอบแบบ 5- Fold Cross Validation

ขั้นตอนสุดท้าย จะเป็นการเปรียบเทียบประสิทธิภาพของโครงสร้างโครงข่ายประสาทเทียมคอนโวลูชัน แบบจำลองแบบ VGG16 และแบบ ResNet-50

3.2 การเก็บรวบรวมข้อมูล (Data Acquisition)

งานวิจัยนี้นำภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) มาศึกษาและจำแนกประเภท (Binary Classification) การเป็นโรคมะเร็งตับและไม่เป็นโรคมะเร็งตับประเภทการจำแนกประเภทแบบ 2 ประเภท (Binary Classification) ซึ่งทดลองกับข้อมูลภาพถ่ายทางรังสีวินิจฉัยจากฐานข้อมูลเปิด (Open database) ของ The Cancer Imaging Archive (TCIA) สำหรับผู้ป่วยที่เป็นโรคมะเร็งตับชนิด Hepatocellular carcinoma ใช้ชุดข้อมูล The Cancer Genome Atlas Liver Hepatocellular Carcinoma Collection (TCGA-LIHC) (Erickson, 2016) และสำหรับกรณีผู้ป่วยที่ไม่เป็นโรคมะเร็งตับชนิด Hepatocellular carcinoma จะใช้ชุดข้อมูลทั้งหมด 2 ชุด คือ The Cancer Genome Atlas Cervical Kidney Renal Papillary Cell Carcinoma Collection (TCGA-KIRP) (Linehan, 2016) ของข้อมูลผู้ป่วยที่เป็นโรคมะเร็งไต และอีกชุดข้อมูลหนึ่งจะเป็นข้อมูลผู้ป่วยที่เป็นโรคมะเร็งตับอ่อน The Clinical Proteomic Tumor Analysis Consortium Pancreatic Ductal Adenocarcinoma Collection (CPTAC-PDA) ((CPTAC), 2018)

โดยจำนวนข้อมูลผู้ป่วยทั้งหมดเริ่มแรกมีทั้งหมด 229 ราย เป็นชุดข้อมูลผู้ป่วยที่เป็นโรคมะเร็งตับชนิด Hepatocellular carcinoma มีจำนวนข้อมูลผู้ป่วยทั้งหมด 98 ราย ใช้ข้อมูลชุด TCGA-LIHC สำหรับผู้ป่วยที่ไม่เป็นโรคมะเร็งตับชนิด Hepatocellular carcinoma ใช้ชุดข้อมูล

TCGA-KIRP และ CPTAC-PDA มีจำนวนข้อมูลผู้ป่วยทั้งหมด 33 ราย และ 98 ราย ตามลำดับ หลังจากที่ได้ทำความสะอาดข้อมูลแล้ว และเลือกเฉพาะในระยะ T1 (Pulse sequence) เท่านั้น ตามวัตถุประสงค์ของงานวิจัย

ดังนั้น ชุดข้อมูลที่จะนำไปใช้ในการทำวิจัยครั้งนี้ คือ มีจำนวนข้อมูลผู้ป่วยทั้งหมด 39 ราย โดยแบ่งเป็นชุดข้อมูลผู้ป่วยที่เป็นโรคมะเร็งตับชนิด Hepatocellular carcinoma คือ TCGA-LIHC จำนวนทั้งหมด 20 ราย และเป็นชุดข้อมูลสำหรับผู้ป่วยที่ไม่เป็นโรคมะเร็งตับชนิด Hepatocellular carcinoma คือ TCGA-KIRP และ CPTAC-PDA จำนวนทั้งหมด 13 ราย และ 6 ราย ตามลำดับ

3.3 การระบุประเภทข้อมูล (Data Labeling)

งานวิจัยนี้นำภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) จากฐานข้อมูลเปิด (Open database) ของ The Cancer Imaging Archive (TCIA) โดยใช้ชุดข้อมูล The Cancer Genome Atlas Liver Hepatocellular Carcinoma Collection (TCGA-LIHC), The Cancer Genome Atlas Cervical Kidney Renal Papillary Cell Carcinoma Collection (TCGA-KIRP) และ The Clinical Proteomic Tumor Analysis Consortium Pancreatic Ductal Adenocarcinoma Collection (CPTAC-PDA) โดยคัดเลือกเฉพาะในระยะ T1 (Pulse sequence) เท่านั้น ในการระบุประเภทข้อมูลทางผู้วิจัยและแพทย์ที่ปรึกษา ได้มีเกณฑ์ในการคัดเลือกและระบุประเภทข้อมูลดังนี้

3.1 ในผู้ป่วยที่เป็นโรคมะเร็งตับชนิด Hepatocellular carcinoma ระบุค่าเป็น HCC และผู้ป่วยที่ไม่เป็นโรคมะเร็งตับชนิด Hepatocellular carcinoma ระบุค่าเป็น Non HCC

3.2 แพทย์จะทำการคัดเลือกภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) ที่มีพื้นที่อวัยวะตับในภาพมากกว่าร้อยละ 50 เนื่องจากหากมีพื้นที่อวัยวะตับน้อยกว่าร้อยละ 50 ในภาพถ่ายการตรวจนั้น ๆ รอยโรคที่ปรากฏในอวัยวะตับจะมีขนาดเล็กค่อนข้างมาก ทำให้สามารถวินิจฉัยโรคได้ยากและไม่ได้อยู่ในขอบเขตงานวิจัยนี้ จึงได้มีการกำหนดเกณฑ์ในการคัดเลือกภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) ให้ที่มีพื้นที่ตับมากกว่าร้อยละ 50 เท่านั้น จึงจะสามารถนำมาใช้ในการสร้างแบบจำลองการจำแนกผู้ป่วยที่เป็นโรคมะเร็งตับและไม่เป็นโรคมะเร็งตับได้

3.4 การทำความสะอาดข้อมูล (Data Cleansing)

จากฐานข้อมูลเปิด (Open database) ของ The Cancer Imaging Archive (TCIA) ใช้ชุดข้อมูลจำนวนทั้งหมด 3 ชุด ประกอบด้วย The Cancer Genome Atlas Liver Hepatocellular Carcinoma Collection (TCGA-LIHC), The Cancer Genome Atlas Cervical Kidney Renal Papillary Cell Carcinoma Collection (TCGA-KIRP) และสุดท้าย คือ The Clinical Proteomic

Tumor Analysis Consortium Pancreatic Ductal Adenocarcinoma Collection (CPTAC-PDA) คัดเลือกเฉพาะในระยะ T1 (Pulse sequence) มาแล้วเท่านั้น ในขั้นตอนถัดมาแพทย์จะคัดเลือกเฉพาะผู้ป่วยที่ไม่เคยได้รับการรักษามาก่อน ตัวอย่างเช่น ต้องไม่เคยได้รับการผ่าตัดมาก่อน และนอกจากนี้ภาพถ่ายจากตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) นั้น ต้องเป็นภาพที่สามารถนำไปใช้ในการสร้างแบบจำลองได้ ไม่เป็นภาพที่มีดจจนเกินไป หรือคุณภาพของภาพถ่ายไม่ดีจนไม่สามารถใช้งานได้ และสุดท้ายมีการกำหนดเกณฑ์การคัดเลือกภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) ให้ที่มีพื้นที่ตับมากกว่าร้อยละ 50 เท่านั้น จึงจะสามารถนำมาใช้ในการสร้างแบบจำลองการจำแนกผู้ป่วยที่เป็นโรคมะเร็งตับและไม่เป็นโรคมะเร็งตับได้

3.5 การเตรียมข้อมูล (Data Preprocessing)

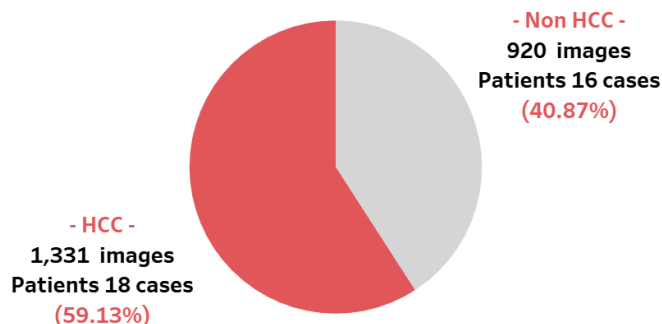
ก่อนจะนำข้อมูลไปใช้ในการสร้างแบบจำลองนั้น ต้องผ่านขั้นตอนการประมวลผลภาพ (Image Preprocessing) เพื่อปรับแต่งภาพก่อนที่จะนำไปใช้ในการวิเคราะห์หรือการประมวลผลอื่น ๆ ของข้อมูลภาพการสอน (Training data) จากข้อมูลภาพเดิม โดยได้มีการปรับแต่งภาพพื้นหลังทั้งหมดให้เป็นสีดำสนิททั้งหมด พร้อมทั้งใช้เทคนิคการค้นหา Contour ที่ใหญ่ที่สุด (Find the largest contour) เพื่อเป็นการเลือกบริเวณในรูปภาพที่เราสนใจ ถัดมาทำการตัดเลือกรูปภาพนั้น (Cropping) มาจัดให้อยู่ตรงกลางในพื้นที่รูปสี่เหลี่ยม แล้วเติมขอบของภาพเพิ่มเติม (Padding) และสุดท้ายใช้การปรับภาพด้วยเทคนิค CLAHE (Contrast Limited Adaptive Histogram Equalization) เพื่อเพิ่มเติมรายละเอียดของภาพให้ชัดเจนมากขึ้น

3.6 การสำรวจข้อมูล (Exploratory Data Analysis: EDA)

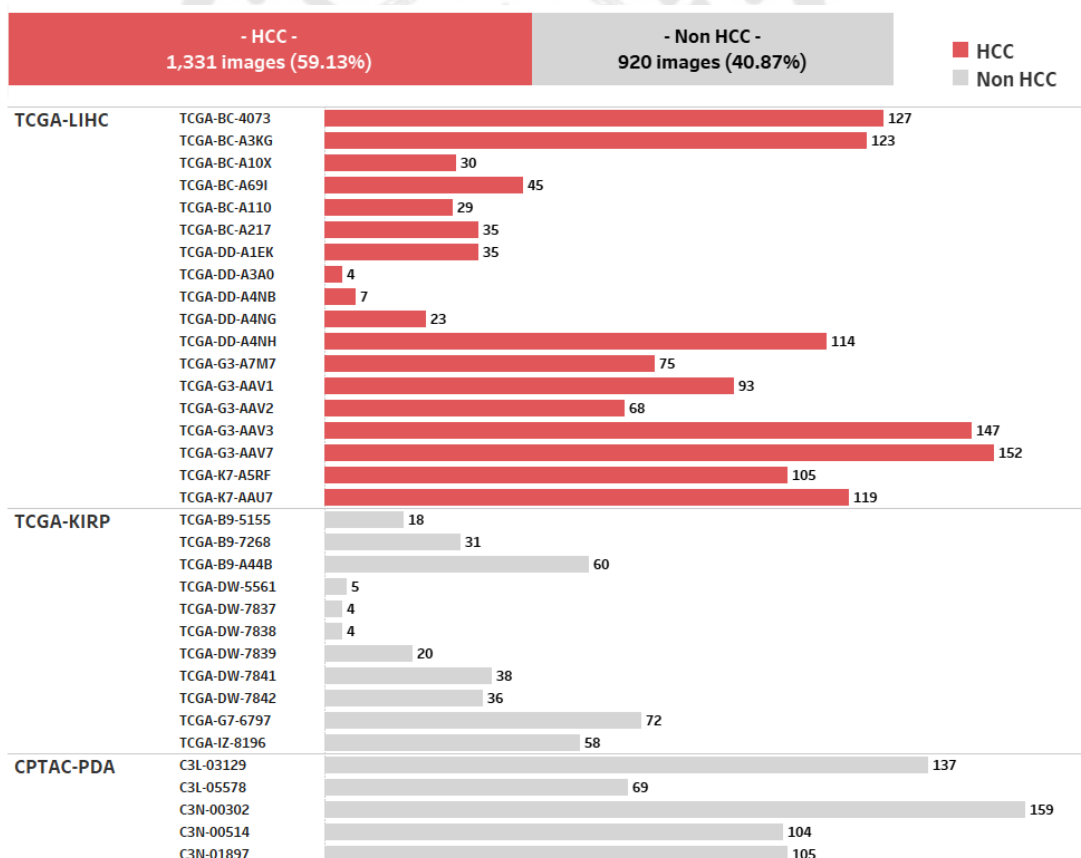
หลังจากเลือกรูปภาพการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) เฉพาะในระยะ T1 (Pulse sequence) มาแล้วตามวัตถุประสงค์ของงานวิจัย มีจำนวนข้อมูลผู้ป่วยทั้งหมด 39 ราย โดยแบ่งเป็นชุดข้อมูลผู้ป่วยที่เป็นโรคมะเร็งตับชนิด Hepatocellular carcinoma คือ TCGA-LIHC จำนวนทั้งหมด 20 ราย และเป็นชุดข้อมูลสำหรับผู้ป่วยที่ไม่เป็นโรคมะเร็งตับชนิด Hepatocellular carcinoma คือ TCGA-KIRP และ CPTAC-PDA จำนวนทั้งหมด 13 ราย และ 6 ราย ตามลำดับ

เมื่อผ่านขั้นตอนการคัดเลือกรูปภาพจากแพทย์แล้ว มีจำนวนผู้ป่วยทั้งหมด 34 ราย โดยมีจำนวนภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) ทั้งหมด 2,251 รูป แบ่งเป็นชุดข้อมูลผู้ป่วยที่เป็นโรคมะเร็งตับชนิด Hepatocellular carcinoma คือ TCGA-LIHC จำนวนทั้งหมด 18 ราย จำนวน 1,331 รูป และเป็นชุดข้อมูลสำหรับผู้ป่วยที่ไม่เป็นโรคมะเร็งตับชนิด

Hepatocellular carcinoma คือ TCGA-KIRP และ CPTAC-PDA จำนวนทั้งหมด 11 ราย จำนวน 346 รูป และ 5 ราย จำนวน 574 รูป ตามลำดับ ดังภาพประกอบ 22



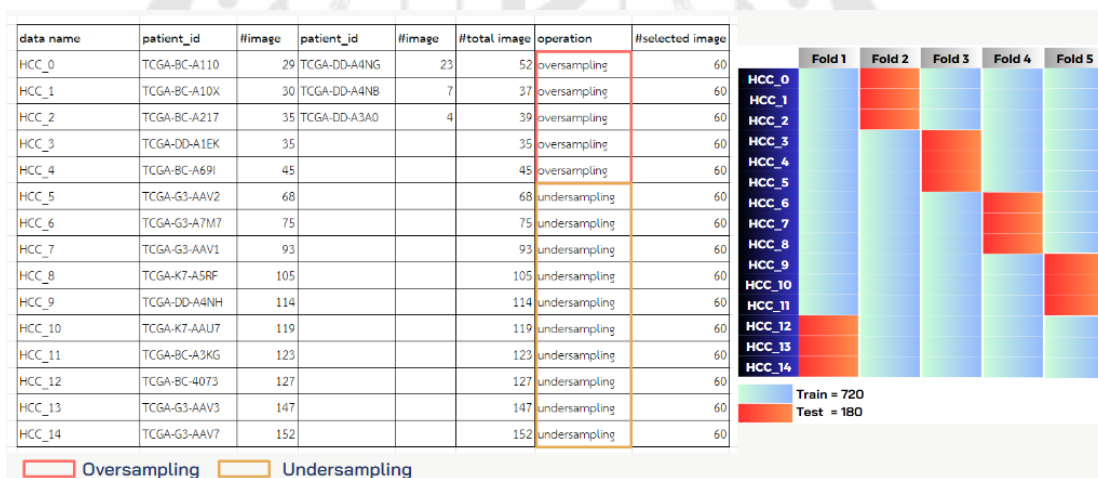
ภาพประกอบ 22 จำนวนชุดข้อมูลผู้ป่วยที่เป็นโรคมะเร็งตับ (Hepatocellular carcinoma : HCC) และผู้ป่วยที่ไม่เป็นโรคมะเร็งตับ (Non Hepatocellular carcinoma : Non HCC)



ภาพประกอบ 23 รายละเอียดผู้ป่วยที่เป็นโรคมะเร็งตับ (Hepatocellular carcinoma : HCC) และผู้ป่วยที่ไม่เป็นโรคมะเร็งตับ (Non Hepatocellular carcinoma : Non HCC)

3.7 การแบ่งข้อมูลสำหรับชุดฝึกสอนและชุดทดสอบ (Train-Test Split)

เมื่อผ่านกระบวนการที่แพทย์คัดเลือกภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) ในระยะ T1 มาเรียบร้อยแล้ว พร้อมทั้งเป็นภาพถ่ายที่มีพื้นที่อวัยวะตับมากกว่าร้อยละ 50 ขั้นตอนการแบ่งข้อมูลชุดฝึกสอน (Training set) และข้อมูลชุดทดสอบ (Validation set) จะใช้หลักการของ K-Fold Cross Validation กำหนดค่า K เท่ากับ 5 (5 folds) และให้สัดส่วนของข้อมูลชุดฝึกสอน (Training set) และข้อมูลชุดทดสอบ (Validation set) ภายในแต่ละส่วน (Fold) เป็นร้อยละ 80 และร้อยละ 20 โดยประมาณ ตามลำดับ เนื่องจากผู้ป่วยที่เป็นโรคมะเร็งตับและไม่เป็นโรคมะเร็งตับนั้น มีสัดส่วนจำนวนรูปภาพที่ไม่เท่ากันในผู้ป่วยแต่ละราย จึงได้ออกแบบการจัดกลุ่มผู้ป่วยทั้ง 2 กลุ่มใหม่ โดยผู้ป่วยแต่ละรายชุดใหม่นี้ ด้วยเทคนิคการลดปริมาณข้อมูล (Undersampling) และเพิ่มปริมาณข้อมูล (Oversampling) เพื่อเป็นการปรับมีจำนวนรูปภาพเท่ากับ 60 รูปเสมอ เกิดความสมดุลกันของข้อมูลรูปภาพที่นำไปใช้ในการสร้างแบบจำลองทั้งในกลุ่มข้อมูลผู้ป่วยที่เป็นโรคมะเร็งตับและไม่เป็นโรคมะเร็งตับ ดังภาพประกอบ 22 และ 23



ภาพประกอบ 24 การจัดกลุ่มผู้ป่วยแต่ละรายใหม่ในกลุ่มผู้ป่วยที่เป็นโรคมะเร็งตับ

(Hepatocellular carcinoma : HCC)

data name	patient_id	#image	patient_id	#image	patient_id	#image	#total image	operation	#selected image
non_HCC0	C3L-03129	137						undersampling	60
non_HCC1	C3L-05578	69						undersampling	60
non_HCC2	C3N-00302	159						undersampling	60
non_HCC3	C3N-00514	104						undersampling	60
non_HCC4	C3N-01897	105						undersampling	60
non_HCC5	TCGA-DW-7842	36->28	TCGA-B9-7268	31->27	TCGA-DW-5561	5		undersampling	60
non_HCC6	TCGA-DW-7841	38->36	TCGA-DW-783	20	TCGA-DW-7838	4		undersampling	60
non_HCC7	TCGA-IZ-8196	58->38	TCGA-B9-5155	18	TCGA-DW-7837	4		undersampling	60
non_HCC8	TCGA-B9-A448	60						select all	60
non_HCC9	TCGA-G7-6797	72						undersampling	60

Note: 36->28 means select 28 from 36.

Legend: Undersampling Select all

ภาพประกอบ 25 การจัดกลุ่มผู้ป่วยแต่ละรายใหม่ในกลุ่มผู้ป่วยที่ไม่เป็นโรคมะเร็งตับ
(Non Hepatocellular carcinoma : Non HCC)

3.8 การสร้างแบบจำลองการจำแนก (Classification Model)

ก่อนนำภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) มาใช้ในการฝึกสอนแบบจำลองนั้น จะต้องผ่านกระบวนการปรับข้อมูลภาพ (Image Augmentation) ก่อน เพื่อสร้างข้อมูลภาพใหม่ จากข้อมูลภาพที่มีอยู่โดยไม่เปลี่ยนแปลงชนิดข้อมูลของภาพ (เช่น ภาพของผู้ป่วยที่เป็นโรคมะเร็ง ตับก็ยังคงเป็นผู้ป่วยโรคมะเร็งตับเหมือนเดิม) แต่ทำให้แบบจำลองสามารถเรียนรู้และทำนายได้ดี ยิ่งขึ้นในสถานการณ์ต่าง ๆ โดยได้มีการปรับข้อมูลรูปภาพด้วยเทคนิค ตามตาราง 1

ตาราง 1 ตารางการปรับข้อมูลรูปภาพด้วยเทคนิคต่าง ๆ ก่อนนำไปใช้ในการสร้างแบบจำลอง

เทคนิค	ค่าที่ใช้
Rotation_range	มีการหมุนภาพได้ในช่วงไม่เกิน 25 องศา
Width_shift_range	มีการเลื่อนภาพซ้าย-ขวาได้ไม่เกิน 10% ของความกว้างของภาพ
Height_shift_range	มีการเลื่อนภาพบน-ล่างได้ไม่เกิน 10% ของความสูงของภาพ
Shear_range	มีการเบียดภาพได้ในช่วงไม่เกิน 10 องศา
Zoom_range	มีการซูมเข้าหรือซูมออกภาพได้ไม่เกิน 10%
Horizontal_flip	มีการพลิกภาพตามแนวนอน (ซ้าย-ขวา) แต่เป็นทิศทางตรงข้าม
Brightness_range	มีการปรับความสว่างของภาพในช่วง 0.5 - 1.0
Fill_mode	เติมพื้นที่ว่างในภาพหลังจากการเปลี่ยนแปลง โดยใช้วิธี 'nearest' ซึ่งจะใช้ค่าพิกเซลใกล้เคียงที่มีอยู่ในข้อมูลภาพเดิม

การสร้างแบบจำลองการเรียนรู้เชิงลึกด้วยโครงข่ายประสาทเทียมคอนโวลูชัน (Deep Learning with Convolutional Neural Network) การจำแนกประเภท (Binary Classification) ของการเป็นโรคมะเร็งตับและไม่เป็นโรคมะเร็งตับจากภาพทางรังสีวินิจฉัยนั้น มีโครงสร้างโครงข่ายประสาทเทียมคอนโวลูชันแบบจำลองที่เราสนใจ ทั้งหมด 2 ประเภท คือ VGG16 และ ResNet-50 โดยได้มีการปรับโครงสร้างแบบจำลอง และค่าพารามิเตอร์ (Parameters) ต่าง ๆ ในการฝึกสอนแบบจำลอง ออกเป็นทั้งหมด 4 ชุดการทดลอง (Version) ดังนี้

3.8.1 การสร้างแบบจำลองด้วยอัลกอริทึม VGG16 Version 1

การสร้างแบบจำลอง VGG16 ในชุดการทดลองลำดับที่ 1 (Version 1) นี้ ใช้การเลือกแบบจำลองต้นแบบ (Pre-trained Model) และทำการปรับชั้นเชื่อมต่อแบบสมบูรณ์ (Fully-Connected Layer) ชั้นสุดท้ายให้เหมาะกับงานการจำแนกประเภทแบบ 2 ชนิด (Binary classification) โดยปรับให้มีจำนวน Neuron = 1 และเลือกฟังก์ชันการกระตุ้น (Activation function) เป็น Sigmoid พร้อมทั้งมีการกำหนดค่า Hyperparameter ต่าง ๆ ดังตาราง 2

ตาราง 2 ตารางแสดงค่าพารามิเตอร์ต่าง ๆ ที่ใช้ในการสร้างแบบจำลอง VGG16 Version 1

Layer (type)	Output Shape
input_1 (Input Layer)	(None, 224, 224, 3)
VGG16 model with Pre-train layer	
Flatten Layer	(None, 25088)
Fully connected Layer 1	(None, 4096)
Fully connected Layer 2	(None, 4096)
Dense Layer	(None, 1)
Activation Function	Sigmoid
Optimizers	Adam
Learning rate	0.0001
Epoch	30
Batch size	64

3.8.2 การสร้างแบบจำลองด้วยอัลกอริทึม VGG16 Version 2

การสร้างแบบจำลอง VGG16 ในชุดการทดลองลำดับที่ 2 (Version 2) นี้ ใช้การเลือกแบบจำลองต้นแบบ (Pre-trained Model) เหมือนเดิม ทำการปรับชั้นเชื่อมต่อแบบสมบูรณ์

(Fully-Connected Layer) ขั้นสุดท้ายให้เหมาะกับการจำแนกประเภทแบบ 2 ชนิด (Binary classification) โดยปรับให้มีจำนวน Neuron = 1 และเลือกฟังก์ชันการกระตุ้น (Activation function) เป็น Sigmoid แต่มีปรับลดจำนวน Neural network ลดลงจากเดิมที่เคยมีจำนวนเท่ากับ 4,096 ในชั้นเชื่อมโยงแบบสมบูรณ์ชั้นที่ 1 (Fully connected Layer 1) และชั้นที่ 2 (Fully connected Layer 2) ปรับลดลงเหลือ 256 และ 128 ตามลำดับจึงพอสรุปการกำหนดค่า Hyperparameter ต่าง ๆ ดังตาราง 3

ตาราง 3 ตารางแสดงค่าพารามิเตอร์ต่าง ๆ ที่ใช้ในการสร้างแบบจำลอง VGG16 Version 2

Layer (type)	Output Shape
input_1 (Input Layer)	(None, 224, 224, 3)
VGG16 model with Pre-train layer	
Flatten Layer	(None, 25088)
Fully connected Layer 1	(None, 256)
Fully connected Layer 2	(None, 128)
Dense Layer	(None, 1)
Activation Function	Sigmoid
Optimizers	Adam
Learning rate	0.0001
Epoch	30
Batch size	64

3.8.3 การสร้างแบบจำลองด้วยอัลกอริทึม VGG16 Version 3

การสร้างแบบจำลอง VGG16 ในชุดการทดลองลำดับที่ 3 (Version 3) นี้ เลียนแบบโครงสร้างมาจาก VGG16 Version 2 แต่มีการปรับเปลี่ยนตรงชั้นเชื่อมโยงแบบสมบูรณ์ชั้นที่ 1 (Fully connected Layer 1) และชั้นที่ 2 (Fully connected Layer 2) ให้มีจำนวน Neural network คงเดิม คือ 4,096 แต่กำหนดให้เป็นการฝึกสอนใหม่โดยเฉพาะเลยในทั้ง 2 ชั้นนี้ โดยการตั้งค่าเชื่อมโยงแบบสมบูรณ์ชั้นที่ 1 และชั้นที่ 2 นี้เป็น trainable = True แทน จึงพอสรุปการกำหนดค่า Hyperparameter ต่าง ๆ ดังตาราง 4

ตาราง 4 ตารางแสดงค่าพารามิเตอร์ต่าง ๆ ที่ใช้ในการสร้างแบบจำลอง VGG16 Version 3

Layer (type)	Output Shape
input_1 (Input Layer)	(None, 224, 224, 3)
VGG16 model with Pre-train layer	
Flatten Layer	(None, 25088)
Fully connected Layer 1 (Trainable = True)	(None, 4096)
Fully connected Layer 2 (Trainable = True)	(None, 4096)
Dense Layer	(None, 1)
Activation Function	Sigmoid
Optimizers	Adam
Learning rate	0.0001
Epoch	30
Batch size	64

3.8.4 การสร้างแบบจำลองด้วยอัลกอริทึม VGG16 Version 4

การสร้างแบบจำลอง VGG16 ในชุดการทดลองลำดับที่ 4 (Version 4) นี้ ใช้การเลือกแบบจำลองต้นแบบ (Pre-trained Model) แต่มีการนำชั้น Flatten Layer ออก แล้วใส่ชั้น GlobalAveragePooling2D แทน ปรับชั้นเชื่อมต่อแบบสมบูรณ์ (Fully-Connected Layer) ให้มี Neuron เท่ากับ 256 เพิ่มขึ้น BatchNormalization เพิ่มขึ้นฟังก์ชันการกระตุ้น (Activation function) ที่ตั้งค่าเป็น ReLU เพิ่มขึ้น Dropout มีค่าเท่ากับ 0.5 สุดท้ายชั้นเชื่อมต่อแบบสมบูรณ์ (Fully-Connected Layer) ปรับให้เหมาะกับการจำแนกประเภทแบบ 2 ชนิด (Binary classification) โดยให้มีจำนวน Neuron = 1 และเลือกฟังก์ชันการกระตุ้น (Activation function) เป็น Sigmoid ดังเดิม จึงสามารถสรุปการกำหนดค่า Hyperparameter ต่าง ๆ ได้ดังตาราง 5

ตาราง 5 ตารางแสดงค่าพารามิเตอร์ต่าง ๆ ที่ใช้ในการสร้างแบบจำลอง VGG16 Version 4

Layer (type)	Output Shape
input_1 (Input Layer)	(None, 224, 224, 3)
VGG16 model with Pre-train layer	

ตาราง 5 (ต่อ)

Layer (type)	Output Shape
GlobalAveragePooling2D	(None, 512)
Fully connected Layer 1	(None, 256)
BatchNormalization	(None, 256)
Activation Function	ReLU
Dropout (0.5)	(None, 256)
Dense Layer	(None, 1)
Activation Function	Sigmoid
Optimizers	Adam
Learning rate	0.0001
Epoch	30
Batch size	64

3.8.5 การสร้างแบบจำลองด้วยอัลกอริทึม ResNet-50 Version 1

การสร้างแบบจำลอง ResNet-50 ในชุดการทดลองลำดับที่ 1 (Version 1) นี้ ใช้การเลือกแบบจำลองต้นแบบ (Pre-trained Model) และทำการปรับชั้นเชื่อมต่อแบบสมบูรณ์ (Fully-Connected Layer) ชั้นสุดท้ายให้เหมาะกับการจำแนกประเภทแบบ 2 ชนิด (Binary classification) โดยปรับให้มีจำนวน Neuron = 1 และเลือกฟังก์ชันการกระตุ้น (Activation function) เป็น Sigmoid พร้อมทั้งมีการกำหนดค่า Hyperparameter ต่าง ๆ ดังตาราง 6

ตาราง 6 ตารางแสดงค่าพารามิเตอร์ต่าง ๆ ที่ใช้ในการสร้างแบบจำลอง ResNet50 Version 1

Layer (type)	Output Shape
input_1 (Input Layer)	(None, 224, 224, 3)
ResNet50 model with Pre-train layer	
Dense Layer	(None, 1)
Activation Function	Sigmoid
Optimizers	Adam
Learning rate	0.0001
Epoch	30
Batch size	64

3.8.6 การสร้างแบบจำลองด้วยอัลกอริทึม ResNet-50 Version 2

การสร้างแบบจำลอง ResNet-50 ในชุดการทดลองลำดับที่ 2 (Version 2) นี้ ใช้การเลือกแบบจำลองต้นแบบ (Pre-trained Model) เหมือนเดิม ทำการปรับชั้นเชื่อมโยงแบบสมบูรณ์ (Fully-Connected Layer) ชั้นสุดท้ายให้เหมาะกับการจำแนกประเภทแบบ 2 ชนิด (Binary classification) โดยปรับให้มีจำนวน Neuron = 1 และเลือกฟังก์ชันการกระตุ้น (Activation function) เป็น Sigmoid มีการเพิ่มชั้นเชื่อมโยงแบบสมบูรณ์ชั้นที่ 1 (Fully connected Layer 1) และชั้นที่ 2 (Fully connected Layer 2) โดยมีจำนวน Neuron เท่ากับ 256 และ 128 ตามลำดับ จึงพอสรุปการกำหนดค่า Hyperparameter ต่าง ๆ ดังตาราง 7

ตาราง 7 ตารางแสดงค่าพารามิเตอร์ต่าง ๆ ที่ใช้ในการสร้างแบบจำลอง ResNet50 Version 2

Layer (type)	Output Shape
input_1 (Input Layer)	(None, 224, 224, 3)
ResNet50 model with Pre-train layer	
Flatten Layer	(None, 100352)
Fully connected Layer 1	(None, 256)
Fully connected Layer 2	(None, 128)
Dense Layer	(None, 1)
Activation Function	Sigmoid
Optimizers	Adam
Learning rate	0.0001
Epoch	30
Batch size	64

3.8.7 การสร้างแบบจำลองด้วยอัลกอริทึม ResNet-50 Version 3

แบบจำลอง ResNet-50 ชุดการทดลองลำดับที่ 3 (Version 3) คล้ายโครงสร้าง ResNet50 Version 2 แต่มีการปรับชั้นเชื่อมโยงแบบสมบูรณ์ชั้นที่ 1 (Fully connected Layer 1) และชั้นที่ 2 (Fully connected Layer 2) ให้มีจำนวน Neuron เท่ากับ 512 และ 256 แต่กำหนดให้เป็นการฝึกสอนใหม่โดยเฉพาะในทั้ง 2 ชั้นนี้ โดยตั้งค่าเชื่อมโยงแบบสมบูรณ์ชั้นที่ 1 และชั้นที่ 2 นี้ เป็น trainable = True และนำ Flatten Layer ออก จึงพอสรุปการกำหนดค่าต่าง ๆ ดังตาราง 8

ตาราง 8 ตารางแสดงค่าพารามิเตอร์ต่าง ๆ ที่ใช้ในการสร้างแบบจำลอง ResNet50 Version 3

Layer (type)	Output Shape
input_1 (Input Layer)	(None, 224, 224, 3)
ResNet50 model with Pre-train layer	
Fully connected Layer 1 (Trainable = True)	(None, 512)
Fully connected Layer 2 (Trainable = True)	(None, 256)
Dense Layer	(None, 1)
Activation Function	Sigmoid
Optimizers	Adam
Learning rate	0.0001
Epoch	30
Batch size	64

3.8.8 การสร้างแบบจำลองด้วยอัลกอริทึม ResNet-50 Version 4

การสร้างแบบจำลอง ResNet-50 ในชุดการทดลองลำดับที่ 4 (Version 4) นี้ใช้การเลือกแบบจำลองต้นแบบ (Pre-trained Model) มีใส่ชั้นGlobalAveragePooling2D ปรับชั้นเชื่อมต่อแบบสมบูรณ์ (Fully-Connected Layer) ให้มี Neuron เท่ากับ 256 เพิ่มชั้นBatchNormalization เพิ่มชั้นฟังก์ชันการกระตุ้น (Activation function) ที่ตั้งค่าเป็น ReLU เพิ่มชั้นDropout มีค่าเท่ากับ 0.5 และสุดท้ายชั้นเชื่อมต่อแบบสมบูรณ์ (Fully-Connected Layer) ชั้นสุดท้ายให้เหมาะกับการจำแนกประเภทแบบ 2 ชนิด (Binary classification) โดยปรับให้มีจำนวน Neuron = 1 และเลือกฟังก์ชันการกระตุ้น (Activation function) เป็น Sigmoid ดังเดิมจึงสามารถสรุปการกำหนดค่า Hyperparameter ต่าง ๆ ได้ดังตาราง 9

ตาราง 9 ตารางแสดงค่าพารามิเตอร์ต่าง ๆ ที่ใช้ในการสร้างแบบจำลอง ResNet50 Version 4

Layer (type)	Output Shape
input_1 (Input Layer)	(None, 224, 224, 3)
ResNet50 model with Pre-train layer	
GlobalAveragePooling2D	(None, 2048)

ตาราง 9 (ต่อ)

Layer (type)	Output Shape
Fully connected Layer 1	(None, 256)
BatchNormalization	(None, 256)
Activation Function	ReLU
Dropout (0.5)	(None, 256)
Dense Layer	(None, 1)
Activation Function	Sigmoid
Optimizers	Adam
Learning rate	0.0001
Epoch	30
Batch size	64

3.9 การวัดประสิทธิภาพของแบบจำลอง (Model Evaluation)

นำแบบจำลองที่ได้ในแต่ละชุดการทดลอง (Version) ของโครงสร้างโครงข่ายประสาทเทียมคอนโวลูชันแบบจำลองทั้งหมด 2 ประเภท คือ VGG16 และ ResNet-50 มาทดสอบกับชุดข้อมูลทดสอบ จากนั้นนำแบบจำลองทั้งหมดมาเปรียบเทียบประสิทธิภาพโดยพิจารณาด้วย Confusion Matrix ซึ่งมีค่าทั้งหมด 4 ค่า ดังนี้ Accuracy, Precision, Recall และ F1-score

บทที่ 4

ผลการดำเนินงานวิจัย

งานวิจัยนี้เป็นการสร้างแบบจำลองการจำแนกประเภท (Binary Classification) การเป็นโรคมะเร็งตับและไม่เป็นโรคมะเร็งตับจากภาพทางรังสีวินิจฉัย โดยใช้เทคนิคการเรียนรู้เชิงลึกด้วยโครงข่ายประสาทเทียมคอนโวลูชัน และเพื่อให้สามารถเลือกโครงสร้างโครงข่ายประสาทเทียมคอนโวลูชันแบบจำลองแบบ VGG16 และแบบ ResNet-50 ที่มีประสิทธิภาพในการจำแนกประเภทได้ โดยผลการทดลองของงานวิจัยจะแบ่งออกเป็น 2 ส่วน ประกอบด้วย

4.1 ผลการประเมินประสิทธิภาพแบบจำลองด้วยอัลกอริทึม VGG16

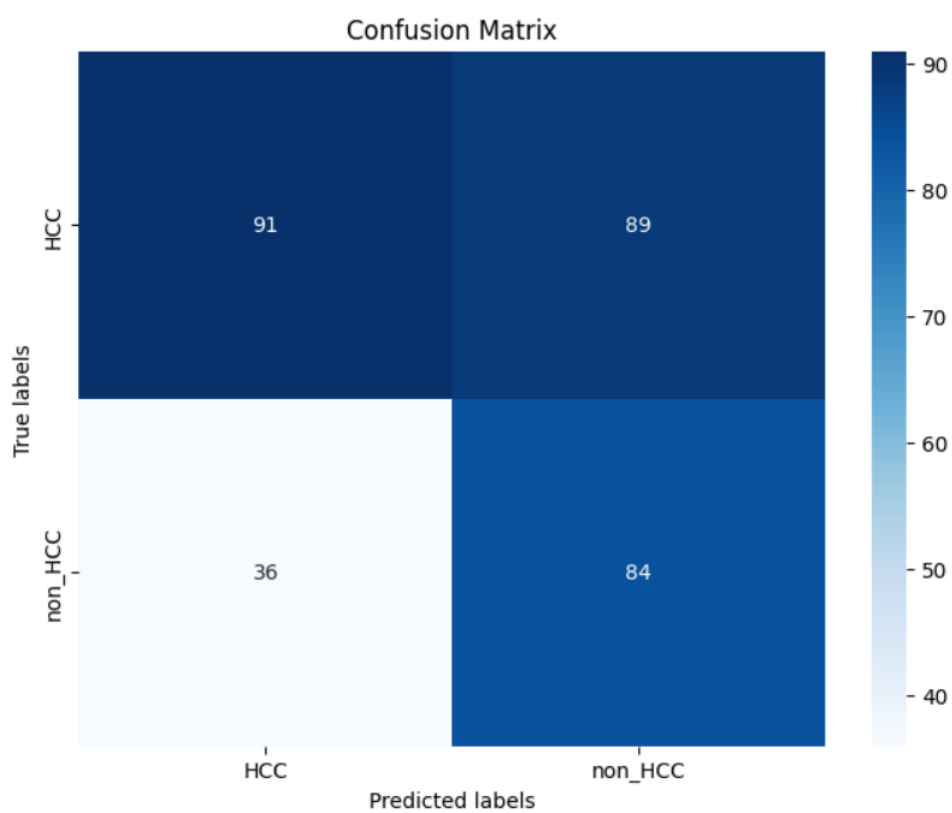
จากทั้งหมด 4 ชุดผลการทดลองด้วยอัลกอริทึม VGG16 นั้น พบว่า เมื่อพิจารณาประสิทธิภาพของแบบจำลองด้วยเมตริกซ์ Accuracy ที่มุ่งเน้นเรื่องอัตราส่วนของตัวอย่างที่แบบจำลองทำนายถูกต้องทั้งหมด เรียกว่า Accuracy จากมากไปน้อยจะได้ VGG16 Version 2 ที่ 0.63 VGG16 Version 3 ที่ 0.59 VGG16 Version 4 ที่ 0.58 และ VGG16 Version 1 ที่ 0.54 ดังตาราง 10

ตาราง 10 ตารางแสดงประสิทธิภาพการจำแนกของแบบจำลอง VGG16 แต่ละ Version

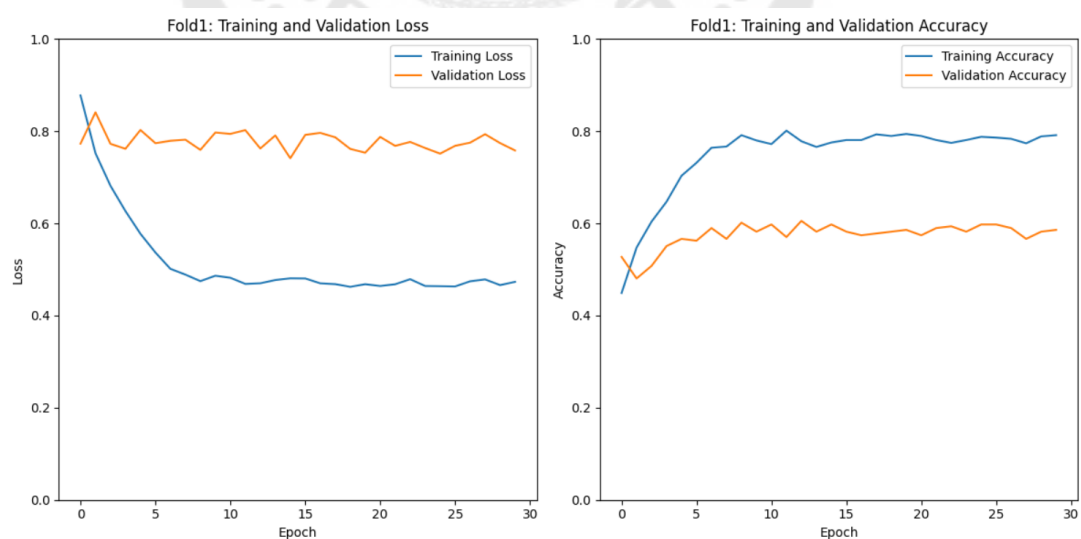
VGG16 Version	Accuracy	Precision	Recall	F1-score
1	0.54	0.53	0.53	0.53
2	0.63	0.61	0.59	0.58
3	0.59	0.59	0.59	0.57
4	0.58	0.59	0.59	0.57

เมื่อเปรียบเทียบผลการทดลอง VGG16 ในแต่ละ Version และตามข้อมูลแต่ละส่วนที่แบ่งการฝึกสอนและทดสอบข้อมูล (Fold) ตามการวัดประสิทธิภาพของแบบจำลองด้วย Confusion Matrix คือ ตารางสำคัญในการวัดความสามารถของ machine learning ในการแก้ปัญหา classification พร้อมทั้งดูกราฟการวิเคราะห์ผลการฝึกสอนแบบจำลอง (Learning Curve) โดยการวัดผลการเรียนรู้ของโมเดลจากข้อมูลฝึกสอน (Training Dataset) และข้อมูลทดสอบ (Validation Dataset) จะได้ผลลัพธ์ ดังนี้

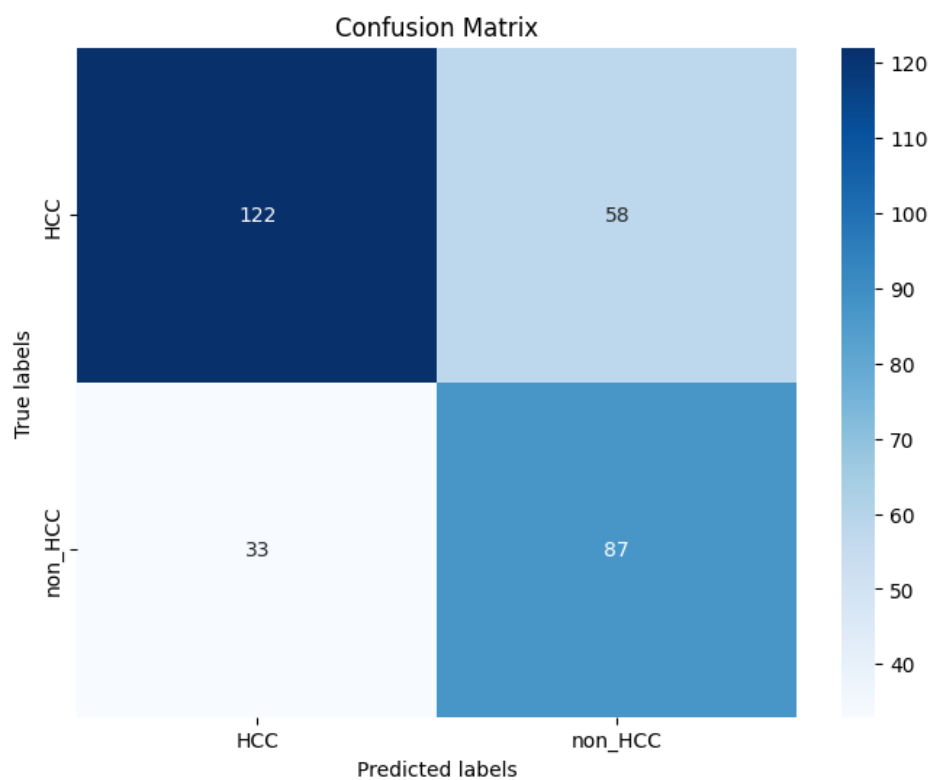
4.1.1 Confusion Matrix และ Learning Curve แบบจำลอง VGG16 Version 1



ภาพประกอบ 26 Confusion Matrix แบบจำลอง VGG16 Version 1 fold 1



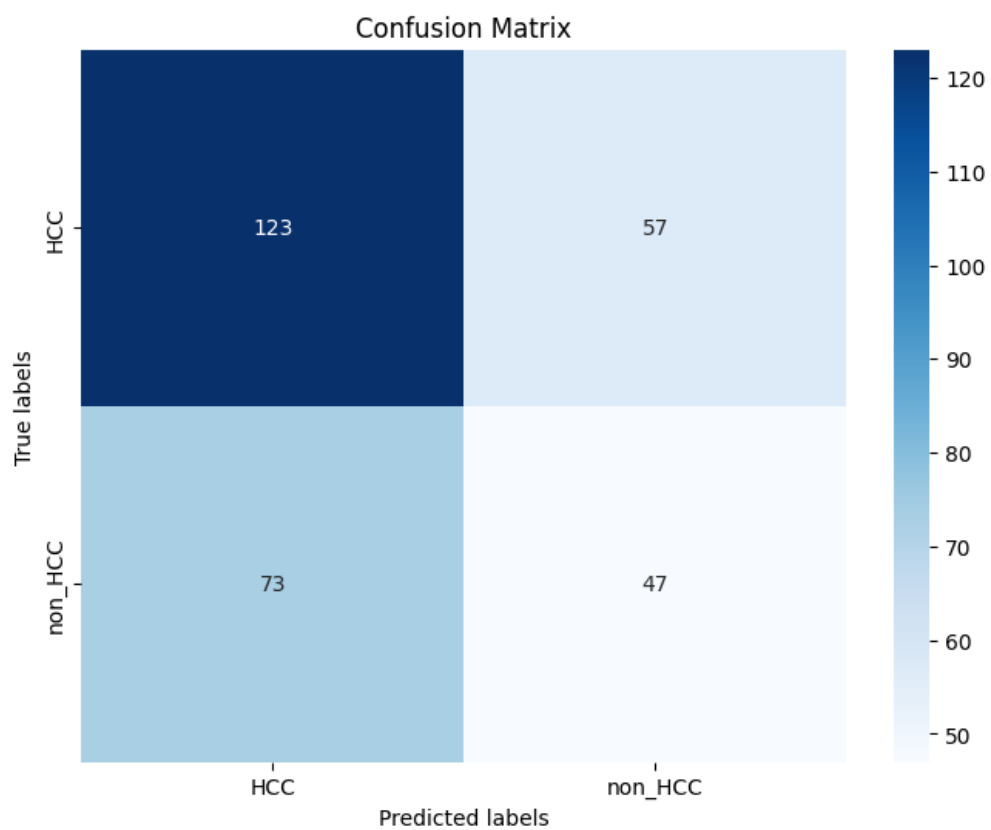
ภาพประกอบ 27 Learning Curve แบบจำลอง VGG16 Version 1 fold 1



ภาพประกอบ 28 Confusion Matrix แบบจำลอง VGG16 Version 1 fold 2



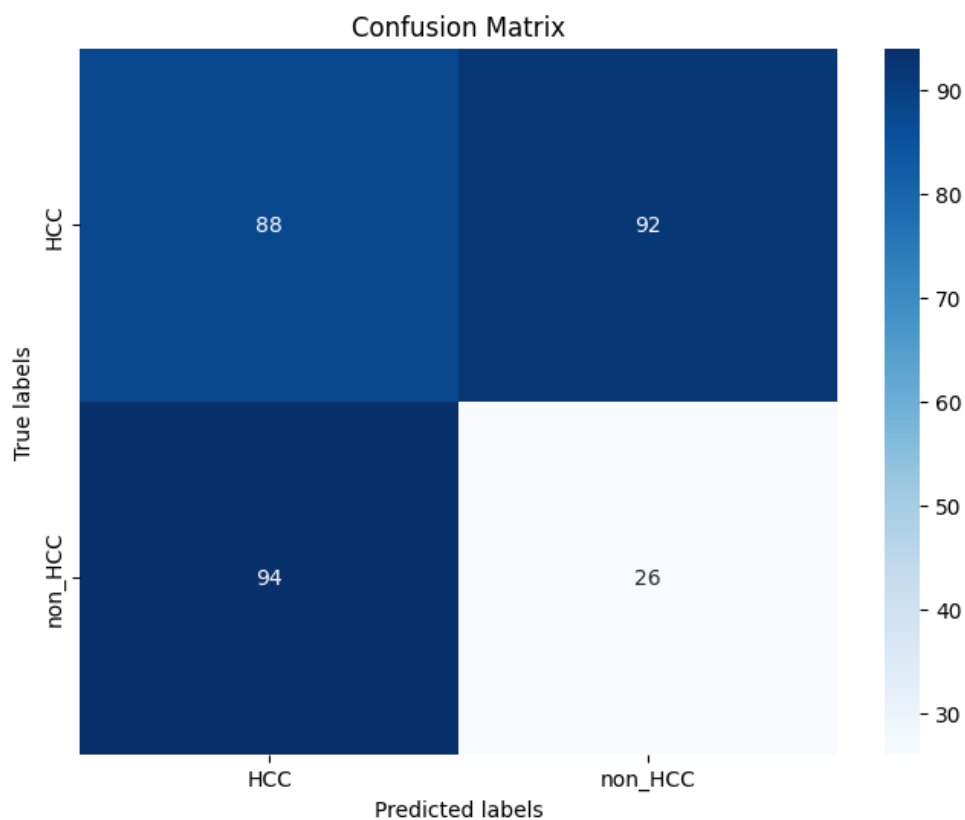
ภาพประกอบ 29 Learning Curve แบบจำลอง VGG16 Version 1 fold 2



ภาพประกอบ 30 Confusion Matrix แบบจำลอง VGG16 Version 1 fold 3



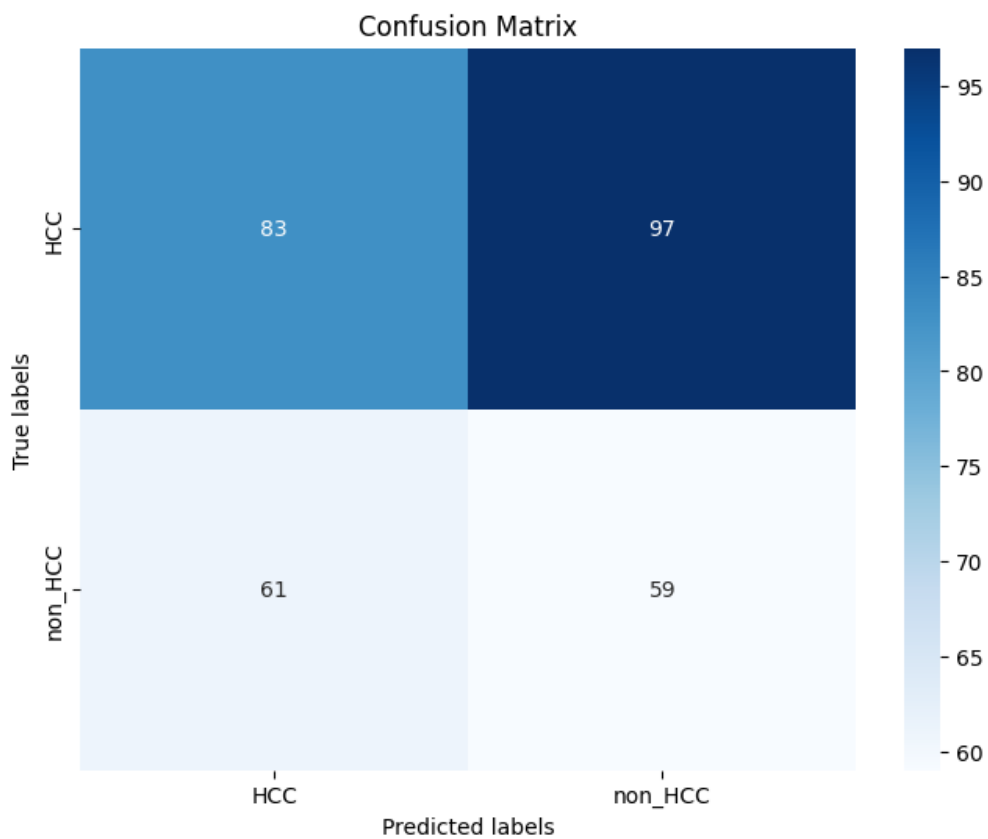
ภาพประกอบ 31 Learning Curve แบบจำลอง VGG16 Version 1 fold 3



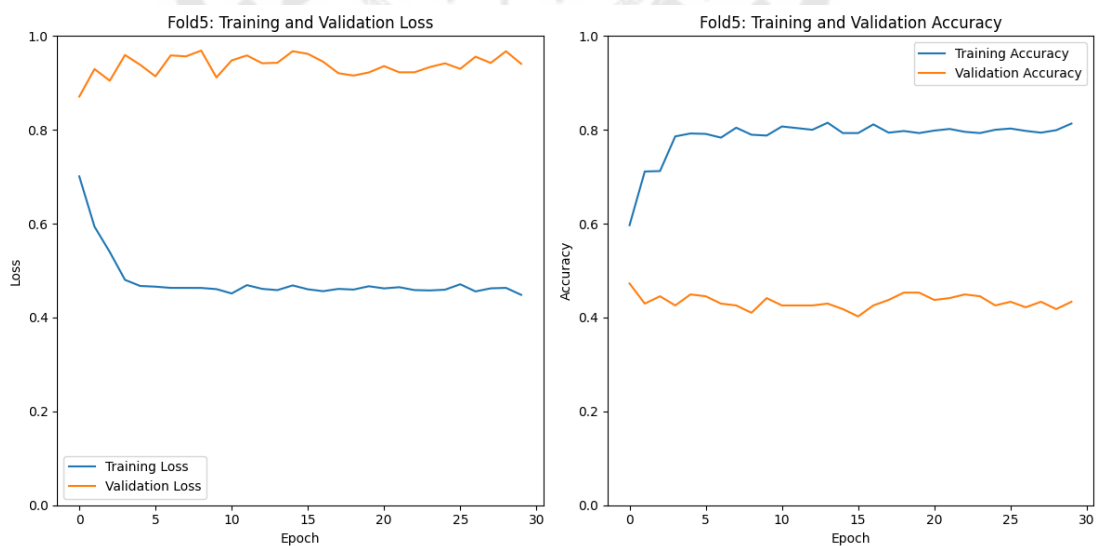
ภาพประกอบ 32 Confusion Matrix แบบจำลอง VGG16 Version 1 fold 4



ภาพประกอบ 33 Learning Curve แบบจำลอง VGG16 Version 1 fold 4

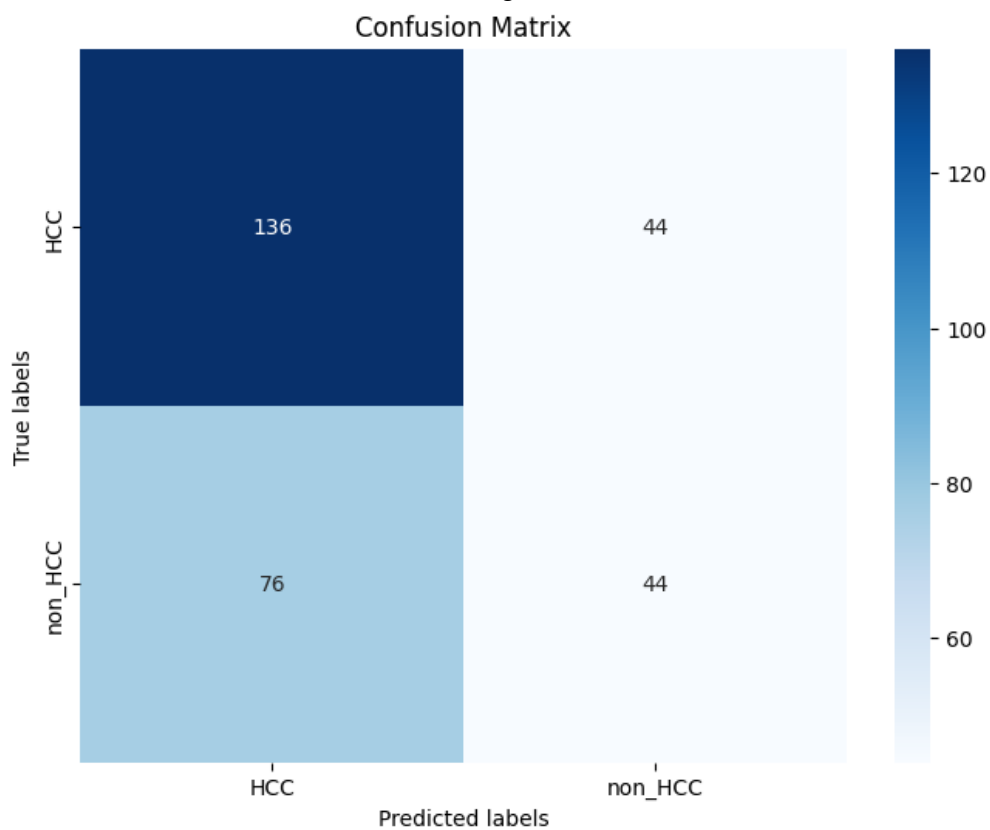


ภาพประกอบ 34 Confusion Matrix แบบจำลอง VGG16 Version 1 fold 5

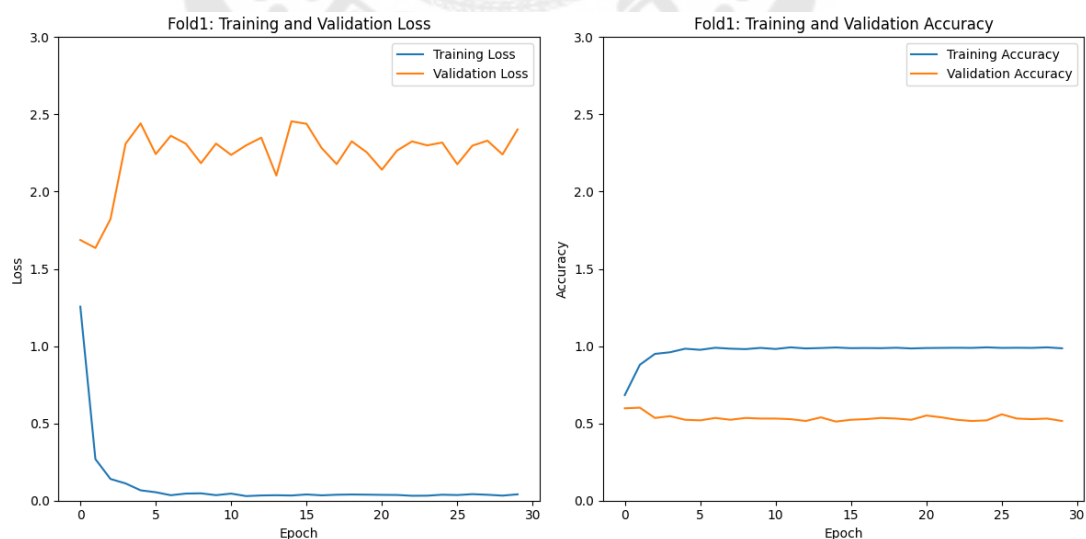


ภาพประกอบ 35 Learning Curve แบบจำลอง VGG16 Version 1 fold 5

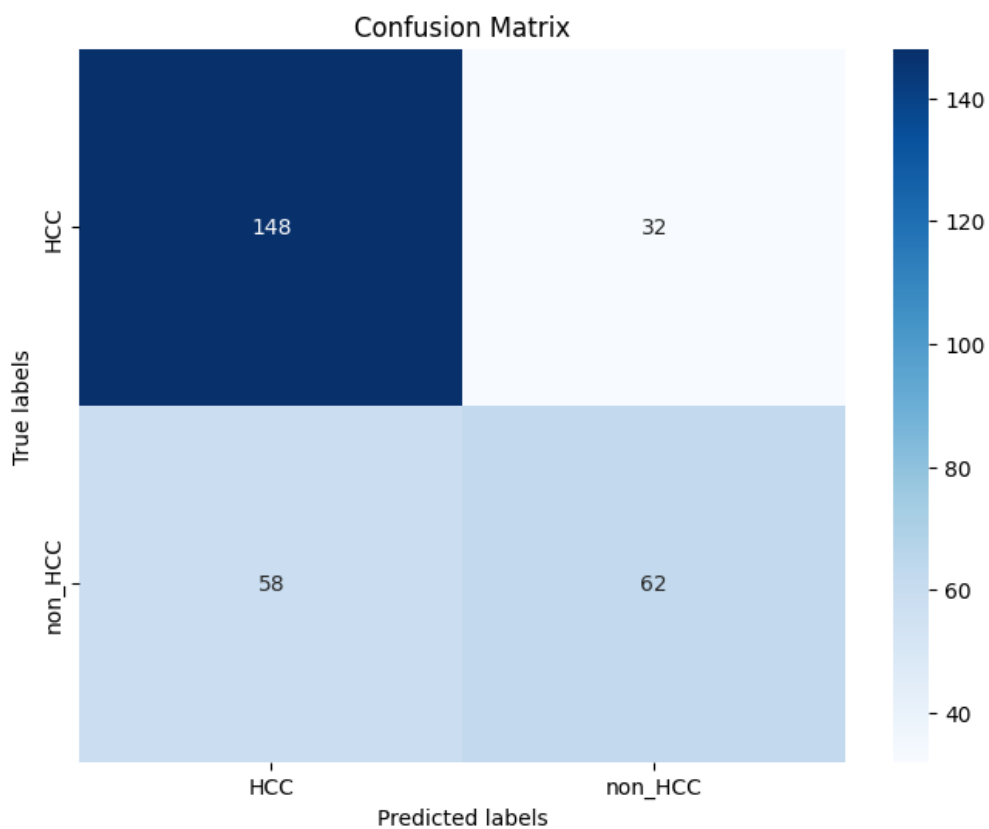
4.1.2 Confusion Matrix และ Learning Curve แบบจำลอง VGG16 Version 2



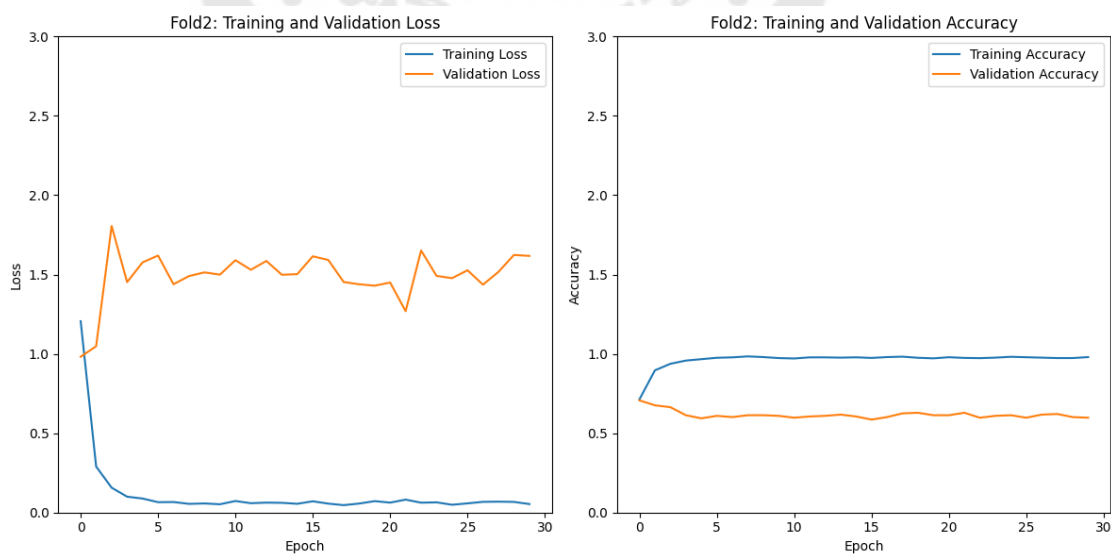
ภาพประกอบ 36 Confusion Matrix แบบจำลอง VGG16 Version 2 fold 1



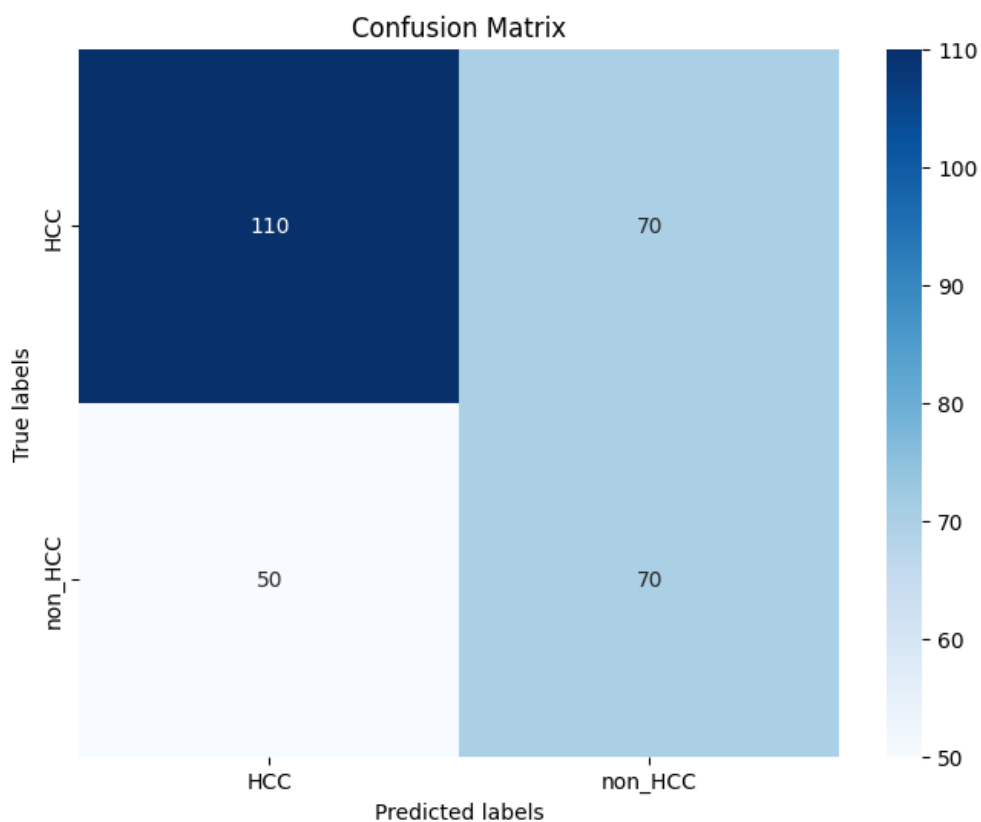
ภาพประกอบ 37 Learning Curve แบบจำลอง VGG16 Version 2 fold 1



ภาพประกอบ 38 Confusion Matrix แบบจำลอง VGG16 Version 2 fold 2



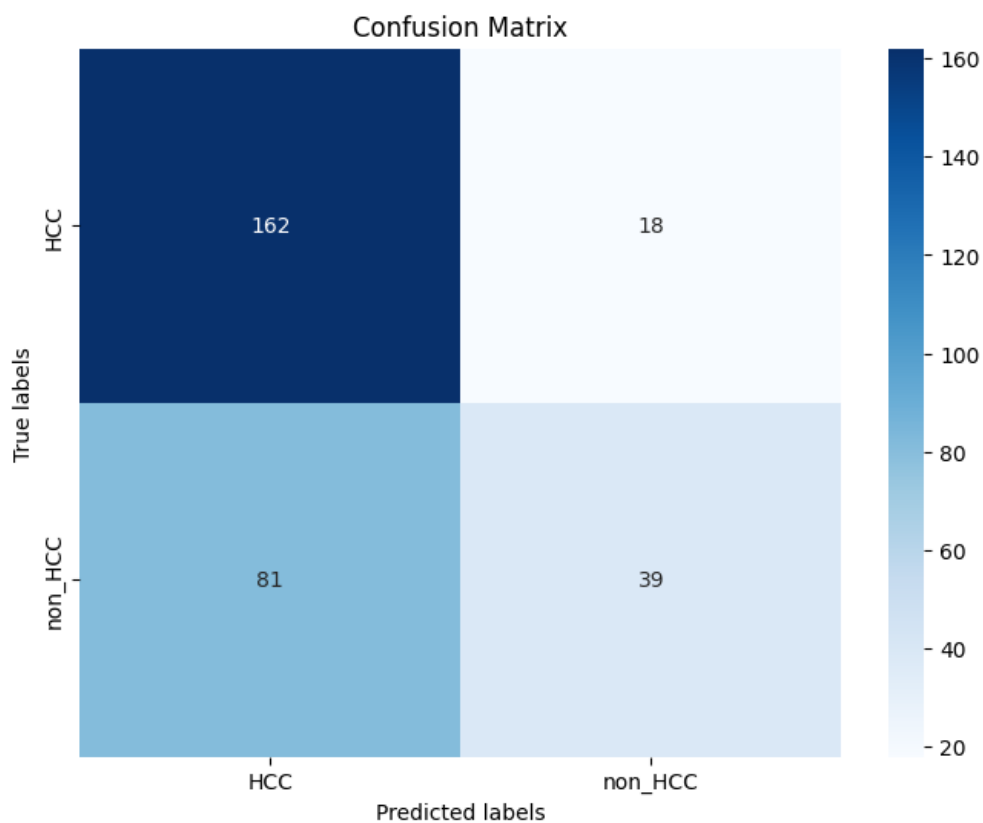
ภาพประกอบ 39 Learning Curve แบบจำลอง VGG16 Version 2 fold 2



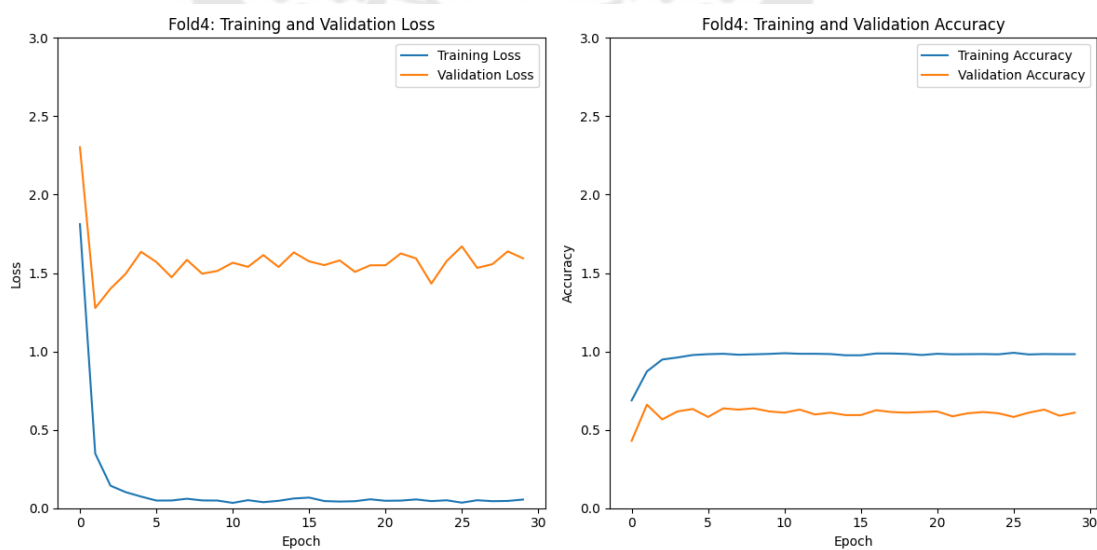
ภาพประกอบ 40 Confusion Matrix แบบจำลอง VGG16 Version 2 fold 3



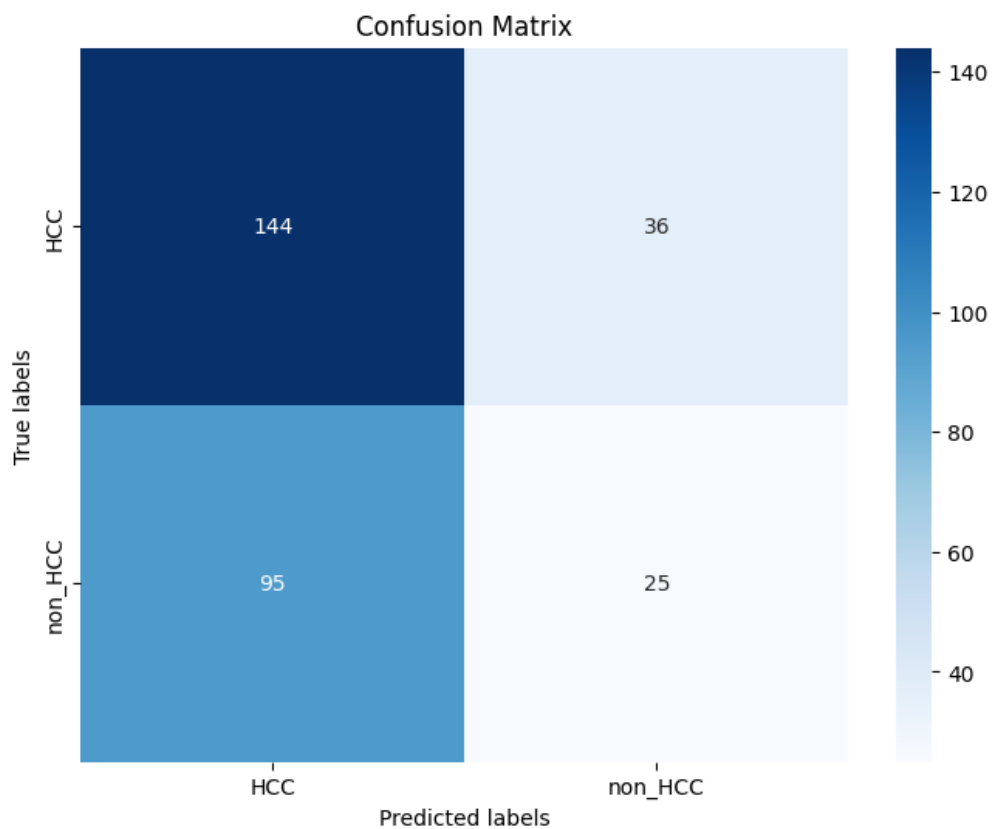
ภาพประกอบ 41 Learning Curve แบบจำลอง VGG16 Version 2 fold 3



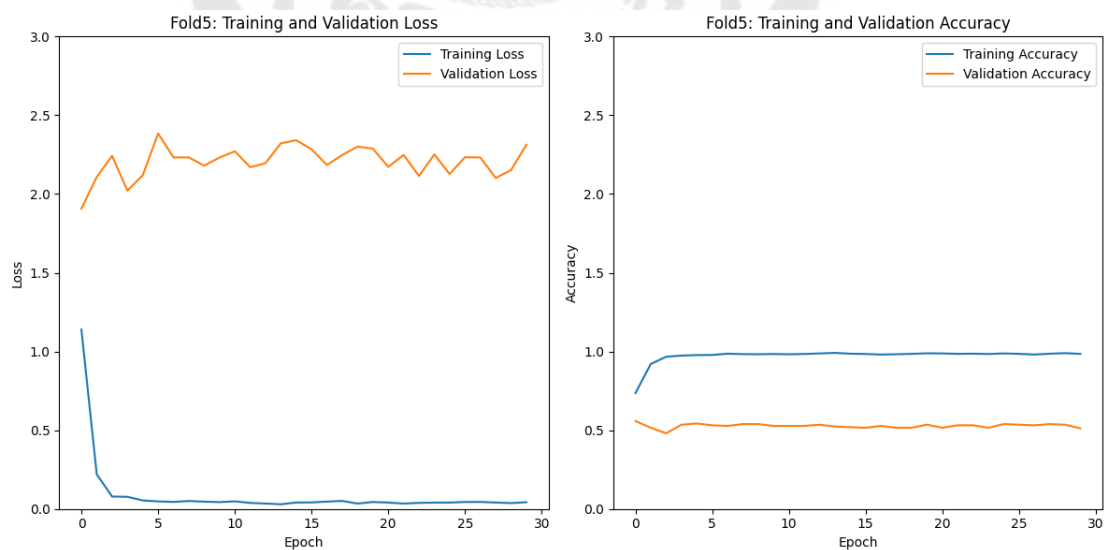
ภาพประกอบ 42 Confusion Matrix แบบจำลอง VGG16 Version 2 fold 4



ภาพประกอบ 43 Learning Curve แบบจำลอง VGG16 Version 2 fold 4

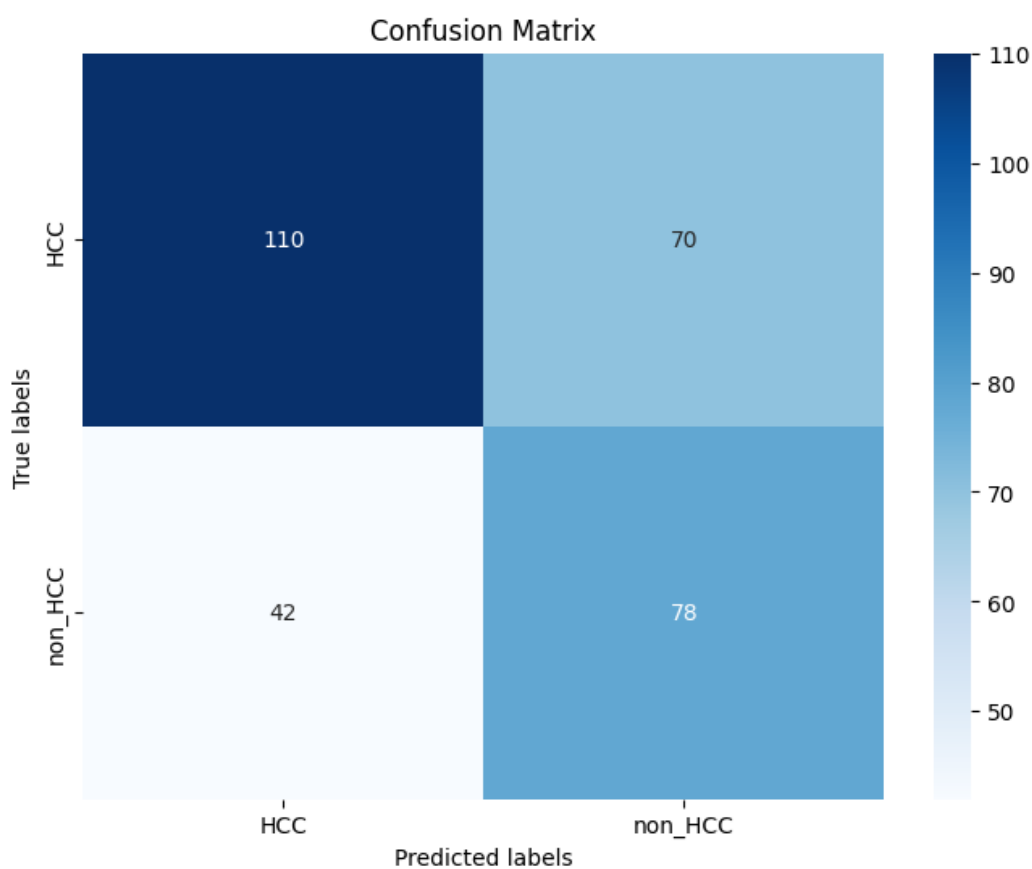


ภาพประกอบ 44 Confusion Matrix แบบจำลอง VGG16 Version 2 fold 5

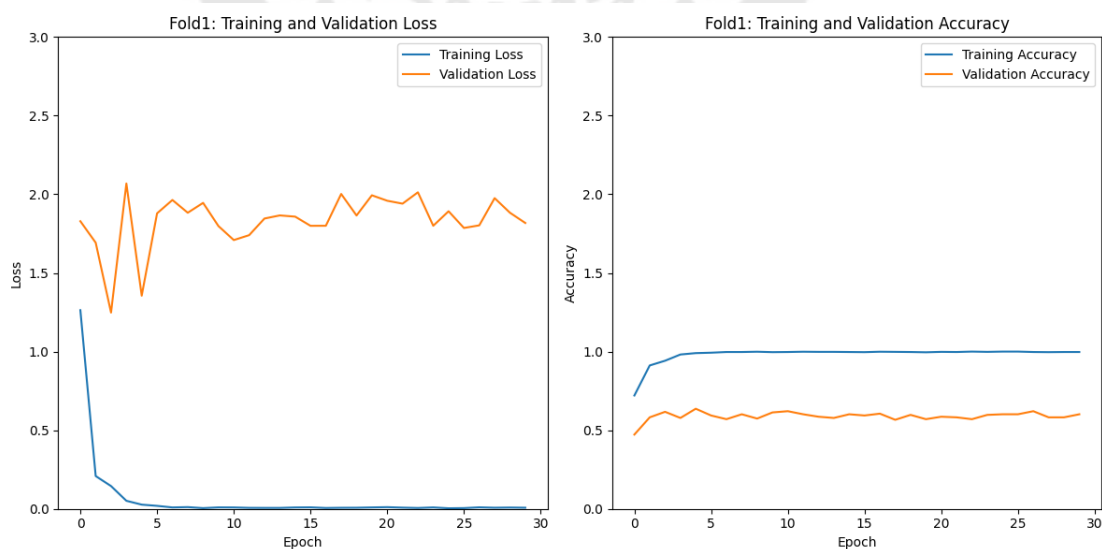


ภาพประกอบ 45 Learning Curve แบบจำลอง VGG16 Version 2 fold 5

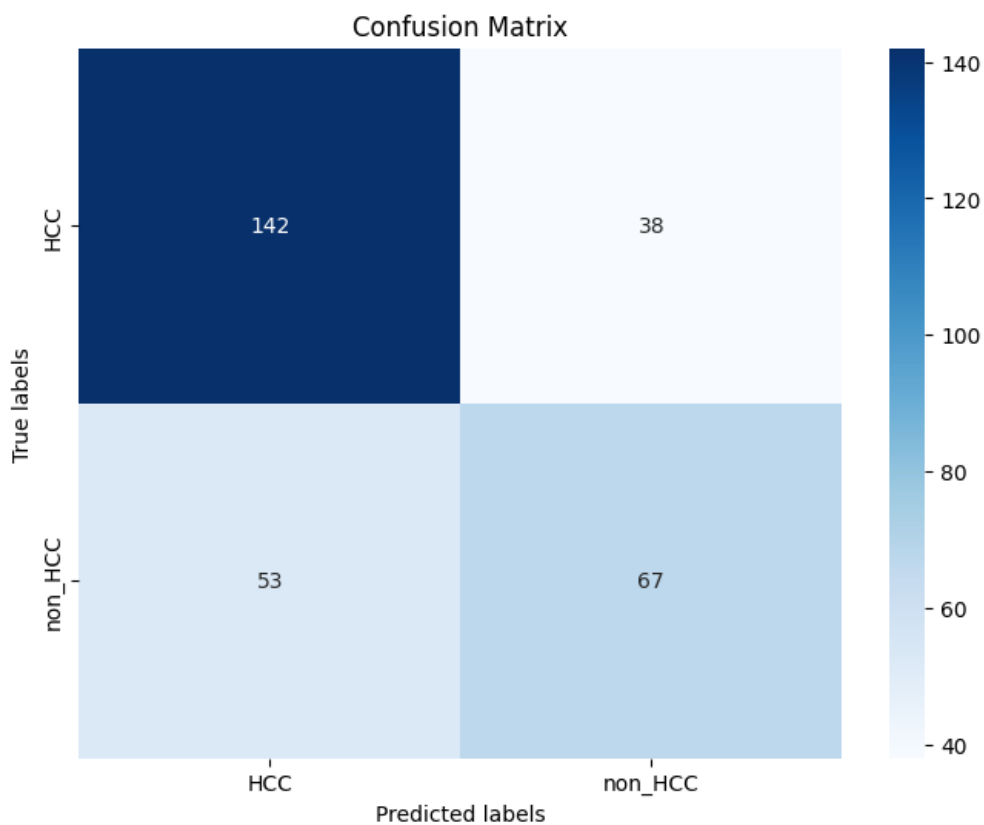
4.1.3 Confusion Matrix และ Learning Curve แบบจำลอง VGG16 Version 3



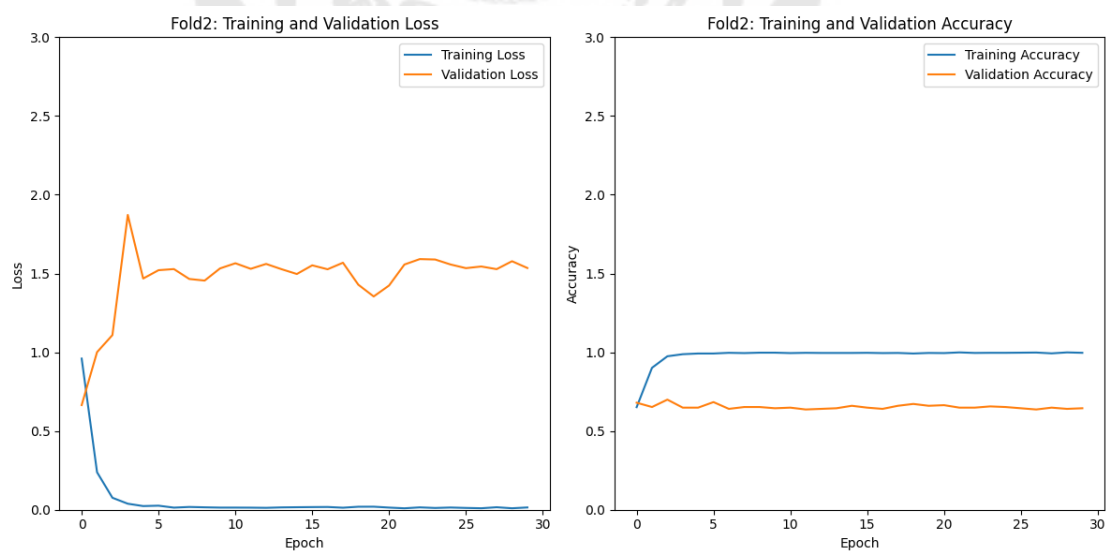
ภาพประกอบ 46 Confusion Matrix แบบจำลอง VGG16 Version 3 fold 1



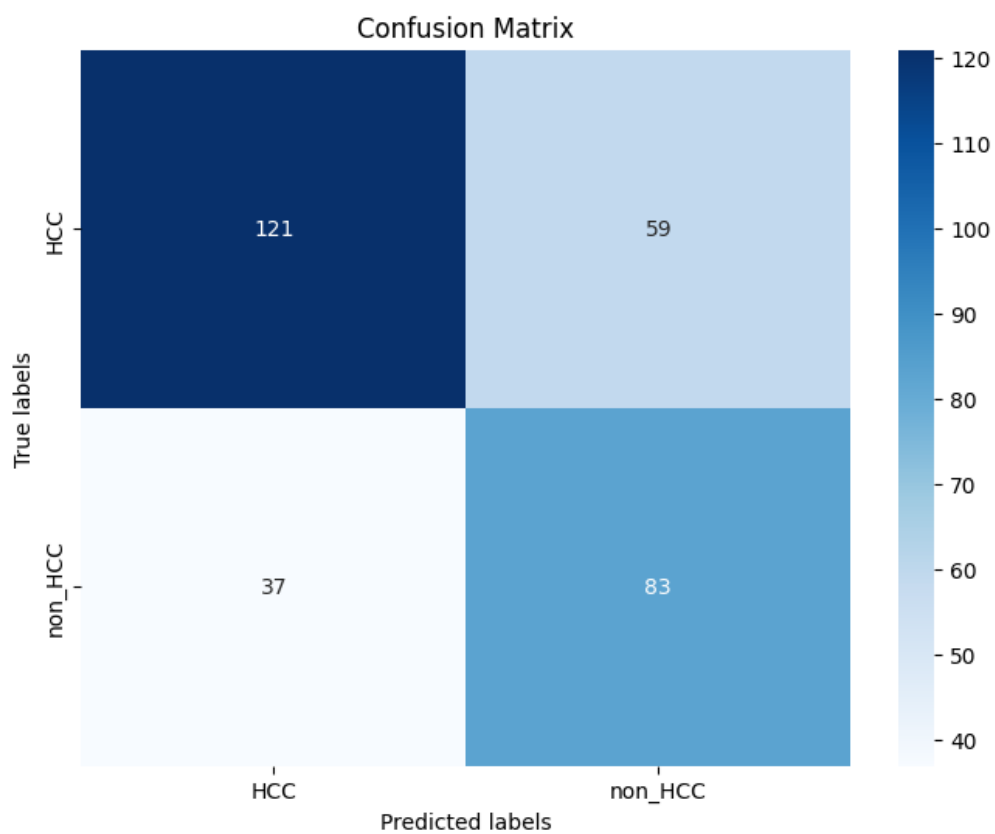
ภาพประกอบ 47 Learning Curve แบบจำลอง VGG16 Version 3 fold 1



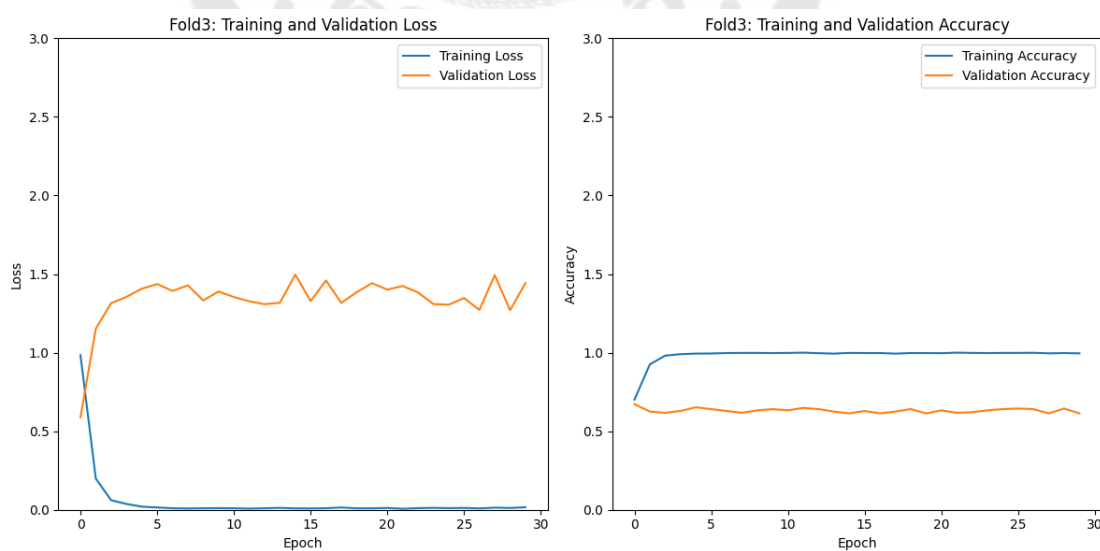
ภาพประกอบ 48 Confusion Matrix แบบจำลอง VGG16 Version 3 fold 2



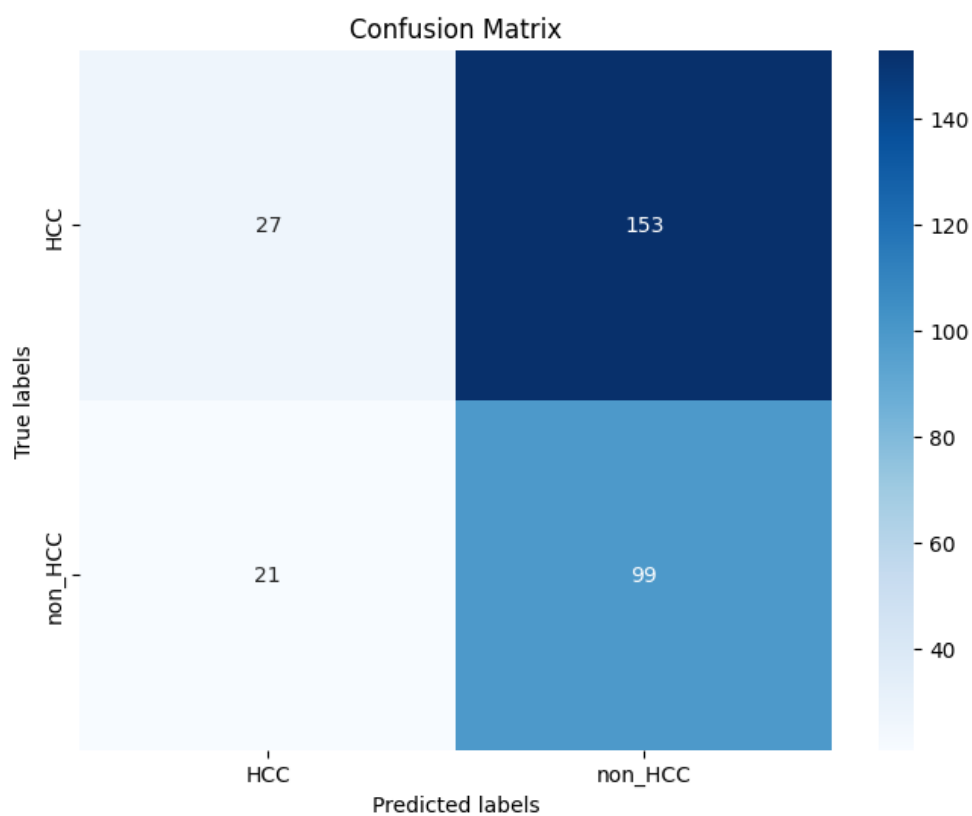
ภาพประกอบ 49 Learning Curve แบบจำลอง VGG16 Version 3 fold 2



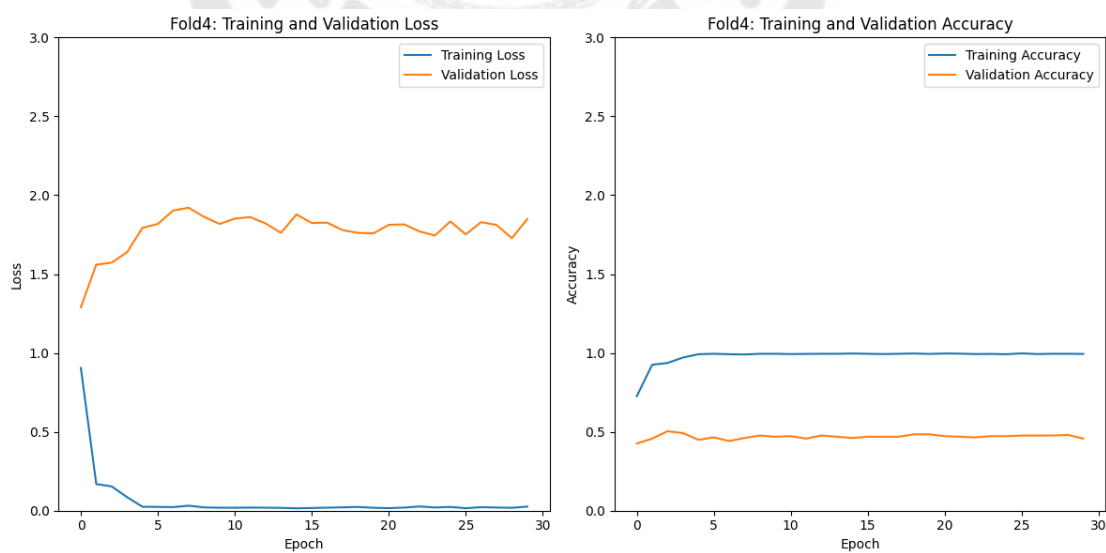
ภาพประกอบ 50 Confusion Matrix แบบจำลอง VGG16 Version 3 fold 3



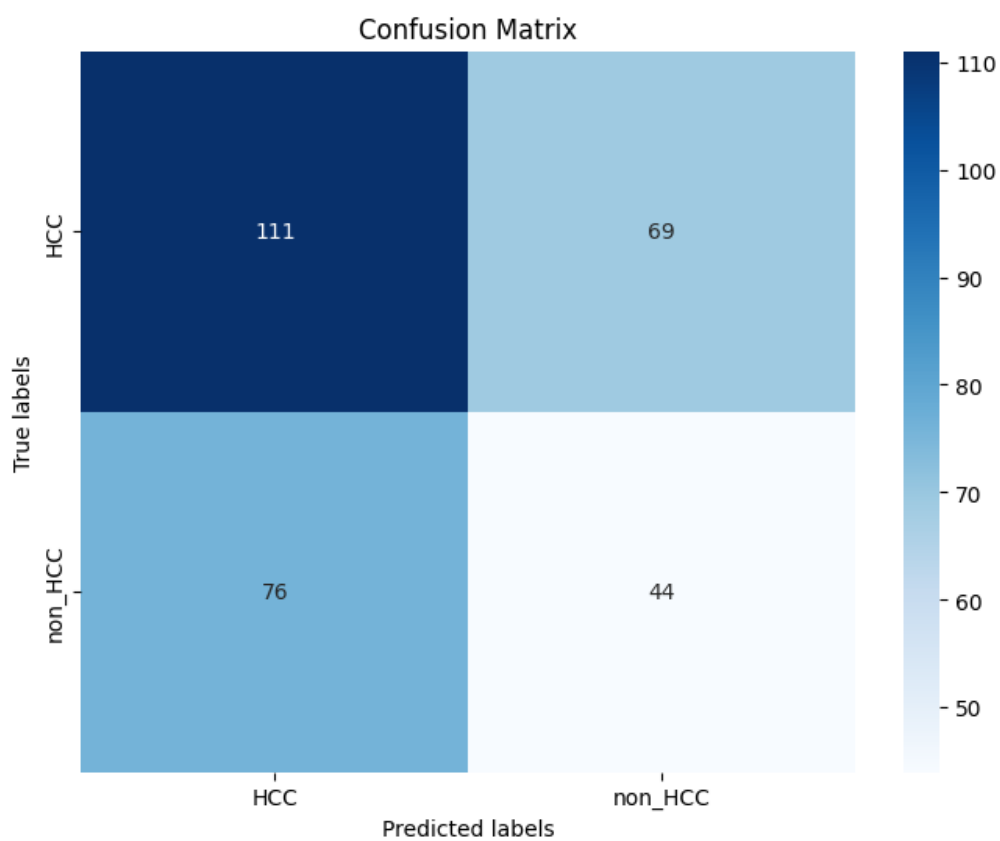
ภาพประกอบ 51 Learning Curve แบบจำลอง VGG16 Version 3 fold 3



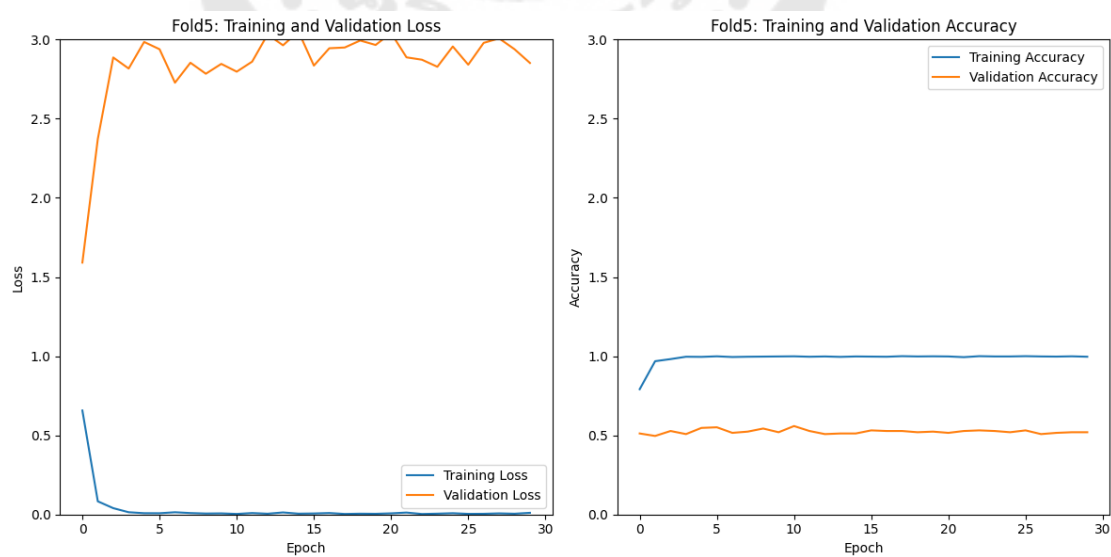
ภาพประกอบ 52 Confusion Matrix แบบจำลอง VGG16 Version 3 fold 4



ภาพประกอบ 53 Learning Curve แบบจำลอง VGG16 Version 3 fold 4

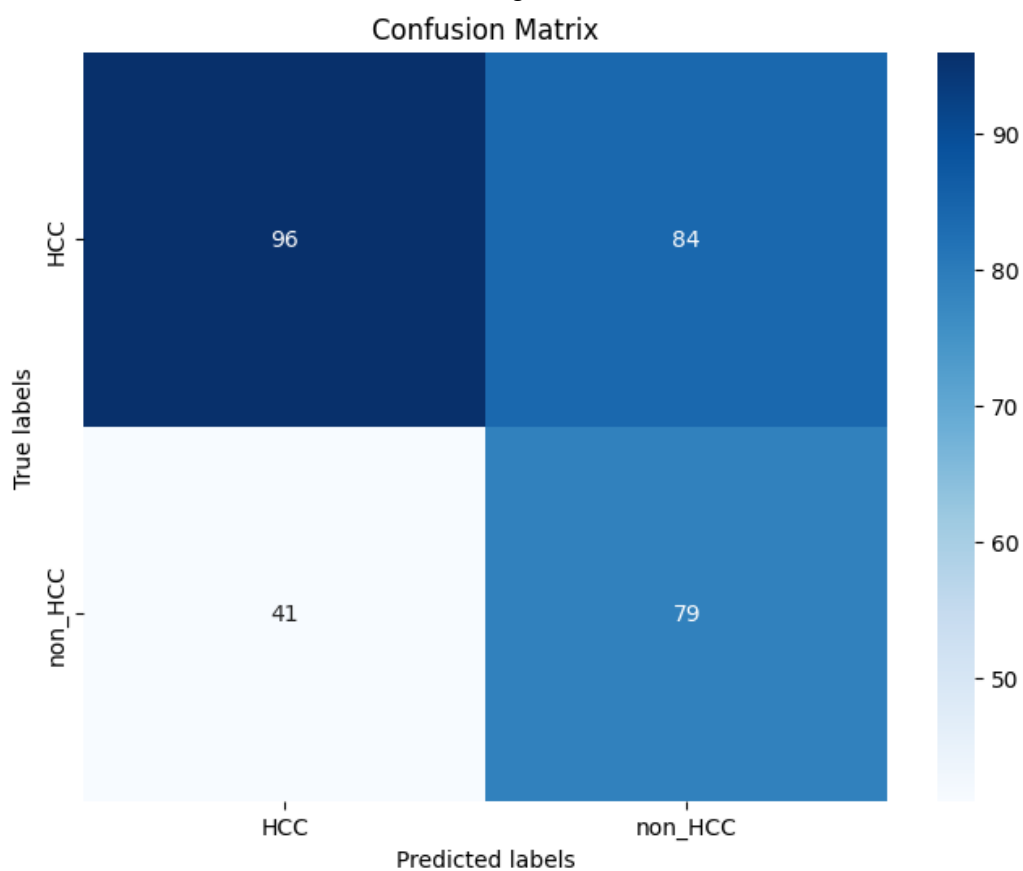


ภาพประกอบ 54 Confusion Matrix แบบจำลอง VGG16 Version 3 fold 5



ภาพประกอบ 55 Learning Curve แบบจำลอง VGG16 Version 3 fold 5

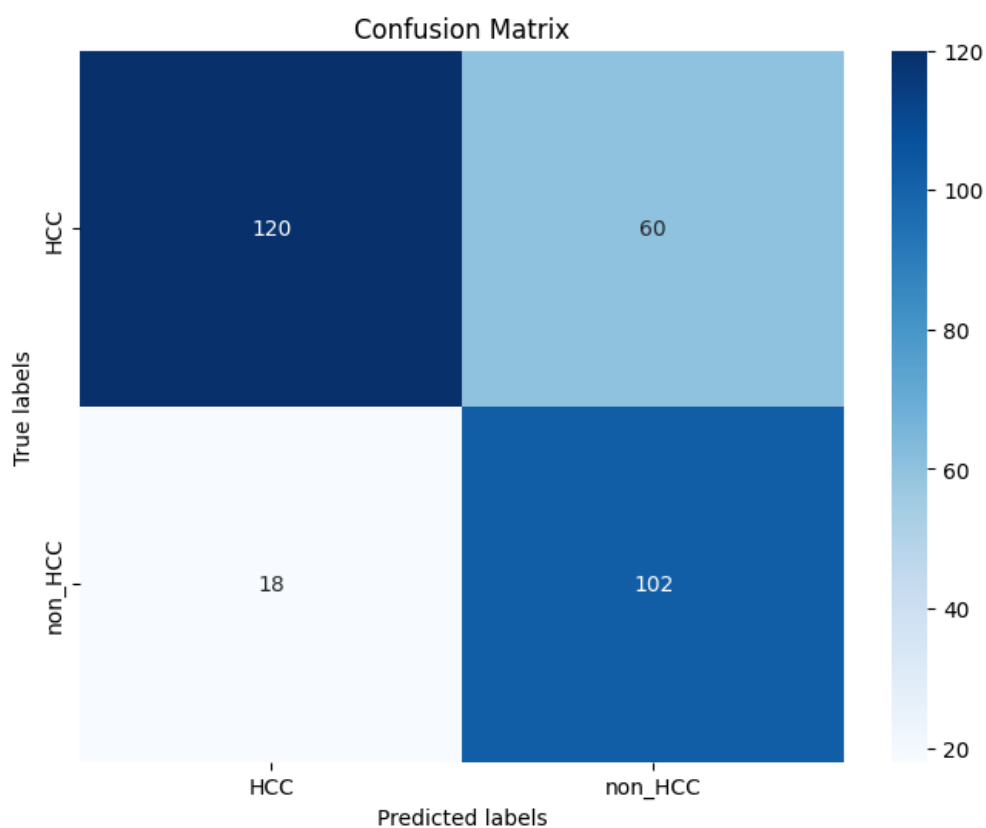
4.1.4 Confusion Matrix และ Learning Curve แบบจำลอง VGG16 Version 4



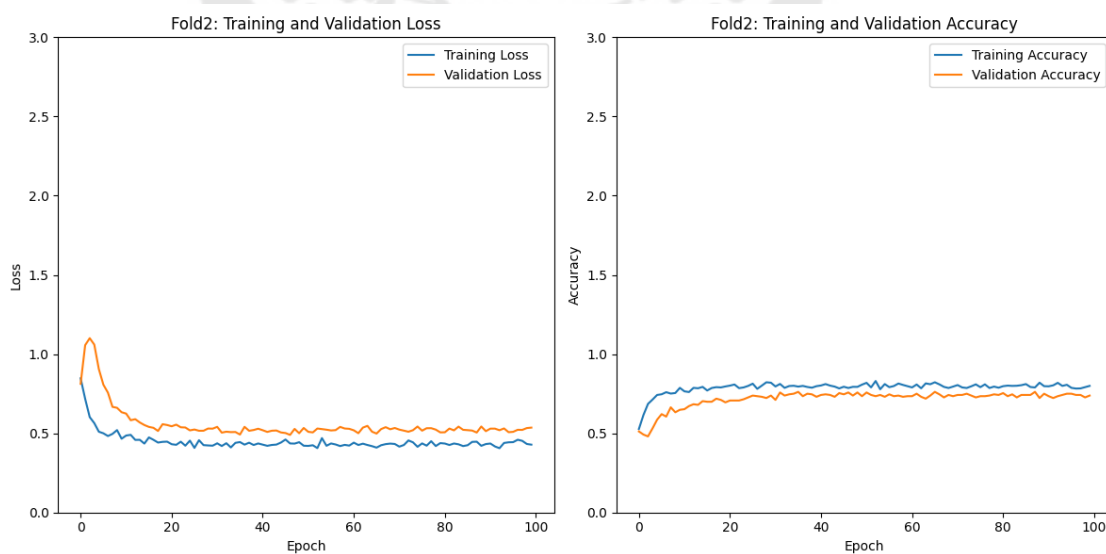
ภาพประกอบ 56 Confusion Matrix แบบจำลอง VGG16 Version 4 fold 1



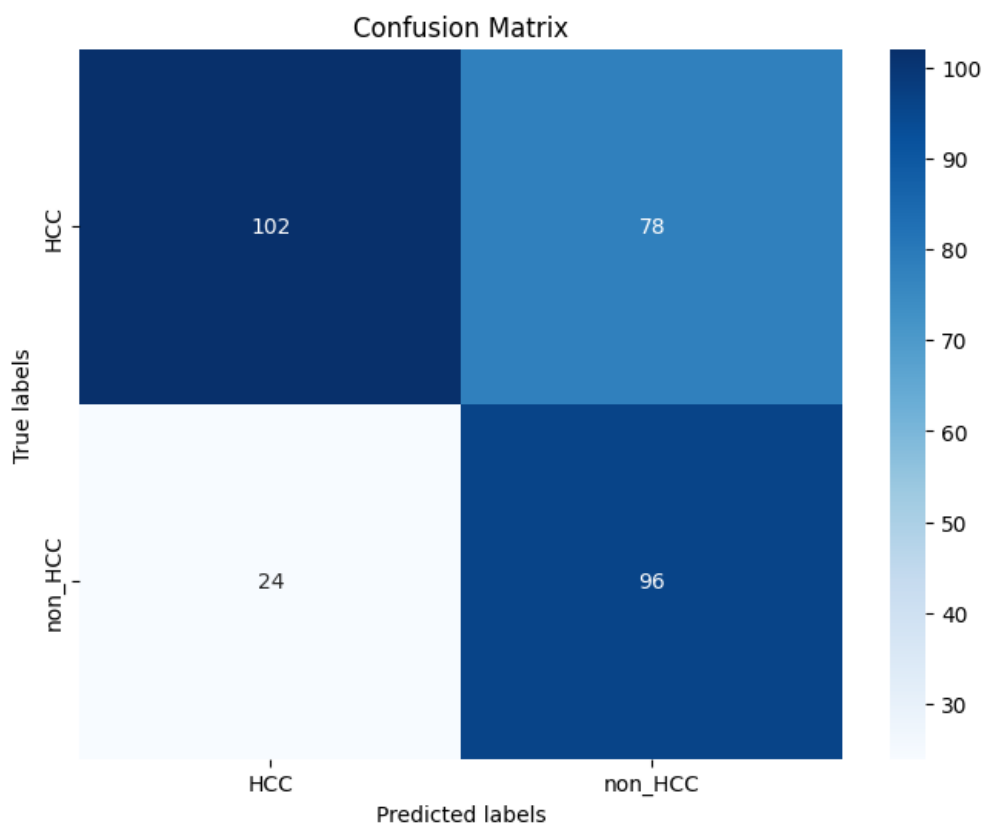
ภาพประกอบ 57 Learning Curve แบบจำลอง VGG16 Version 4 fold 1



ภาพประกอบ 58 Confusion Matrix แบบจำลอง VGG16 Version 4 fold 2



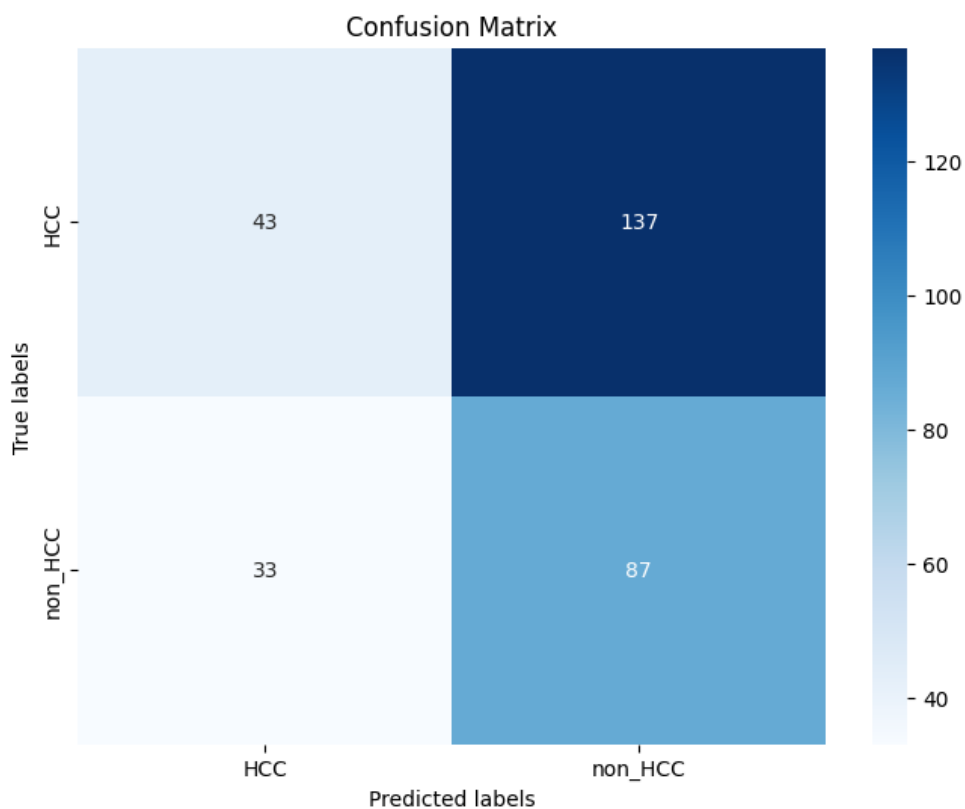
ภาพประกอบ 59 Learning Curve แบบจำลอง VGG16 Version 4 fold 2



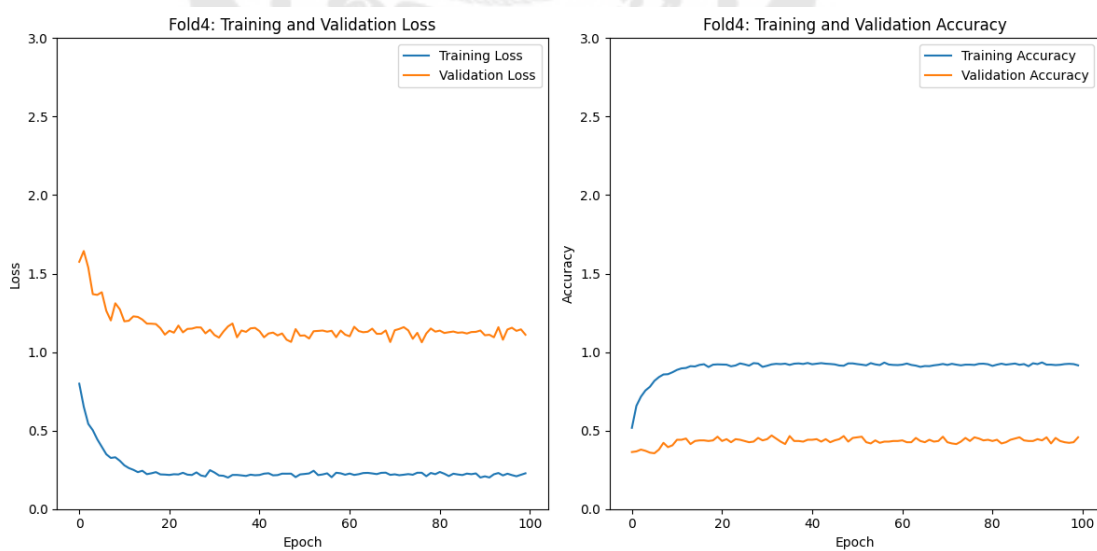
ภาพประกอบ 60 Confusion Matrix แบบจำลอง VGG16 Version 4 fold 3



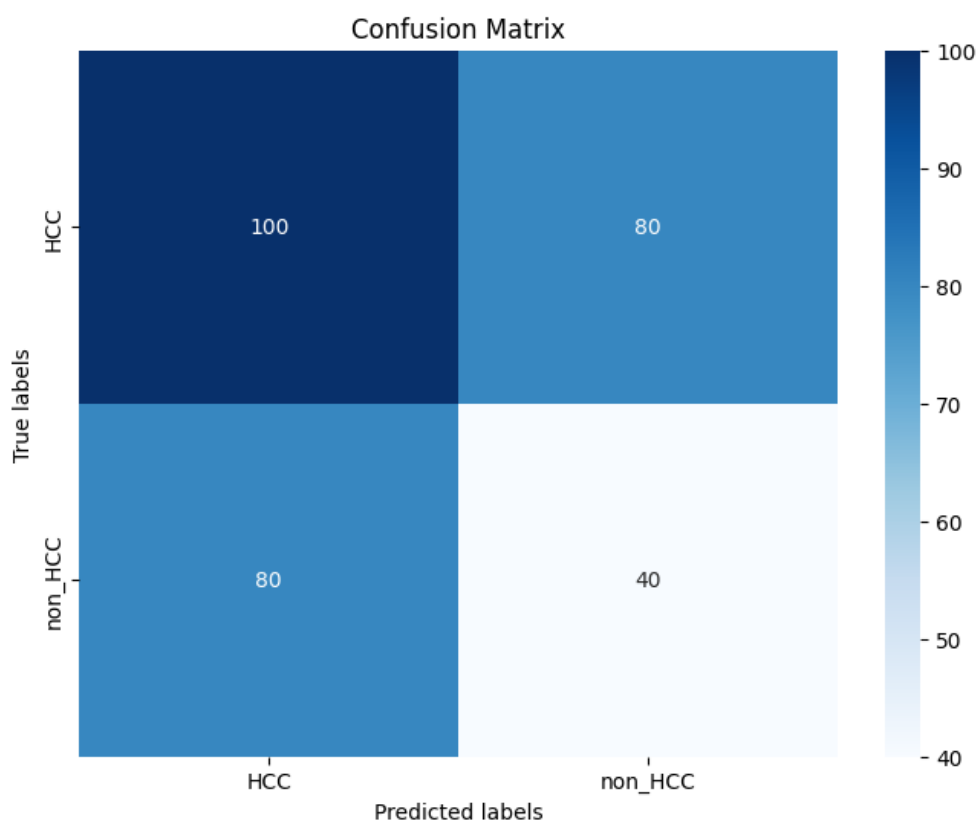
ภาพประกอบ 61 Learning Curve แบบจำลอง VGG16 Version 4 fold 3



ภาพประกอบ 62 Confusion Matrix แบบจำลอง VGG16 Version 4 fold 4



ภาพประกอบ 63 Learning Curve แบบจำลอง VGG16 Version 4 fold 4



ภาพประกอบ 64 Confusion Matrix แบบจำลอง VGG16 Version 4 fold 5



ภาพประกอบ 65 Learning Curve แบบจำลอง VGG16 Version 4 fold 5

นอกจากนี้เมื่อเปรียบเทียบตามการทำ Cross Validation โดยการสุ่มตัวอย่าง (Resampling) การแบ่งข้อมูลออกเป็นกลุ่มหลาย ๆ กลุ่ม เรียกว่า folds ซึ่งในที่นี้กำหนดให้มีค่าเท่ากับ 5 (5 folds) เรียงจากมากไปน้อยมีค่า Accuracy เฉลี่ย คือ fold 2 สูงสุด ที่ 0.71 fold 3 ที่ 0.63 fold 1 ที่ 0.60 fold 5 ที่ 0.51 และ fold 4 ที่ 0.48 ดังตาราง 11

ตาราง 11 ตารางเปรียบเทียบค่า Accuracy เฉลี่ยแต่ละ fold ข้อมูลของแบบจำลอง VGG16

Fold No.	ค่า Accuracy เฉลี่ย
1	0.60
2	0.71
3	0.63
4	0.48
5	0.51



ภาพประกอบ 66 สรุปผลการประเมินประสิทธิภาพแบบจำลองด้วยอัลกอริทึม VGG16

4.2 ผลการประเมินประสิทธิภาพแบบจำลองด้วยอัลกอริทึม ResNet-50

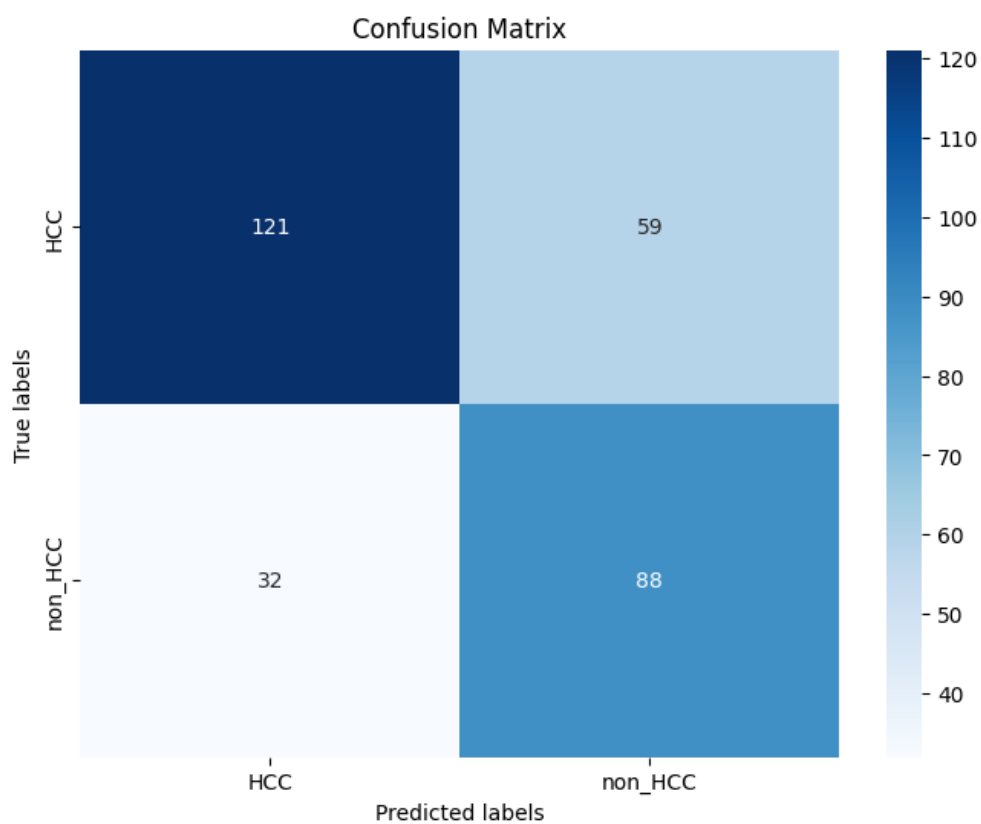
จากทั้งหมด 4 ชุดผลการทดลองด้วยอัลกอริทึม ResNet-50 นั้น พบว่า เมื่อพิจารณาประสิทธิภาพของแบบจำลองด้วยเมตริกซ์ Accuracy ที่มุ่งเน้นเรื่องอัตราส่วนของตัวอย่างที่แบบจำลองทำนายถูกต้องทั้งหมด เรียงค่า Accuracy จากมากไปน้อยจะได้ ResNet-50 Version 2 ที่ 0.69 ResNet-50 Version 4 ที่ 0.66 ResNet-50 Version 3 ที่ 0.64 และ ResNet-50 Version 1 ที่ 0.63 ดังตาราง 12

ตาราง 12 ตารางแสดงประสิทธิภาพการจำแนกของแบบจำลอง ResNet-50 แต่ละ Version

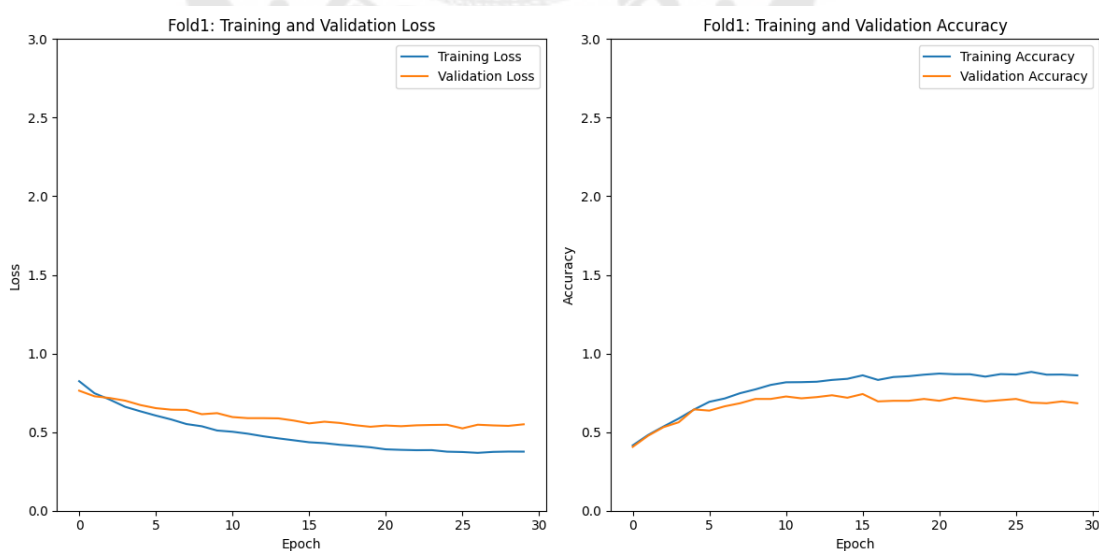
ResNet-50 Version	Accuracy	Precision	Recall	F1-score
1	0.63	0.62	0.62	0.61
2	0.69	0.68	0.68	0.68
3	0.64	0.63	0.63	0.62
4	0.66	0.65	0.65	0.65

เมื่อเปรียบเทียบผลการทดลอง ResNet-50 ในแต่ละ Version และตามข้อมูลแต่ละส่วนที่แบ่งการฝึกสอนและทดสอบข้อมูล (Fold) ตามการวัดประสิทธิภาพของแบบจำลองด้วย Confusion Matrix พร้อมทั้งดูกราฟการวิเคราะห์ผลการฝึกสอนแบบจำลอง (Learning Curve) โดยการวัดผลการเรียนรู้ของโมเดลจากข้อมูลฝึกสอน (Training Dataset) และข้อมูลทดสอบ (Validation Dataset) จะได้ผลลัพธ์ ดังนี้

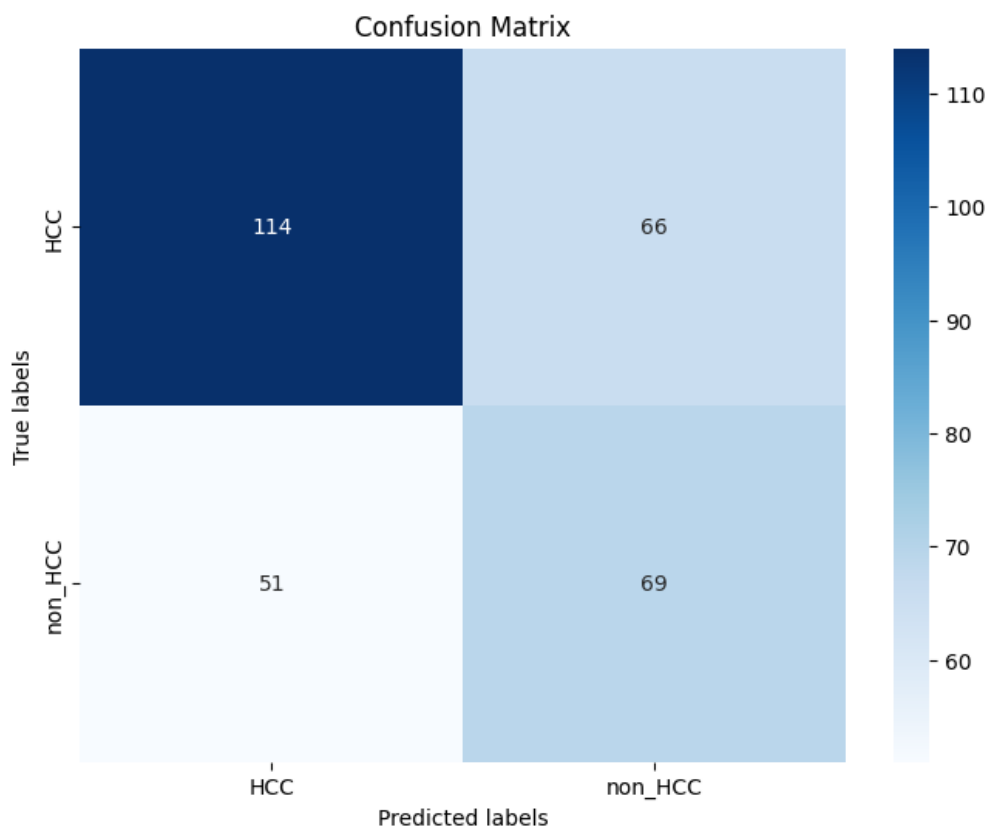
4.2.1 Confusion Matrix และ Learning Curve แบบจำลอง ResNet-50 Version 1



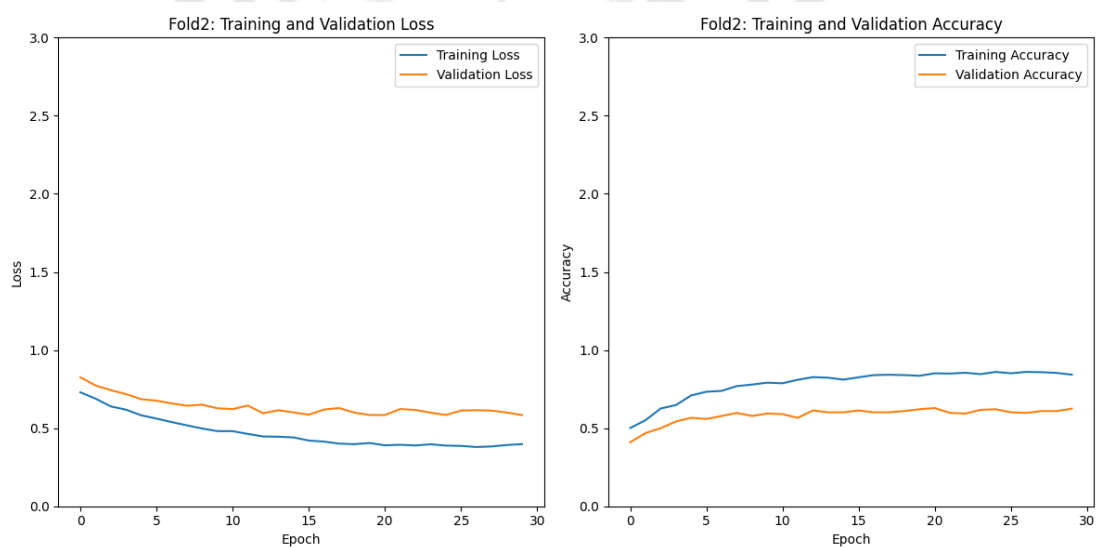
ภาพประกอบ 67 Confusion Matrix แบบจำลอง ResNet-50 Version 1 fold 1



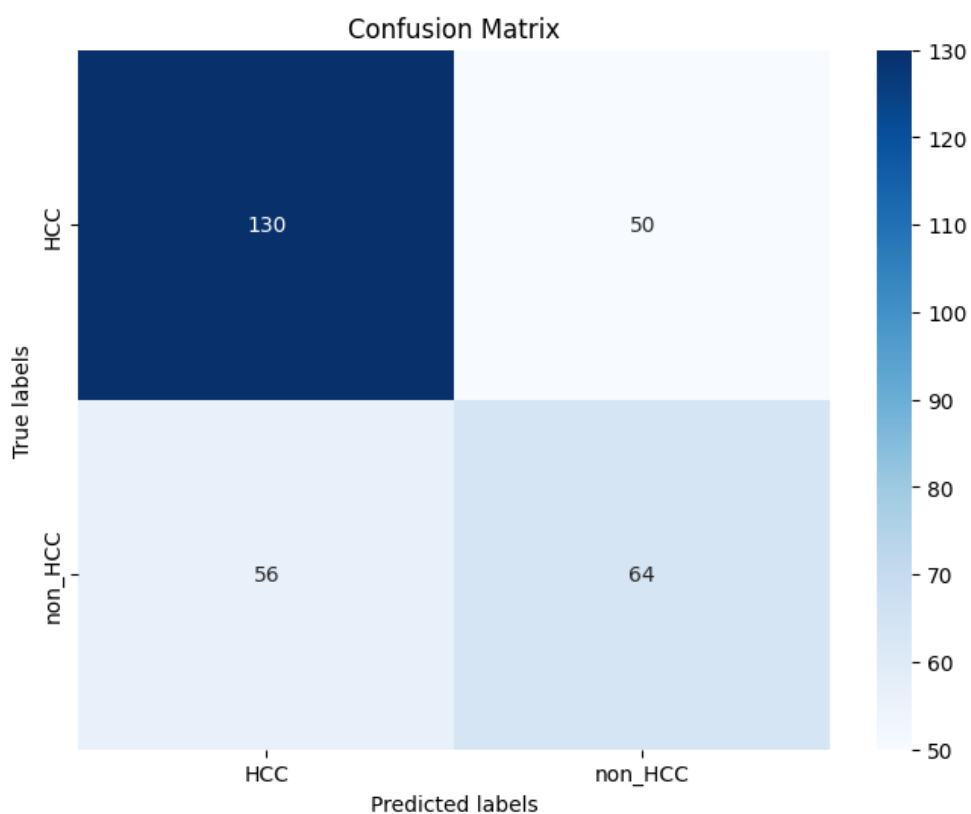
ภาพประกอบ 68 Learning Curve แบบจำลอง ResNet-50 Version 1 fold 1



ภาพประกอบ 69 Confusion Matrix แบบจำลอง ResNet-50 Version 1 fold 2



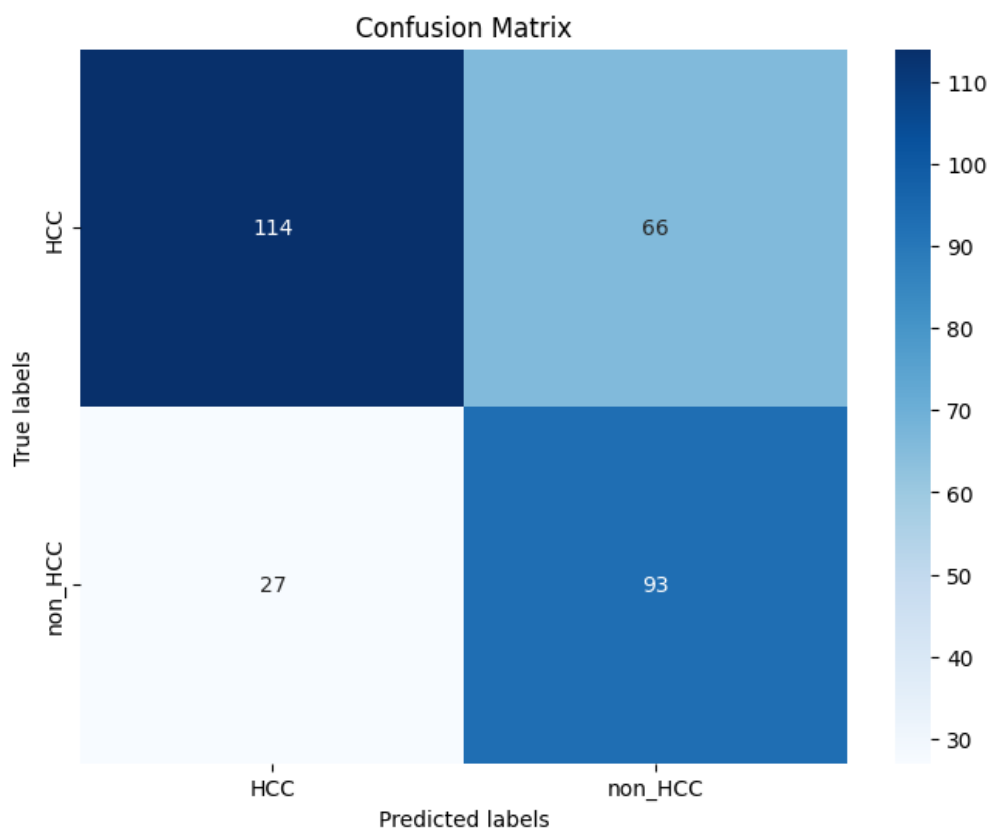
ภาพประกอบ 70 Learning Curve แบบจำลอง ResNet-50 Version 1 fold 2



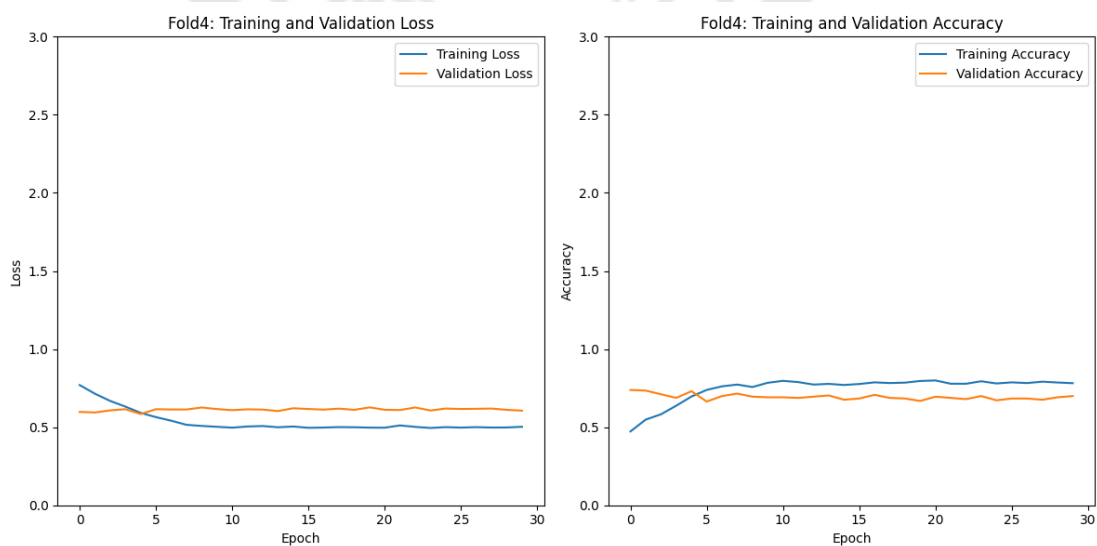
ภาพประกอบ 71 Confusion Matrix แบบจำลอง ResNet-50 Version 1 fold 3



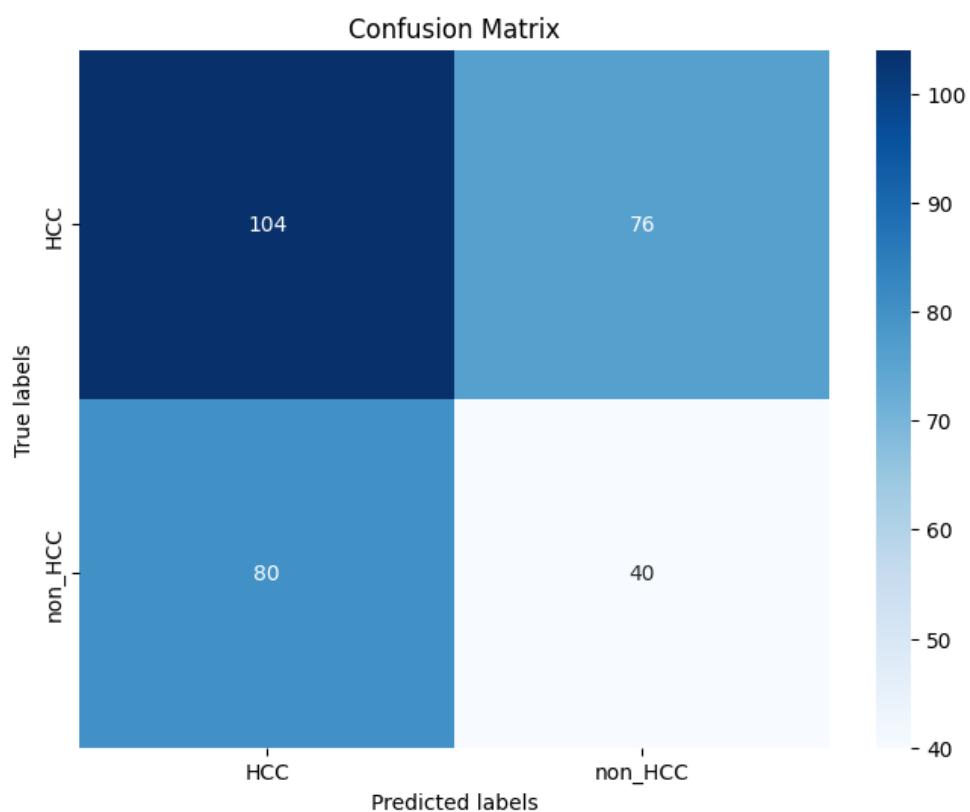
ภาพประกอบ 72 Learning Curve แบบจำลอง ResNet-50 Version 1 fold 3



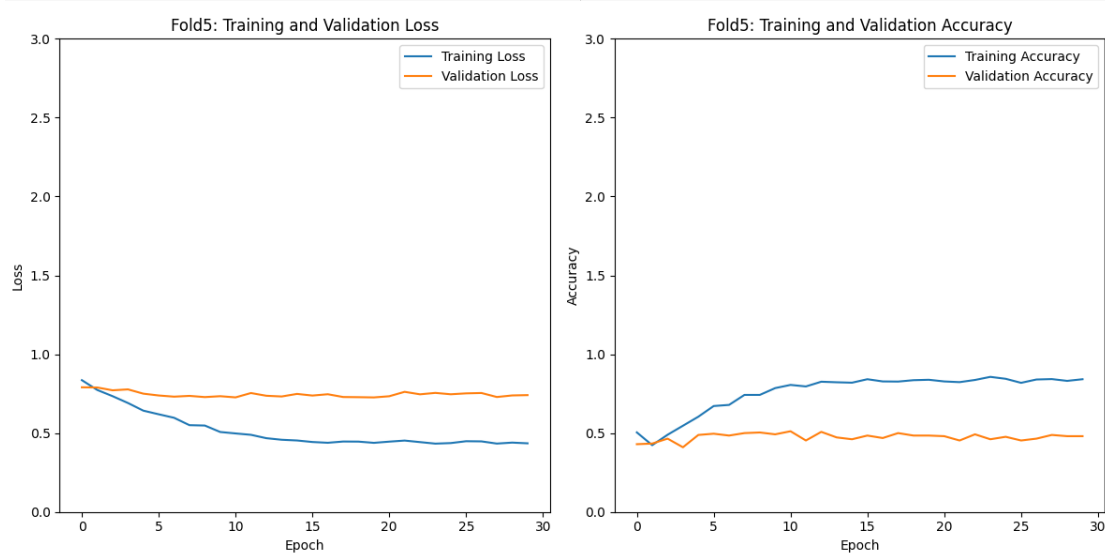
ภาพประกอบ 73 Confusion Matrix แบบจำลอง ResNet-50 Version 1 fold 4



ภาพประกอบ 74 Learning Curve แบบจำลอง ResNet-50 Version 1 fold 4

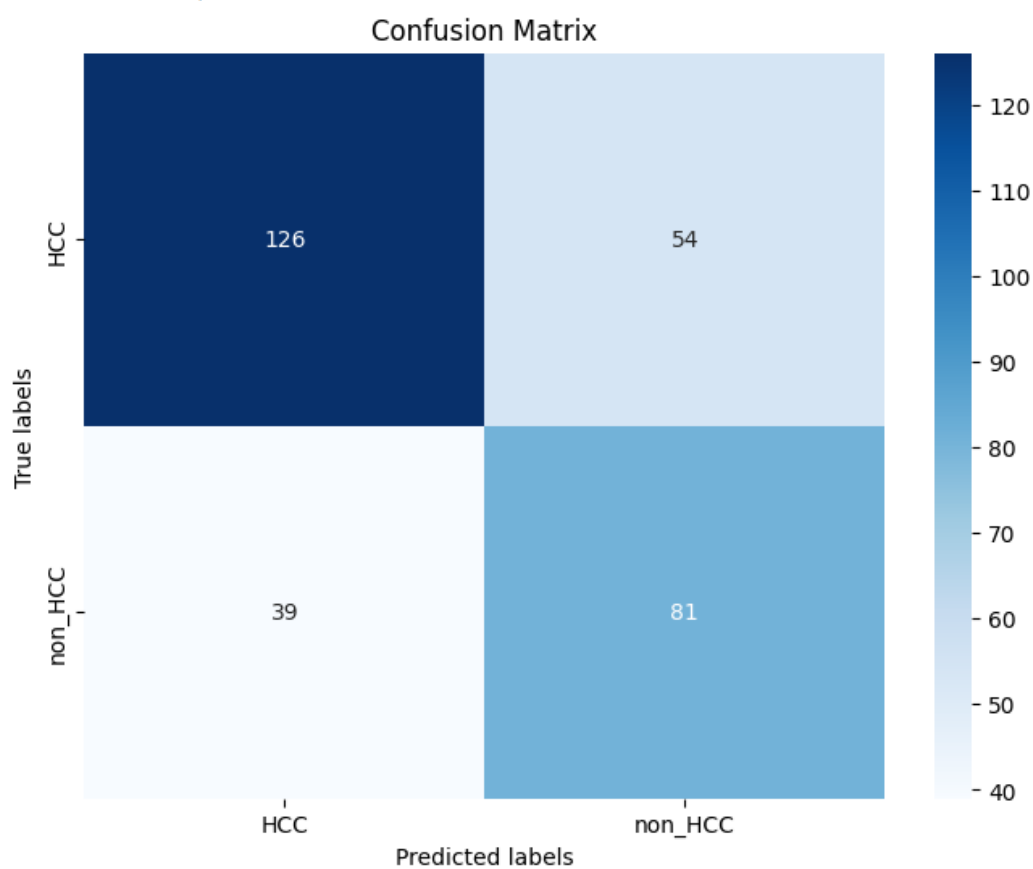


ภาพประกอบ 75 Confusion Matrix แบบจำลอง ResNet-50 Version 1 fold 5

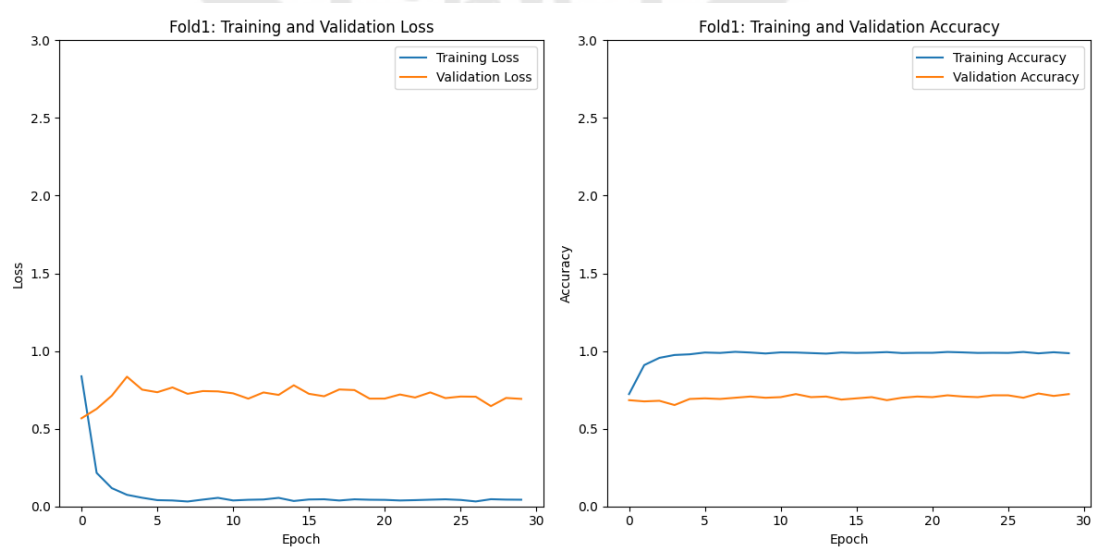


ภาพประกอบ 76 Learning Curve แบบจำลอง ResNet-50 Version 1 fold 5

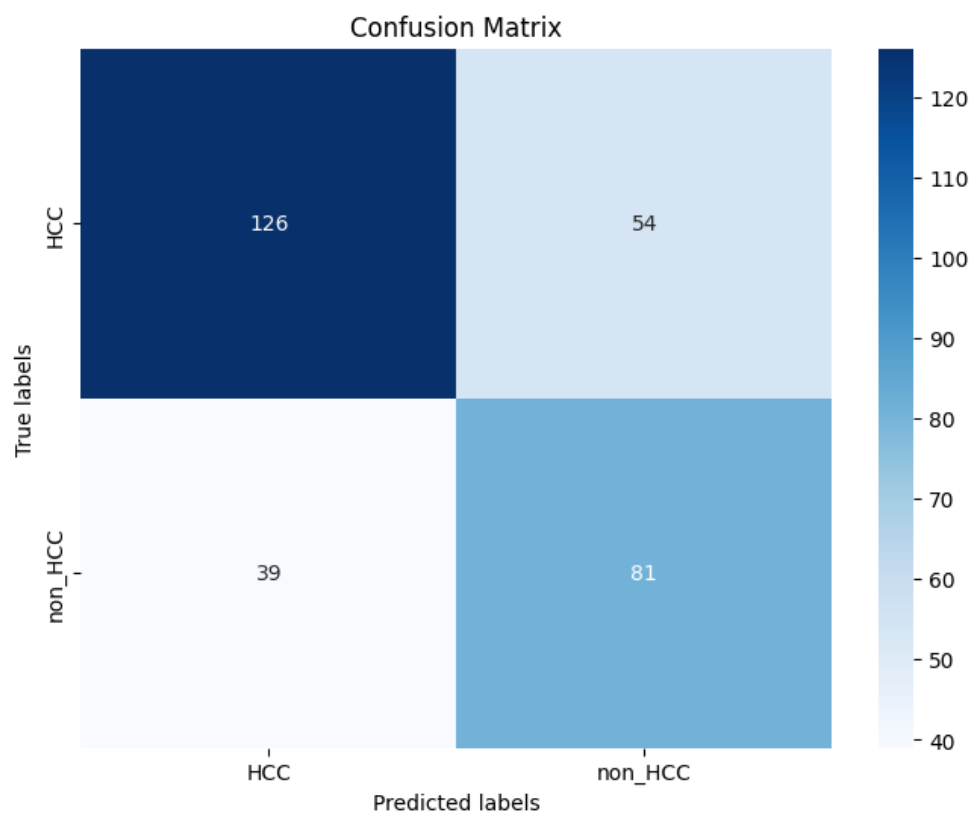
4.2.2 Confusion Matrix และ Learning Curve แบบจำลอง ResNet-50 Version 2



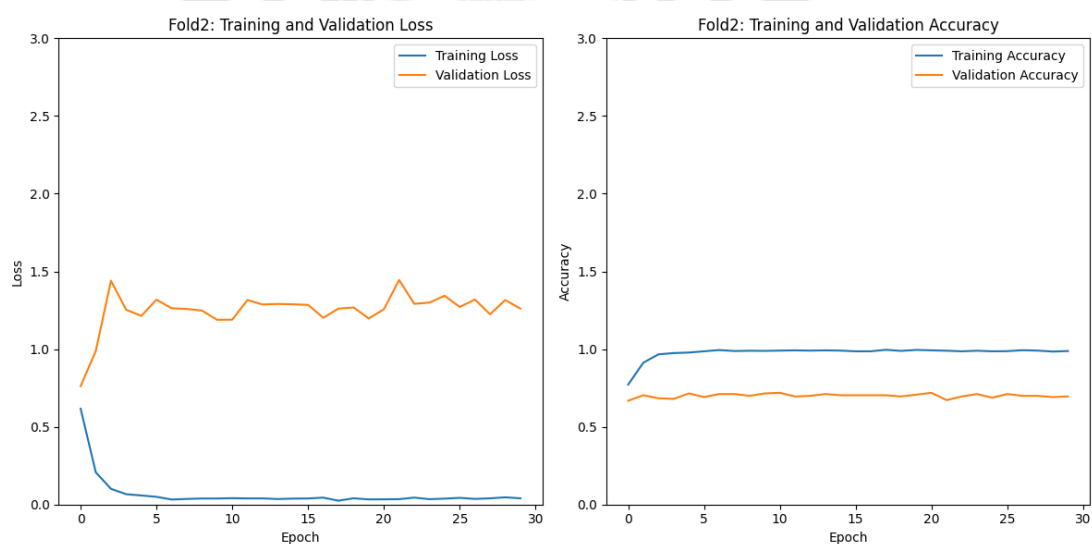
ภาพประกอบ 77 Confusion Matrix แบบจำลอง ResNet-50 Version 2 fold 1



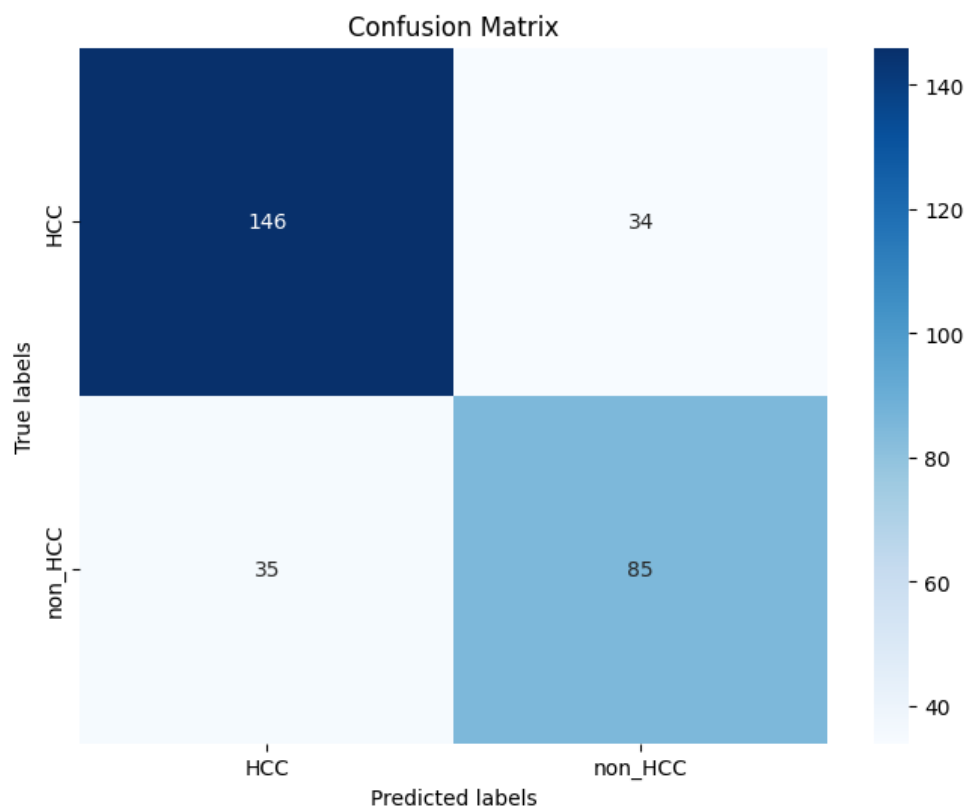
ภาพประกอบ 78 Learning Curve แบบจำลอง ResNet-50 Version 2 fold 1



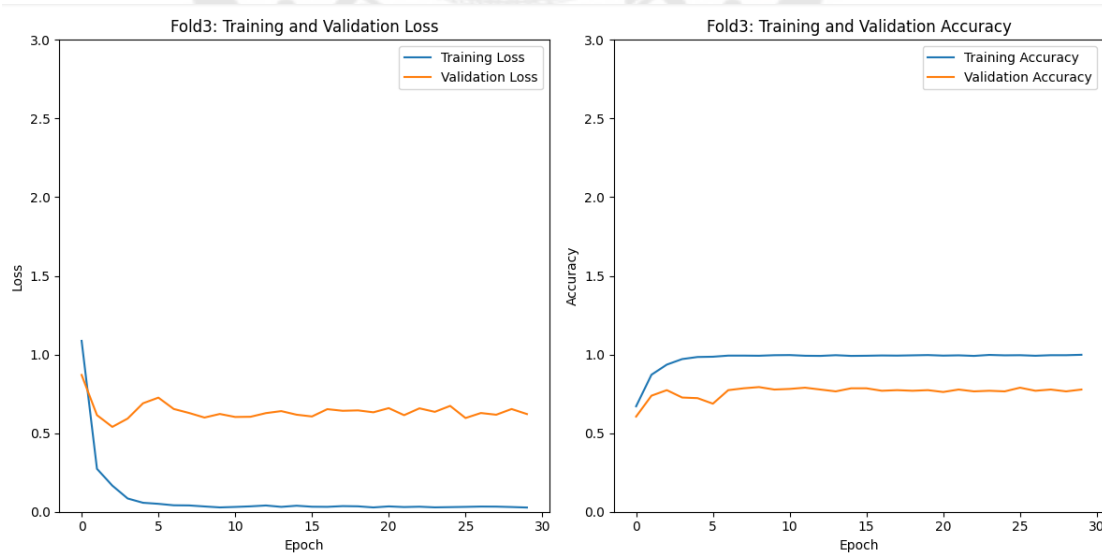
ภาพประกอบ 79 Confusion Matrix แบบจำลอง ResNet-50 Version 2 fold 2



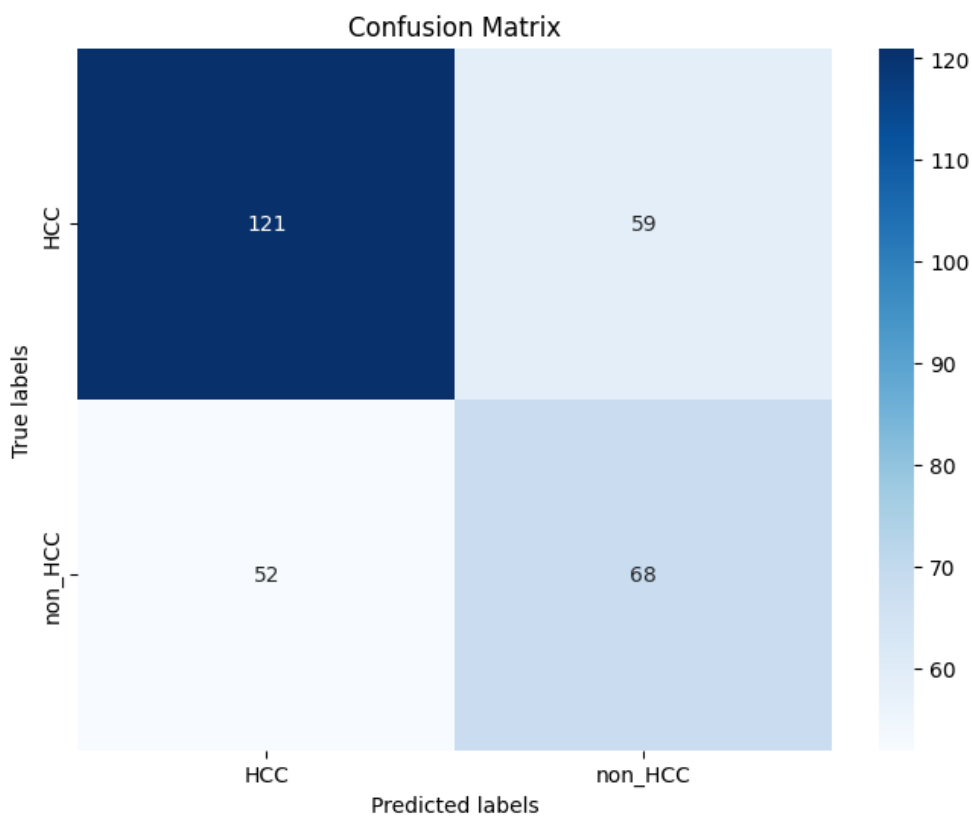
ภาพประกอบ 80 Learning Curve แบบจำลอง ResNet-50 Version 2 fold 2



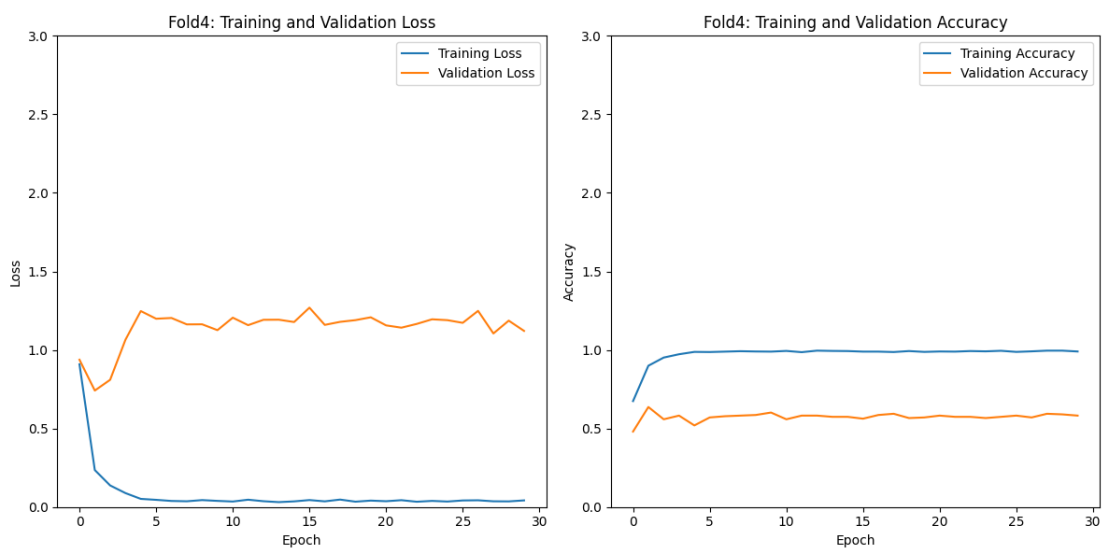
ภาพประกอบ 81 Confusion Matrix แบบจำลอง ResNet-50 Version 2 fold 3



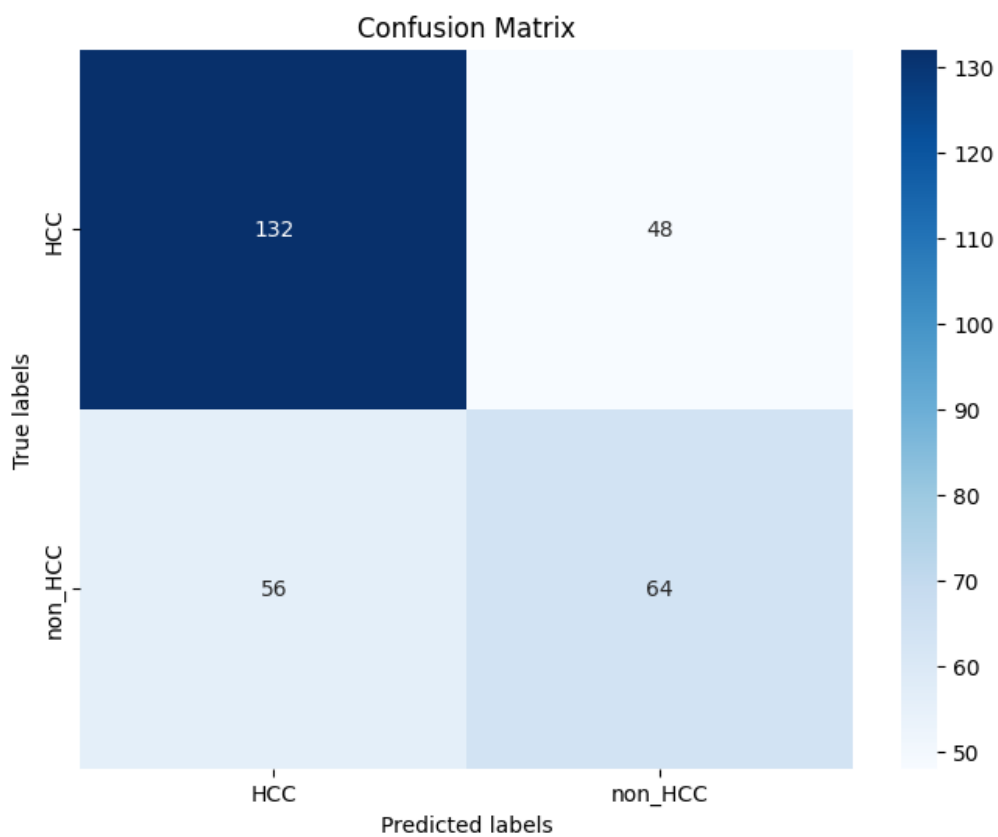
ภาพประกอบ 82 Learning Curve แบบจำลอง ResNet-50 Version 2 fold 3



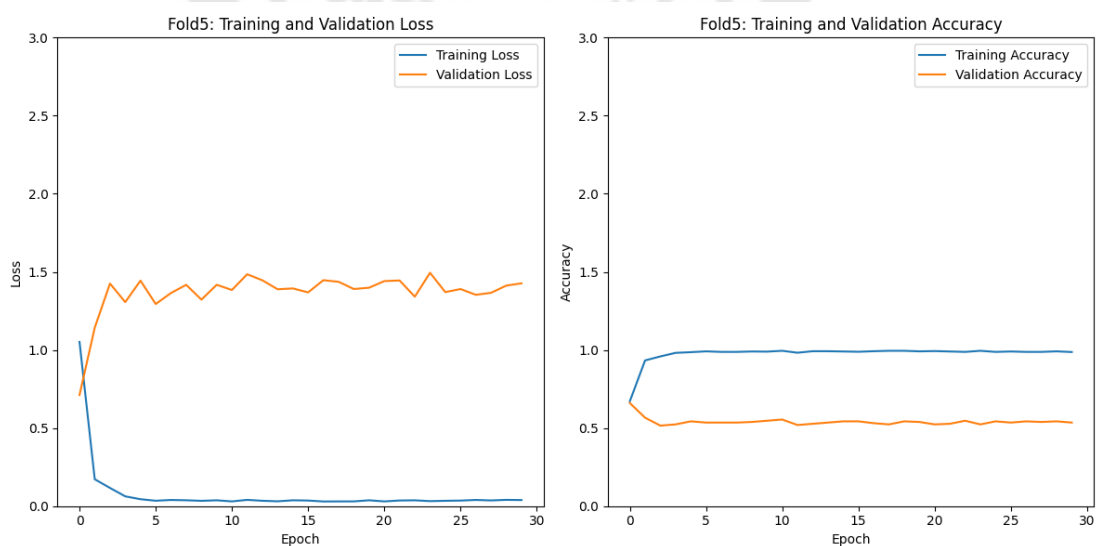
ภาพประกอบ 83 Confusion Matrix แบบจำลอง ResNet-50 Version 2 fold 4



ภาพประกอบ 84 Learning Curve แบบจำลอง ResNet-50 Version 2 fold 4

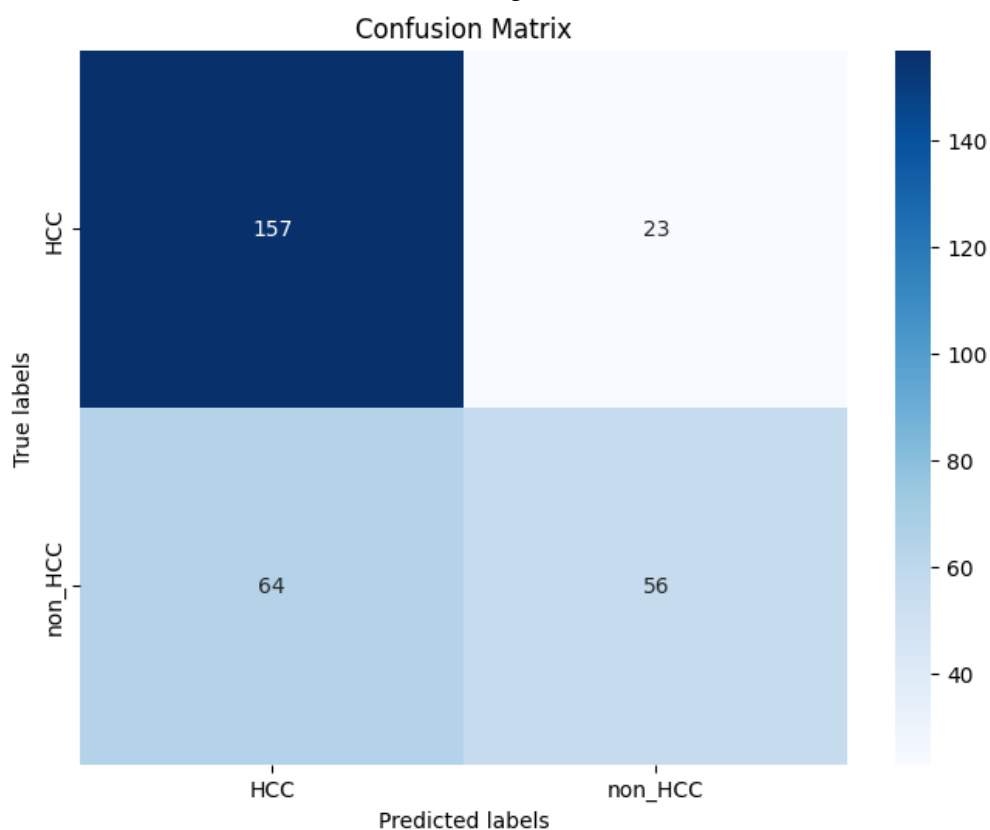


ภาพประกอบ 85 Confusion Matrix แบบจำลอง ResNet-50 Version 2 fold 5

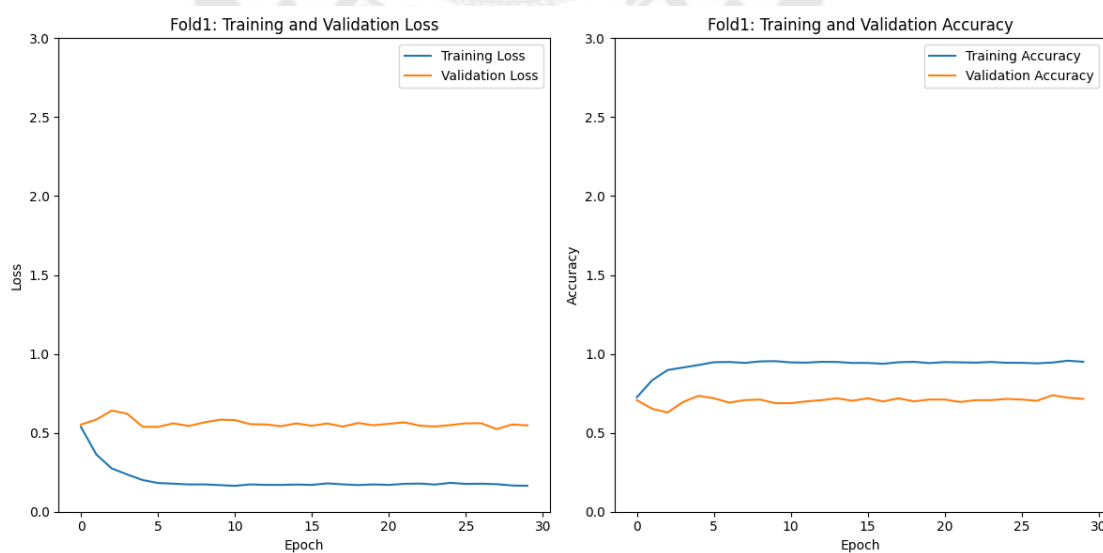


ภาพประกอบ 86 Learning Curve แบบจำลอง ResNet-50 Version 2 fold 5

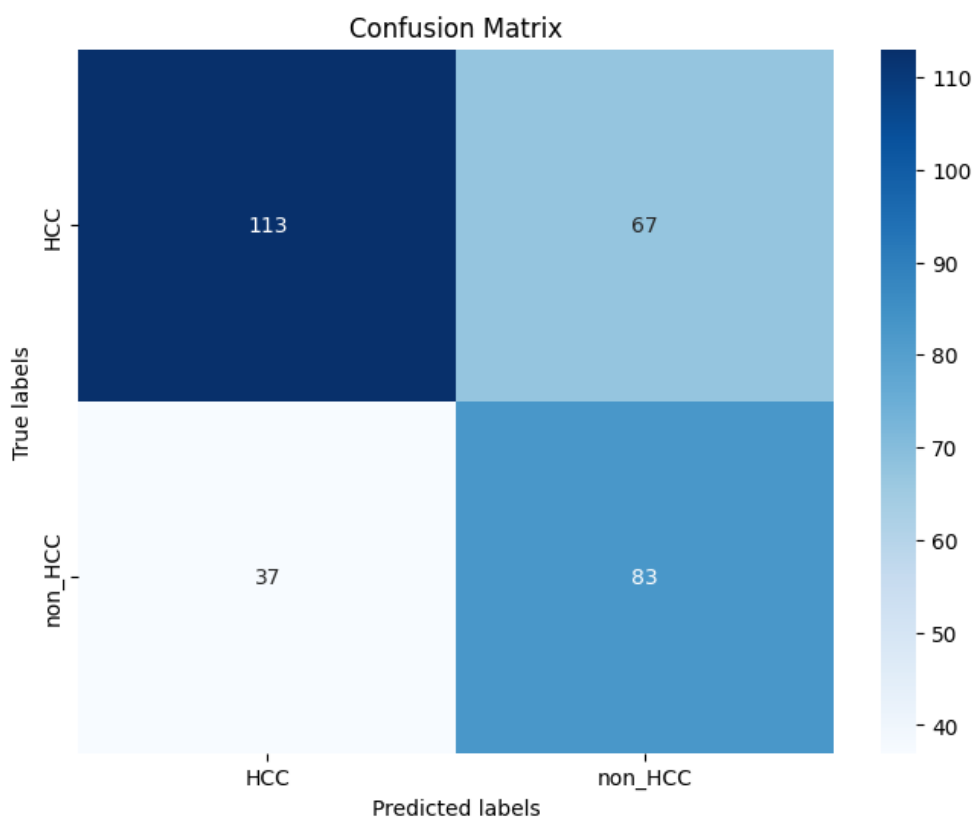
4.2.3 Confusion Matrix และ Learning Curve แบบจำลอง ResNet-50 Version 3



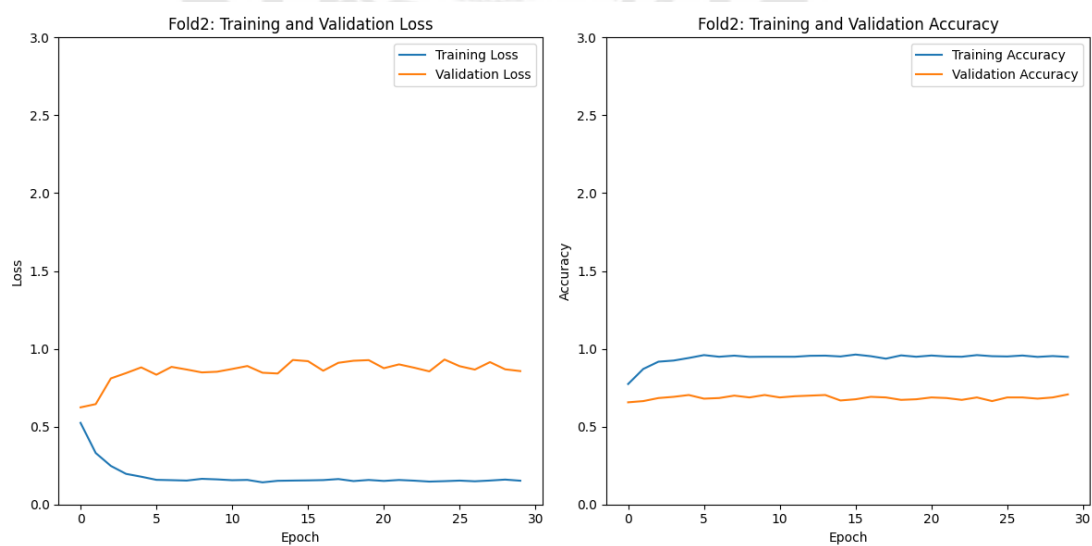
ภาพประกอบ 87 Confusion Matrix แบบจำลอง ResNet-50 Version 3 fold 1



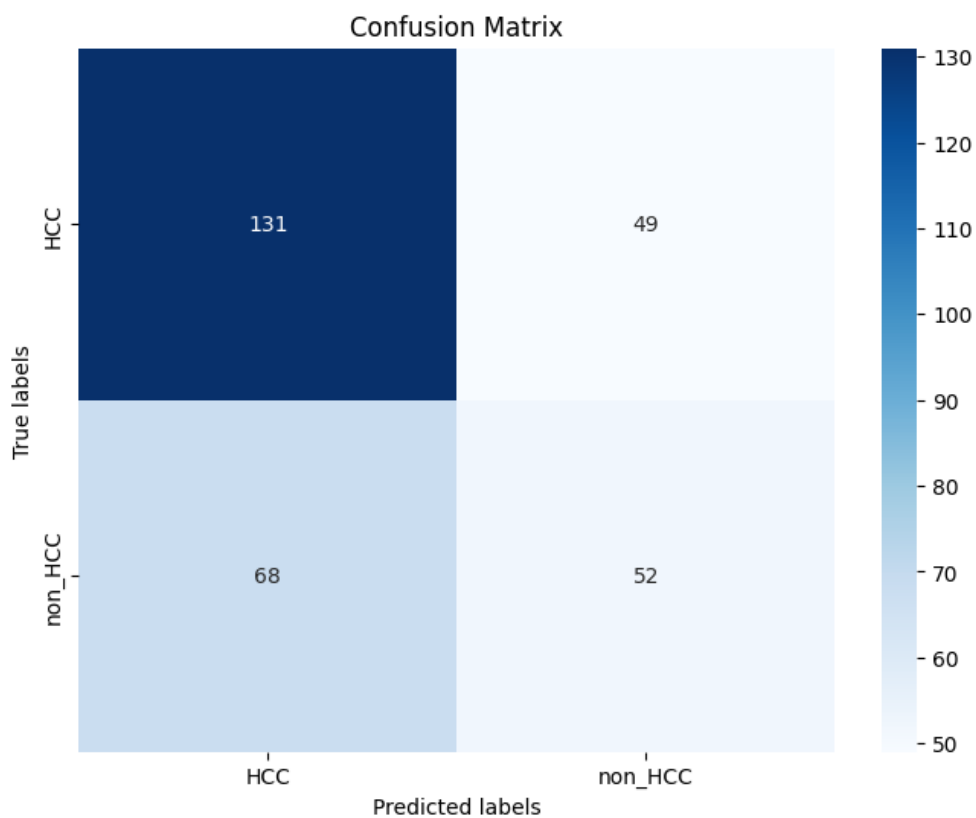
ภาพประกอบ 88 Learning Curve แบบจำลอง ResNet-50 Version 3 fold 1



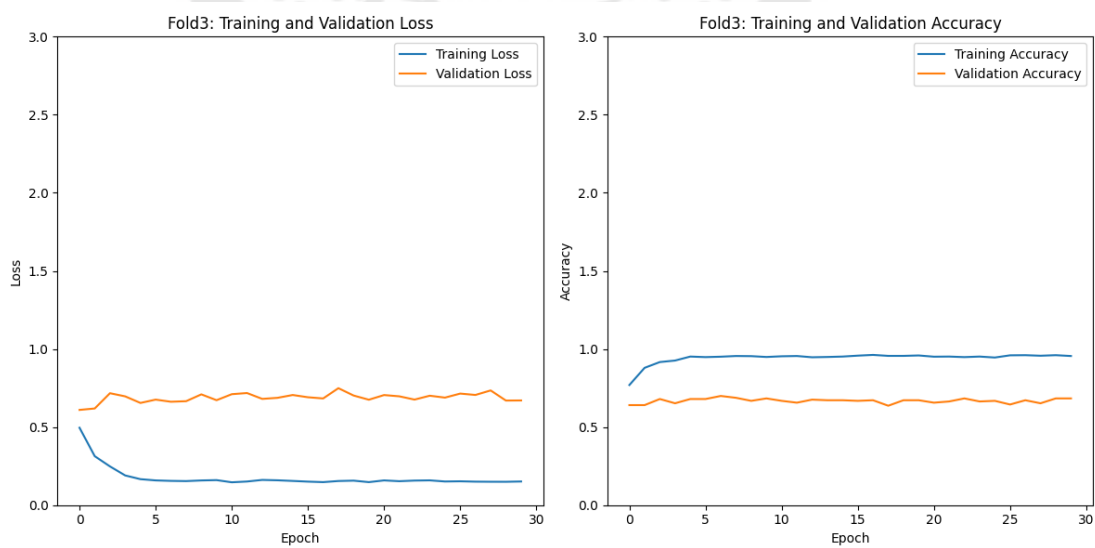
ภาพประกอบ 89 Confusion Matrix แบบจำลอง ResNet-50 Version 3 fold 2



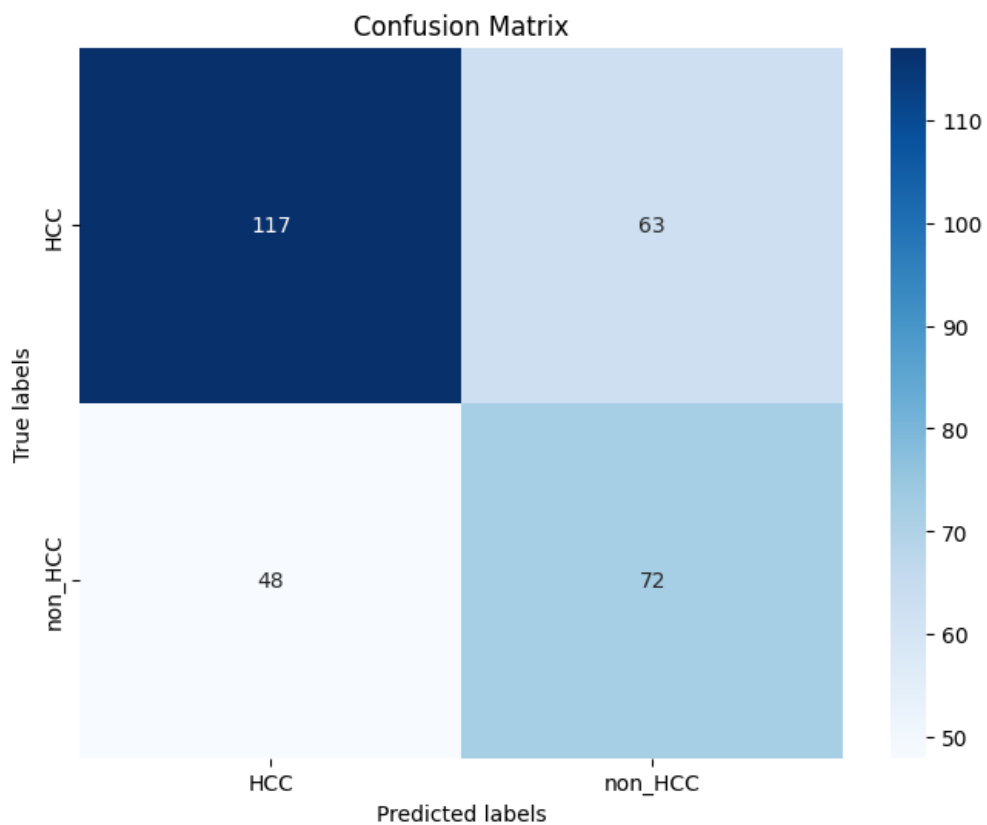
ภาพประกอบ 90 Learning Curve แบบจำลอง ResNet-50 Version 3 fold 2



ภาพประกอบ 91 Confusion Matrix แบบจำลอง ResNet-50 Version 3 fold 3



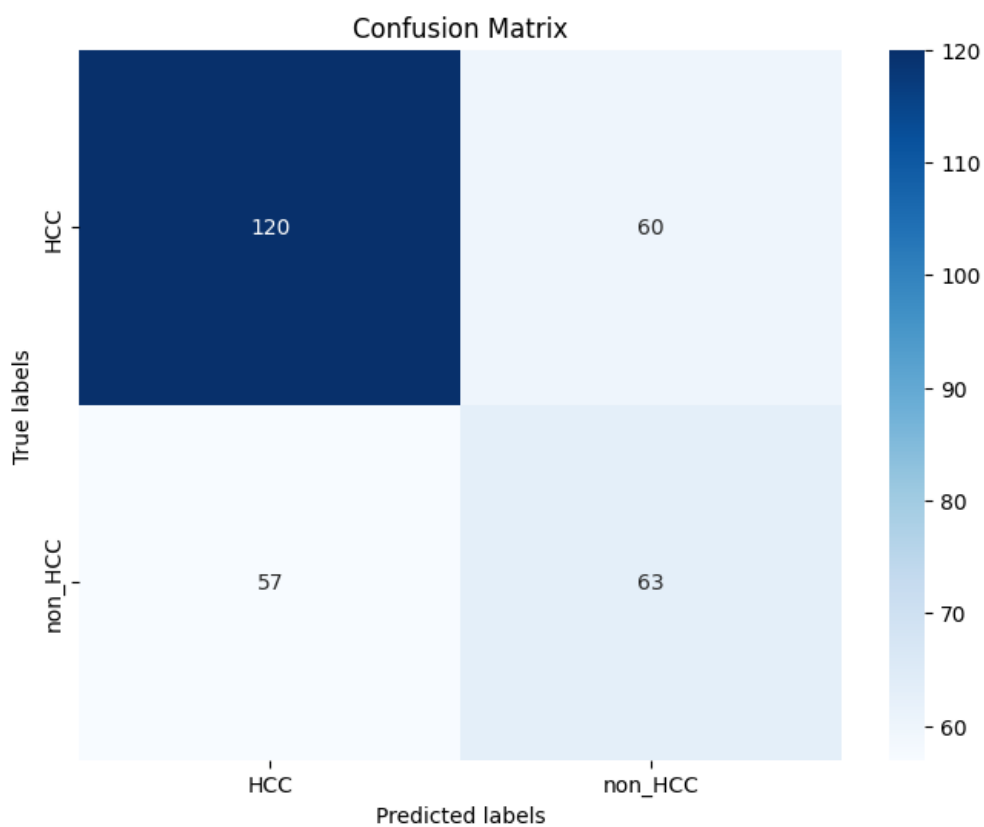
ภาพประกอบ 92 Learning Curve แบบจำลอง ResNet-50 Version 3 fold 3



ภาพประกอบ 93 Confusion Matrix แบบจำลอง ResNet-50 Version 3 fold 4



ภาพประกอบ 94 Learning Curve แบบจำลอง ResNet-50 Version 3 fold 4

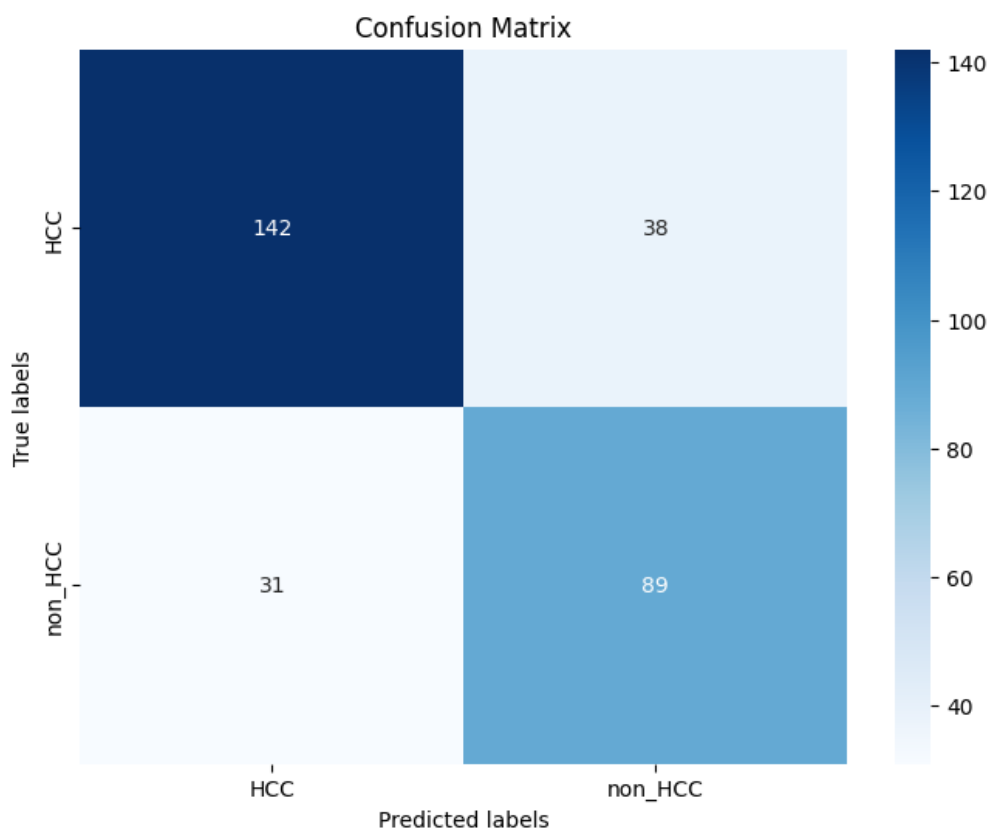


ภาพประกอบ 95 Confusion Matrix แบบจำลอง ResNet-50 Version 3 fold 5

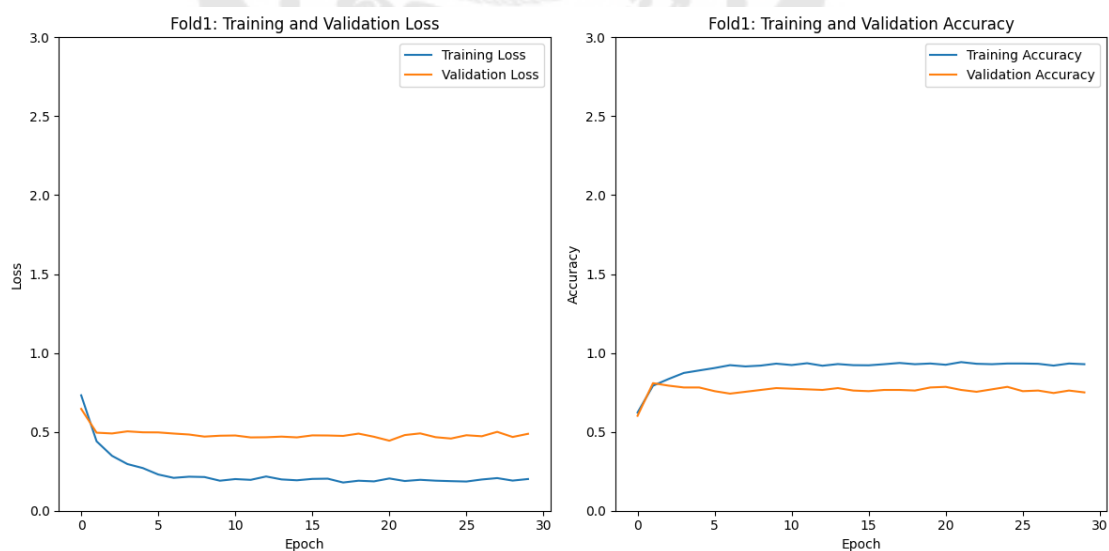


ภาพประกอบ 96 Learning Curve แบบจำลอง ResNet-50 Version 3 fold 5

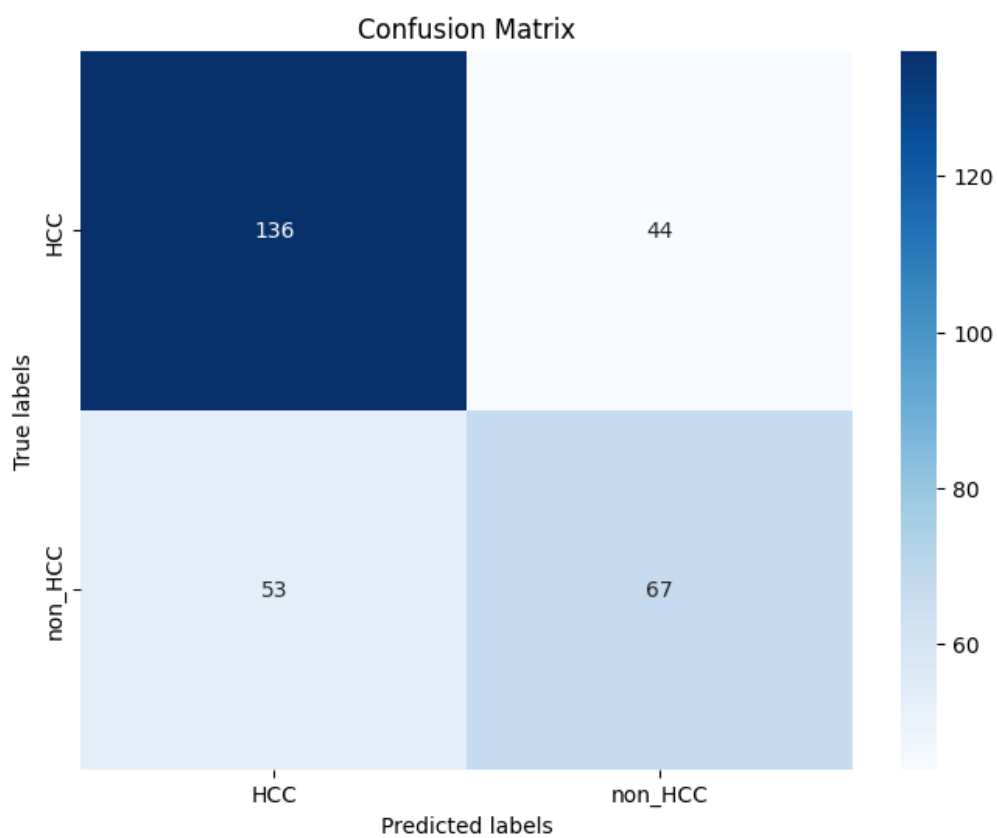
4.2.4 Confusion Matrix และ Learning Curve แบบจำลอง ResNet-50 Version 4



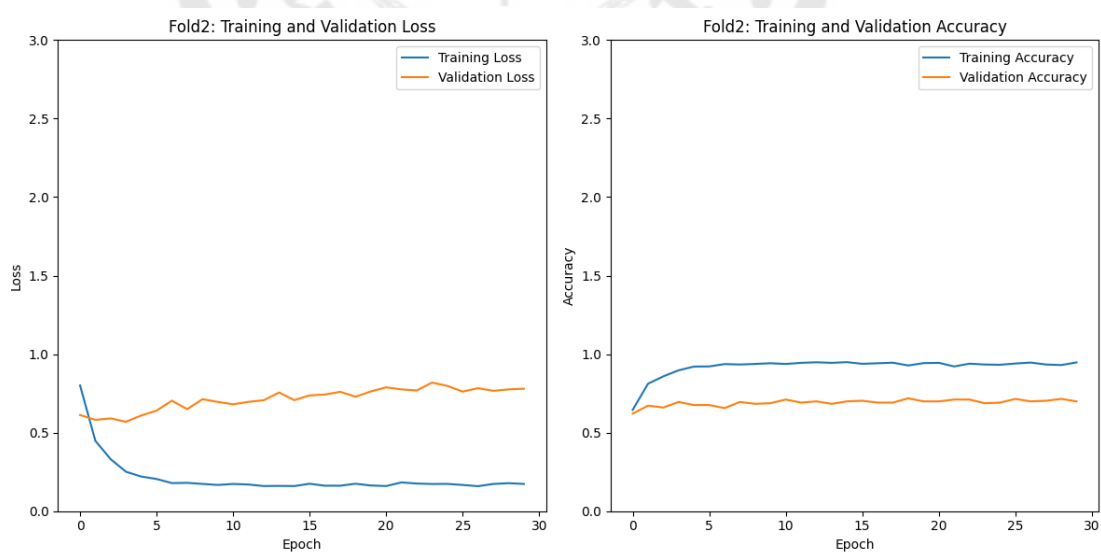
ภาพประกอบ 97 Confusion Matrix แบบจำลอง ResNet-50 Version 4 fold 1



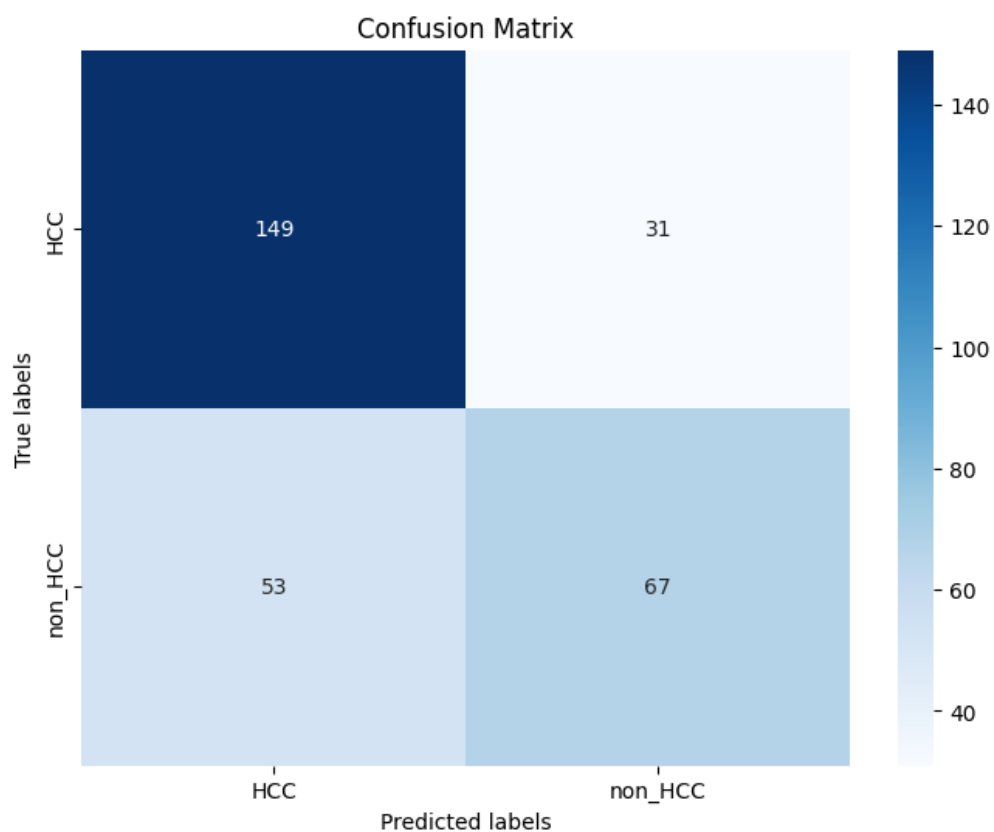
ภาพประกอบ 98 Learning Curve แบบจำลอง ResNet-50 Version 4 fold 1



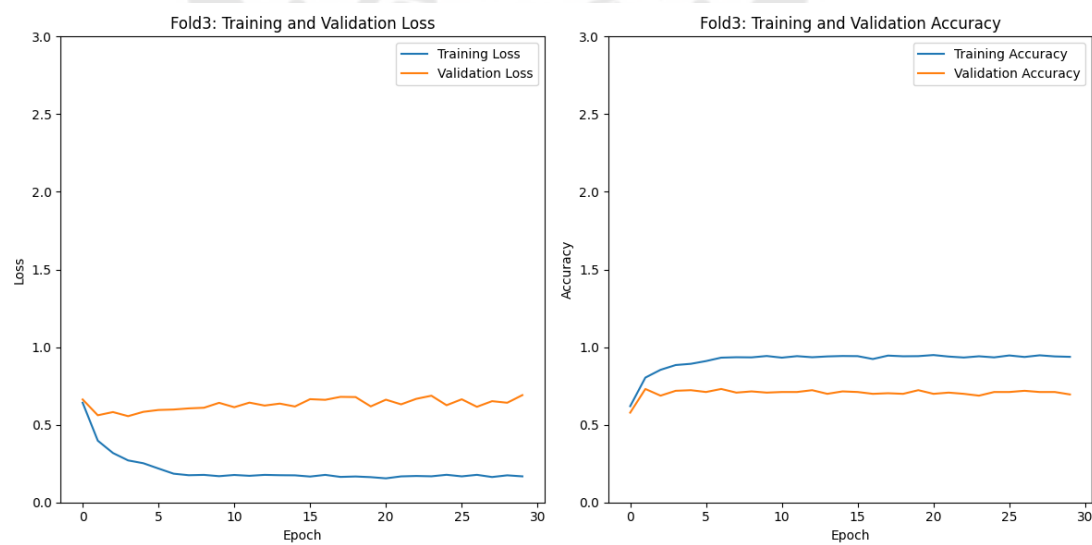
ภาพประกอบ 99 Confusion Matrix แบบจำลอง ResNet-50 Version 4 fold 2



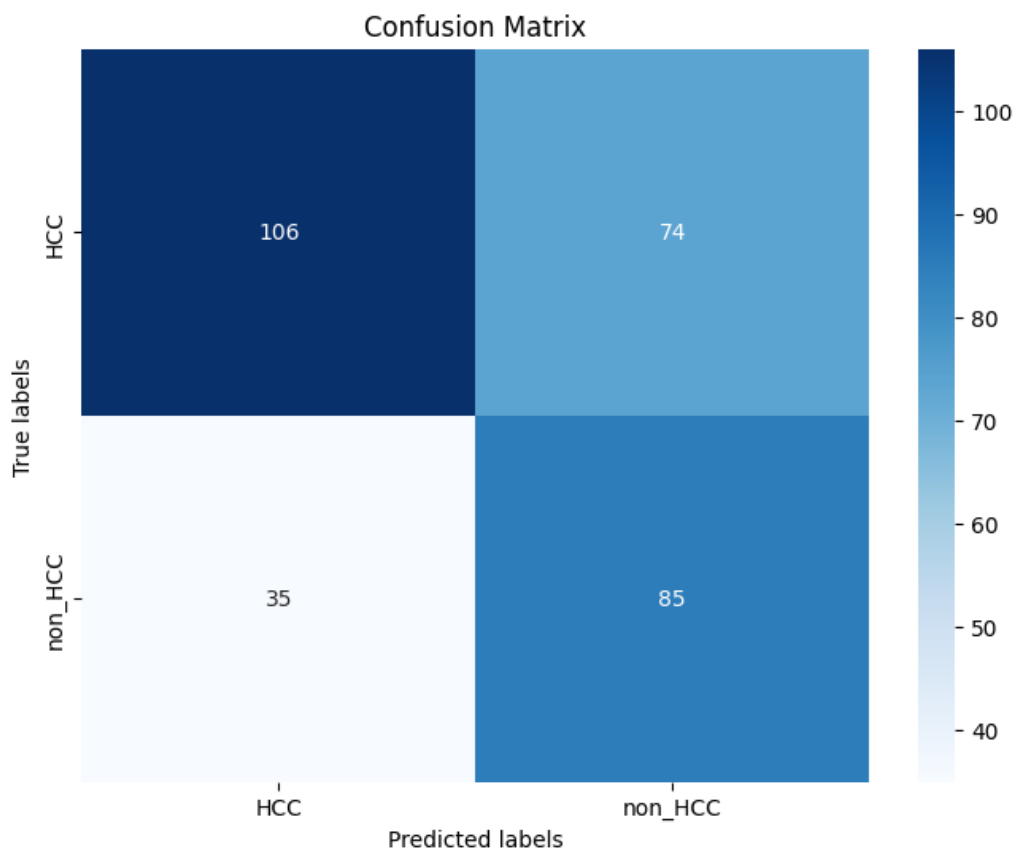
ภาพประกอบ 100 Learning Curve แบบจำลอง ResNet-50 Version 4 fold 2



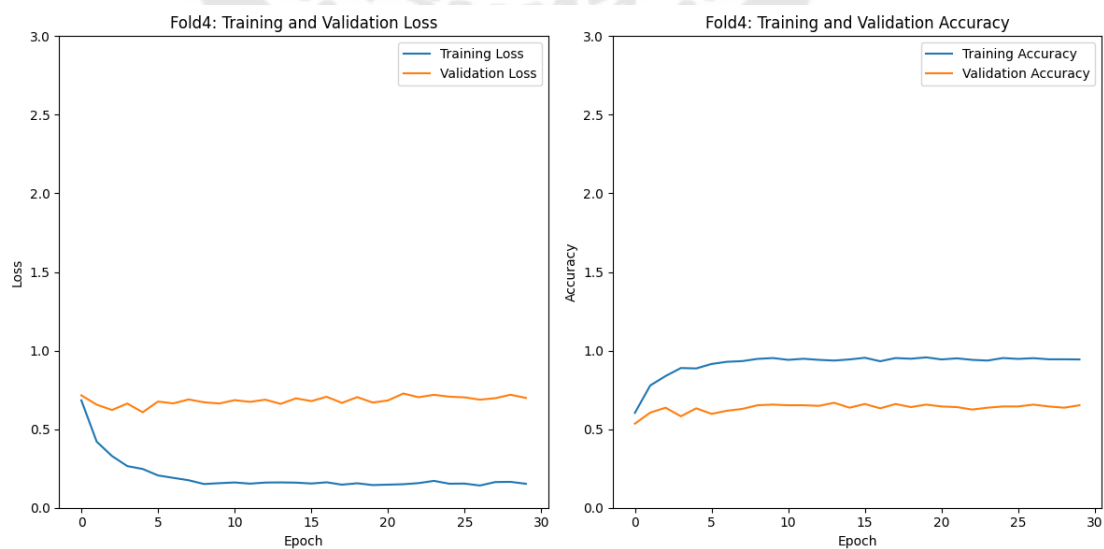
ภาพประกอบ 101 Confusion Matrix แบบจำลอง ResNet-50 Version 4 fold 3



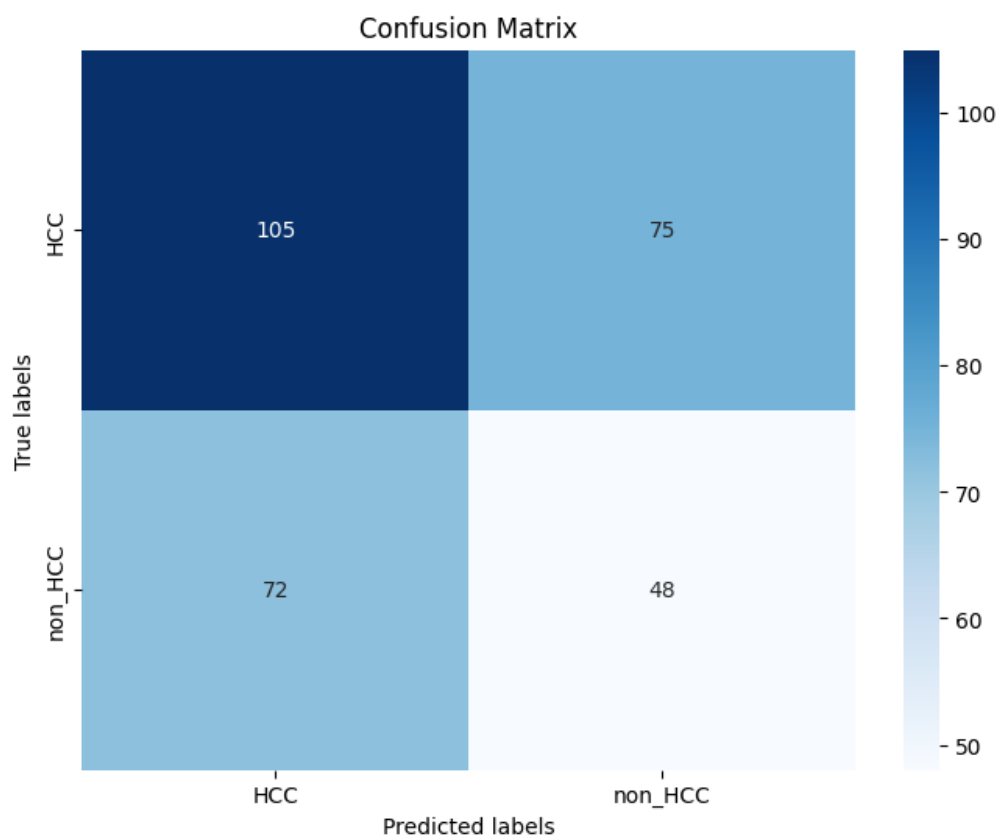
ภาพประกอบ 102 Learning Curve แบบจำลอง ResNet-50 Version 4 fold 3



ภาพประกอบ 103 Confusion Matrix แบบจำลอง ResNet-50 Version 4 fold 4



ภาพประกอบ 104 Learning Curve แบบจำลอง ResNet-50 Version 4 fold 4



ภาพประกอบ 105 Confusion Matrix แบบจำลอง ResNet-50 Version 4 fold 5



ภาพประกอบ 106 Learning Curve แบบจำลอง ResNet-50 Version 4 fold 5

นอกจากนี้เมื่อเปรียบเทียบตามการทำ Cross Validation โดยการสุ่มตัวอย่าง (Resampling) การแบ่งข้อมูลออกเป็นกลุ่มหลาย ๆ กลุ่ม เรียกว่า folds ซึ่งในที่นี้กำหนดให้มีค่าเท่ากับ 5 (5 folds) เรียงจากมากไปน้อยมีค่า Accuracy เฉลี่ย คือ fold 1 สูงสุด ที่ 0.72 fold 3 ที่ 0.69 fold 2 ที่ 0.66 fold 4 ที่ 0.65 และ fold 5 ที่ 0.56 ดังตาราง 13

ตาราง 13 ตารางเปรียบเทียบค่า Accuracy เฉลี่ยแต่ละ fold ข้อมูลของแบบจำลอง ResNet-50

Fold No.	ค่า Accuracy เฉลี่ย
1	0.72
2	0.66
3	0.69
4	0.65
5	0.56



ภาพประกอบ 107 สรุปผลการประเมินประสิทธิภาพแบบจำลองด้วยอัลกอริทึม ResNet-50

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

ในการทำวิจัยผู้วิจัยได้มีการประเมินประสิทธิภาพของแบบจำลองใน 2 โครงสร้าง คือ โครงสร้างโครงข่ายประสาทเทียมคอนโวลูชันแบบจำลองแบบ VGG16 และแบบ ResNet-50 เพื่อนำแบบจำลองทั้งสองมาเปรียบเทียบและสรุปผล โดยสามารถแบ่งหัวข้อในการสรุปผลได้ ดังนี้

5.1 สรุปผลการวิจัย

5.2 อภิปรายผลการวิจัย

5.3 ข้อเสนอแนะ

5.1 สรุปผลการวิจัย

ในงานวิจัยนี้ได้ศึกษาการสร้างแบบจำลองเพื่อจำแนกประเภท (Binary Classification) การเป็นโรคมะเร็งตับและไม่ใช่โรคมะเร็งตับจากภาพทางรังสีวินิจฉัย โดยใช้เทคนิคการเรียนรู้เชิงลึกด้วยโครงข่ายประสาทเทียมคอนโวลูชัน (Deep Learning with Convolutional Neural Network) ข้อมูลที่นำมาใช้ในการศึกษาวิจัยครั้งนี้มาจากฐานข้อมูลเปิด (Open database) ข้อมูลภาพถ่ายทางรังสีวินิจฉัยของ The Cancer Imaging Archive (TCIA) ซึ่งสามารถเข้าถึงได้แบบสาธารณะ โดยกลุ่มตัวอย่างผู้ป่วยที่เป็นโรคมะเร็งตับชนิด Hepatocellular carcinoma ใช้ชุดข้อมูล The Cancer Genome Atlas Liver Hepatocellular Carcinoma Collection (TCGA-LIHC) และสำหรับกลุ่มตัวอย่างผู้ป่วยที่ไม่เป็นโรคมะเร็งตับชนิด Hepatocellular carcinoma ใช้ชุดข้อมูลทั้งหมด 2 ชุด คือ The Cancer Genome Atlas Cervical Kidney Renal Papillary Cell Carcinoma Collection (TCGA-KIRP) ซึ่งเป็นข้อมูลผู้ป่วยที่เป็นโรคมะเร็งไต และอีกชุดข้อมูลหนึ่งจะเป็นข้อมูลผู้ป่วยที่เป็นโรคมะเร็งตับอ่อน The Clinical Proteomic Tumor Analysis Consortium Pancreatic Ductal Adenocarcinoma Collection (CPTAC-PDA)

จากการพัฒนาแบบจำลองโครงสร้างโครงข่ายประสาทเทียมคอนโวลูชันแบบ VGG16 และแบบ ResNet-50 นั้น พบว่า ผลการทดลองของโครงสร้างโครงข่ายประสาทเทียมคอนโวลูชันของแบบ VGG16 ทั้งหมด 4 ชุด เมื่อพิจารณาประสิทธิภาพของแบบจำลองด้วยเมตริกซ์ Accuracy ที่มุ่งเน้นเรื่องอัตราส่วนของตัวอย่างที่แบบจำลองทำนายถูกต้องทั้งหมด เรียงค่า Accuracy จากมากไปน้อยจะได้ VGG16 Version 2 ที่ 0.63 VGG16 Version 3 ที่ 0.59 VGG16 Version 4 ที่ 0.58 และ VGG16 Version 1 ที่ 0.54 และในผลจากการทดลองโครงสร้างโครงข่ายประสาทเทียมคอนโวลูชันแบบ ResNet-50 จากทั้งหมด 4 ชุดผลการทดลอง พบว่า

เมื่อประเมินประสิทธิภาพของแบบจำลองด้วยเมตริกซ์ Accuracy เช่นเดียวกัน โดยเรียงค่า Accuracy จากมากไปน้อยจะได้ ResNet-50 Version 2 ที่ 0.69 ResNet-50 Version 4 ที่ 0.66 ResNet-50 Version 3 ที่ 0.64 และ ResNet-50 Version 1 ที่ 0.63

และหากวิเคราะห์ด้วยเมตริกซ์ False Negatives (FN) ซึ่งในทางการแพทย์ เราต้องการลดค่า False Negatives (FN) ให้น้อยที่สุดเท่าที่จะเป็นไปได้ เนื่องจาก False Negatives คือการระบุว่าผู้ป่วยไม่มีโรค ทั้งที่จริงแล้วผู้ป่วยเป็นโรค ซึ่งอาจทำให้พลาดการรักษาหรือไม่ได้รับการรักษาตามที่ต้องการ ส่งผลกระทบต่อสุขภาพของผู้ป่วยและสังคมโดยรวมอย่างมากนั้น พบว่า VGG16 Version 2 มีค่า False Negatives (FN) น้อยกว่า ResNet-50 Version 2 โดยมีผู้ป่วยที่ถูกทายเป็น False Negatives (FN) อยู่ที่ 32 รูป และ 54 รูป ตามลำดับ

จึงสามารถสรุปผลได้ว่า สามารถนำแบบจำลองจากงานวิจัยนี้ไปใช้ในการจำแนกการเป็นโรคมะเร็งตับและไม่ใช่โรคมะเร็งตับจากภาพการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) ได้ เพื่อเป็นข้อมูลสนับสนุนการตัดสินใจสำหรับรังสีแพทย์ให้สามารถวินิจฉัย วางแนวทางการรักษาผู้ป่วยได้อย่างเหมาะสม และรวดเร็วทันต่อการรักษา โดยโครงสร้างโครงข่ายประสาทเทียมคอนโวลูชันแบบ ResNet-50 มีประสิทธิภาพมากกว่าแบบ VGG16 ในทุกผลการทดลองทั้งหมด 4 ชุด และผลการทดลองชุดที่ 2 (ResNet-50 Version 2) มีค่า Accuracy มากที่สุด โดยมีค่าเท่ากับ 0.69 และหากวิเคราะห์ด้วยเมตริกซ์ False Negatives (FN) นั้นแบบจำลอง VGG16 Version 2 มีค่า False Negatives น้อยกว่า ResNet-50 Version 2 โดยมีผู้ป่วยที่ถูกทายเป็น False Negatives อยู่ที่ 32 รูป และ 54 รูป ตามลำดับ

ปัจจัยและข้อจำกัดที่ส่งผลต่อการศึกษางานวิจัยครั้งนี้ คือ ข้อมูลที่นำมาใช้ในการศึกษานี้มาจากฐานข้อมูลภาพการแพทย์ขนาดใหญ่ที่สามารถเข้าถึงได้โดยสาธารณะของ The Cancer Imaging Archive (TCIA) ที่ได้รวบรวมข้อมูลภาพถ่ายทางการแพทย์ต่าง ๆ ในผู้ป่วยที่เป็นโรคมะเร็ง มีความน่าเชื่อถือ โดยมีวัตถุประสงค์ในการสนับสนุนการศึกษาและการวิจัยทางการแพทย์ซึ่งเก็บรวบรวมข้อมูลภาพการแพทย์จากหลายแหล่งทั่วโลก เพื่อให้ผู้ใช้สามารถเข้าถึงและใช้ข้อมูลเหล่านี้ในงานวิจัยได้อย่างสะดวก แต่ทั้งนี้ ตามวัตถุประสงค์ของงานวิจัยที่ได้มีการกำหนดเกณฑ์ในการเลือกใช้เฉพาะในระยะ T1 (Pulse sequence) เท่านั้น ต้องเป็นผู้ป่วยที่ไม่เคยได้รับการผ่าตัดก่อน เนื่องจากสนใจกลุ่มตัวอย่างผู้ป่วยที่มีการคัดกรองและติดตามโรค และได้มีการกำหนดเกณฑ์การคัดเลือกภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) ให้ที่มีพื้นที่อวัยวะตับภายในรูปภาพถ่ายการตรวจนั้น ๆ มากกว่าร้อยละ 50 เท่านั้น จึงจะสามารถนำรูปภาพถ่ายมาใช้ในการสร้างแบบจำลองการจำแนกผู้ป่วยที่เป็นโรคมะเร็งตับและไม่ใช่โรคมะเร็งตับได้ ส่งผลให้

ภาพถ่ายที่ใช้ในการสร้างแบบจำลองในตอนสุดท้ายมีปริมาณรูปภาพที่น้อยลงจากเกณฑ์ดังกล่าว จึงอาจส่งผลกระทบต่อแบบจำลองที่ได้มีประสิทธิภาพลดลงตามจำนวนรูปภาพที่ลดลง เพราะในเทคนิคการเรียนรู้เชิงลึกด้วยโครงข่ายประสาทเทียมคอนโวลูชัน (Deep Learning with Convolutional Neural Network) นั้น จำเป็นต้องใช้ข้อมูลปริมาณรูปภาพจำนวนมาก เพื่อใช้ในการฝึกสอนการเรียนรู้ให้กับแบบจำลองมากขึ้น

5.2 อภิปรายผล

งานวิจัยนี้ได้ผลการประเมินประสิทธิภาพของแบบจำลองโครงสร้างโครงข่ายประสาทเทียมคอนโวลูชันแบบ VGG16 และแบบ ResNet-50 ด้วยเมตริกซ์ Accuracy สามารถสรุปผลได้ว่า โครงสร้างโครงข่ายประสาทเทียมคอนโวลูชันแบบ ResNet-50 ของการทดลองชุดที่ 2 (ResNet-50 Version 2) นั้น มีประสิทธิภาพมากกว่าโครงสร้างโครงข่ายประสาทเทียมคอนโวลูชันแบบ VGG16 ในทุกชุดการทดลองทั้งหมด 4 ชุดที่ค่า Accuracy เท่ากับ 0.69 และหากวิเคราะห์ด้วยเมตริกซ์ False Negatives (FN) ตามหลักทางการแพทย์ที่มองว่าต้องการลดค่า False Negatives (FN) ให้น้อยที่สุดเท่าที่จะเป็นไปได้นั้น VGG16 Version 2 มีค่า False Negatives (FN) น้อยกว่า ResNet-50 Version 2 โดยมีผู้ป่วยที่ถูกทายเป็น False Negatives (FN) อยู่ที่ 32 รูป และ 54 รูป ตามลำดับ

จากผลการทดลองมีความสอดคล้องกับเอกสารและงานวิจัยที่เกี่ยวข้องเรื่อง Deep convolutional neural network applied to the liver imaging reporting and data system (LI-RADS) version 2014 category classification: a pilot study โดยในงานวิจัยนี้มีการสร้างแบบจำลองด้วยโครงข่ายประสาทเทียม (CNN) ใช้โครงสร้างเป็น VGG16 และใช้ Weight เป็น "Imagenet" เช่นเดียวกับการศึกษาวิจัยในครั้งนี้ ในข้อมูล LR-1/2 และ LR-3 เทียบได้กับกลุ่มผู้ป่วยที่ไม่ได้เป็นโรคมะเร็งตับ ส่วน LR-4 และ LR-5 เทียบได้กับกลุ่มผู้ป่วยที่เป็นโรคมะเร็งตับ อ้างอิงการจัดกลุ่มตามงานวิจัยเรื่อง Deep learning LI-RADS grading system based on contrast enhanced multiphase MRI for differentiation between LR-3 and LR-4/LR-5 liver tumors มีการวัดประสิทธิภาพของแบบจำลอง (Model Evaluation) ที่ประเมินประสิทธิภาพของแบบจำลองด้วยเมตริกซ์ Accuracy เช่นเดียวกัน ได้ค่า Accuracy โดยรวมทุก LI-RADS (Overall) เท่ากับ 0.64 ซึ่งมีความใกล้เคียงกันกับค่า Accuracy ของแบบจำลอง VGG16 Version 2 ที่ได้จากการศึกษาวิจัยในครั้งนี้เท่ากับ 0.63 และนอกจากนี้ในงานวิจัยเรื่อง Comparison of fine-tuning strategies for transfer learning in medical image classification ที่ระบุว่าแบบจำลองโครงสร้าง ResNet สามารถใช้ในงานภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า

(MRI) ได้ ถึงแม้ว่างานที่ได้ทดลองใช้ในงานวิจัยนั้นจะเป็นโรคอื่น ๆ ที่ไม่ใช่โรคมะเร็งระดับ แต่จากการศึกษานี้ก็ได้แสดงให้เห็นแล้วว่าโครงสร้างแบบจำลอง ResNet-50 Version 2 นั้น สามารถใช้ในการจำแนกการเป็นโรคมะเร็งระดับได้เช่นเดียวกัน

ทั้งนี้ ในแบบจำลองทั้ง VGG16 Version 2 และ ResNet-50 Version 2 นั้น มีจุดร่วมที่เหมือนกันอย่างหนึ่ง คือ มีการเพิ่มชั้นเชื่อมต่อแบบสมบูรณ์ชั้นที่ 1 (Fully connected Layer 1) และชั้นที่ 2 (Fully connected Layer 2) เช่นเดียวกัน ตรงนี้อาจเป็นปัจจัยสำคัญส่วนหนึ่งที่ทำให้แบบจำลองที่ได้ในครั้งนี้นี้มีความแม่นยำมากขึ้น ดังนั้น ในการศึกษาครั้งถัดสามารถนำโครงสร้างแบบจำลองในลักษณะนี้ไปประยุกต์ใช้ได้เช่นกันในงานการจำแนกการเป็นโรคมะเร็งระดับ

และในการสร้างแบบจำลองในครั้งนี้จะพบว่า ขั้นตอนที่ใช้เวลามากที่สุดในการดำเนินงาน คือ การเตรียมรูปภาพซึ่งในการศึกษานี้ได้มีการเพิ่มขั้นตอนกระบวนการปรับข้อมูลภาพ (Image Augmentation) ด้วย เพื่อเป็นการเสริมความแม่นยำของแบบจำลองให้สามารถจำแนกการเป็นโรคมะเร็งระดับและไม่เป็นโรคมะเร็งระดับได้จากข้อมูลภาพที่หลากหลายรูปแบบ เพราะในการทำงานจริงเครื่องที่ใช้ในการตรวจอาจมาจากหลายยี่ห้อ และอาจมีการตั้งค่าการตรวจที่แตกต่างกันตามขั้นตอนของการตรวจรักษาของแพทย์ในครั้งนั้น ๆ (Protocol) กระบวนการที่เพิ่มมาตรงนี้จะทำให้แบบจำลองมีประสิทธิภาพในการจำแนกข้อมูลทดสอบที่ไม่เคยฝึกสอนมาก่อนได้มากยิ่งขึ้น

นอกจากทั้งหมดที่กล่าวมานี้ผลลัพธ์ที่ได้ของแบบจำลองนี้ยังสามารถพัฒนา และต่อยอดในการศึกษาวิจัยครั้งถัดไปได้ ดังนั้น ผู้วิจัยจึงทำการวิเคราะห์ข้อผิดพลาด (Error Analysis) ในเชิงลึกที่มากขึ้น ดังนี้

5.2.1 เพิ่มเทคนิคการปรับปรุงคุณภาพรูปภาพ (Image Enhancement)

จากการศึกษางานวิจัยที่เกี่ยวข้องเพิ่มเติม พบว่า ในงานวิจัยที่มีการเพิ่มขั้นตอนการปรับปรุงคุณภาพรูปภาพนั้น สามารถส่งผลให้แบบจำลองจำแนกประเภทข้อมูลรูปภาพได้ดีมากยิ่งขึ้น อีกทั้งในชุดข้อมูลกลุ่มตัวอย่างที่นำมาศึกษาในงานวิจัยครั้งนี้ เนื่องจากเป็นฐานข้อมูลสาธารณะขนาดใหญ่ ดังนั้น คุณภาพภาพถ่ายที่ได้นั้น ย่อมมีความหลากหลายเป็นอย่างมาก เพราะมีการเก็บรวบรวมมาจากหลายสถานที่การให้บริการตรวจรักษาผู้ป่วย หากสามารถเพิ่มเทคนิคนี้เข้ามาก่อนเข้าขั้นตอนการสร้างแบบจำลองน่าจะส่งผลทำให้แบบจำลองมีประสิทธิภาพที่ดีขึ้น เพราะสามารถสกัดคุณลักษณะของข้อมูลภายในภาพได้ดีมากยิ่งขึ้น

5.2.2 การแบ่งข้อมูลภาพ (Image segmentation)

ด้วยวัตถุประสงค์ของงานวิจัยนี้ เป็นการจำแนกการเป็นโรคมะเร็งระดับและไม่เป็นโรคมะเร็งระดับ ข้อมูลภาพถ่ายในการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) ที่ใช้จะเป็นการตรวจ

บริเวณช่องท้องส่วนบน (MRI Upper Abdomen) ซึ่งจะมีอวัยวะอื่น ๆ ที่ไม่ใช่อวัยวะตับอยู่ในภาพถ่ายการตรวจครั้งนั้น ๆ ด้วย เช่น อวัยวะกระเพาะอาหาร (Stomach) อวัยวะลำไส้ใหญ่ (Colon) อวัยวะไต (Kidney) เป็นต้น หรือในบางกรณีที่เป็นผู้ป่วยมีภาวะอ้วนร่วมด้วยอาจมีชั้นไขมันในช่องท้องที่ค่อนข้างหนาและมีปริมาณมากก็อาจทำให้ถูกบดบังการมองเห็นอวัยวะตับในช่องท้องได้ยากขึ้นด้วย

ดังนั้น หากสามารถเพิ่มขึ้นขั้นตอนการเตรียมรูปภาพก่อนนำไปใช้ในการสร้างแบบจำลองได้ โดยการตัดแบ่งข้อมูลรูปภาพให้เหลือเฉพาะอวัยวะตับตามที่สนใจ จะช่วยให้แบบจำลองสามารถเรียนรู้ในการตรวจจับรูปแบบ (Pattern) ของการเป็นโรคมะเร็งตับและไม่เป็นโรคมะเร็งตับได้ชัดเจนและดีมากยิ่งขึ้นได้ เพราะได้ตัดสิ่งรบกวนต่าง ๆ คือ อวัยวะอื่นๆ ในบริเวณช่องท้องส่วนบนภายในภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้าออกไปแล้ว

5.2.3 การปรับแต่ง (Fine-tuning)

จากผลการทดลองพบว่า โครงสร้างแบบ ResNet-50 ของผลการทดลองชุดที่ 2 (ResNet-50 Version 2) นั้น มีประสิทธิภาพมากกว่าโครงสร้างแบบ VGG16 จริง แต่ทั้งนี้ยังสามารถปรับปรุงประสิทธิภาพของแบบจำลองให้ดีขึ้นได้ เพราะยังสามารถปรับแต่งโครงสร้างของแบบจำลองต้นแบบ (Architecture model) โดยการเพิ่มหรือลดชั้นโครงสร้างต่าง ๆ ได้ หรืออาจเพิ่มเติมเรื่องการปรับค่าพารามิเตอร์ต่าง ๆ (Hyperparameter) ที่ใช้ในการควบคุมกระบวนการฝึกสอนแบบจำลอง โดยเป็นตัวแปรที่ไม่ได้ถูกเรียนรู้โดยตรงจากข้อมูล แต่จะถูกกำหนดค่าโดยผู้พัฒนาแบบจำลอง ก่อนที่จะเริ่มการฝึกโมเดล ซึ่งล้วนมีผลต่อประสิทธิภาพของโมเดลที่ถูกฝึกขึ้นมา โดยพอจะยกตัวอย่างคร่าว ๆ ได้ดังนี้

5.2.3.1 การปรับขนาดของกลุ่มข้อมูลในการฝึกสอน (Batch size)

เป็นจำนวนข้อมูลที่ถูกนำเข้าไปในแบบจำลองแต่ละครั้งในช่วงการฝึกสอน การกำหนดขนาดของกลุ่มข้อมูลในการฝึกสอน (Batch size) ซึ่งการเลือกขนาดที่เหมาะสมสามารถช่วยให้การฝึกแบบจำลองมีประสิทธิภาพที่ดีมากขึ้น โดยทั่วไปแล้วมีมักการเลือกขนาดของกลุ่มข้อมูลในการฝึกสอน (Batch size) ให้เป็นค่าที่ เช่น 32, 64, 128 หรือ 256 เป็นต้น ซึ่งจำนวนนี้อาจมีการปรับเปลี่ยนตามขนาดของข้อมูลและขนาดของแบบจำลองด้วย

5.2.3.2 การปรับรอบของการฝึก (Epoch)

เป็นจำนวนรอบที่แบบจำลองถูกฝึกสอนด้วยชุดข้อมูลทั้งหมด (Training data) โดยใช้ข้อมูลซ้ำไปหลายรอบ เพื่อเป็นการปรับค่าพารามิเตอร์ให้เหมาะสม การกำหนดจำนวนรอบการฝึกสอนนี้มีผลต่อประสิทธิภาพและความแม่นยำของแบบจำลองด้วยเช่นกัน เพราะรอบของ

การฝึกสอนที่มากเกินไปอาจทำให้เกิดปัญหา Overfitting ที่แบบจำลองทำงานได้ดี มีความแม่นยำสูง ทำนายแม่นยำกับข้อมูลฝึกสอน (Training data) แต่ทำงานได้แย่ ไม่แม่นยำ เมื่อเจอข้อมูลใหม่ ๆ หรือ ข้อมูลที่ไม่เคยเห็นมาก่อน (Testing data) ขึ้น

5.2.3.3 การปรับอัตราการเรียนรู้ (Learning rate)

เป็นค่าที่ใช้ในการปรับค่าพารามิเตอร์ (Parameters) ของแบบจำลองในแต่ละรอบการฝึกสอน เป็นตัวกำหนดว่าแบบจำลองจะปรับค่าพารามิเตอร์เท่าใดในแต่ละรอบ การเลือกอัตราการเรียนรู้ (Learning rate) ที่เหมาะสม มีความสำคัญมาก เนื่องจากอัตราการเรียนรู้ (Learning rate) ที่สูงเกินไปอาจทำให้การฝึกสอนไม่สามารถมุ่งเข้าหาคำตอบที่เราต้องการ (Converge) ได้อย่างเหมาะสม ในขณะที่อัตราการเรียนรู้ (Learning rate) ที่ต่ำเกินไป อาจทำให้การฝึกสอนช้าลงและไปไม่ถึงคำตอบที่เราต้องการ การปรับอัตราการเรียนรู้ (Learning rate) ในระหว่างการฝึกสอนแบบจำลองจึงเป็นอีกเรื่องที่สำคัญในการพัฒนาโมเดลให้มีประสิทธิภาพสูงสุด

5.3 ข้อเสนอแนะ

5.3.1 การปรับปรุงคุณภาพของแบบจำลอง โดยแบบจำลองที่พัฒนาในครั้งนี้อาจได้ครอบคลุมกลุ่มผู้ป่วยโรคมะเร็งตับที่มีรอยโรค (Lesion) ขนาดเล็กมาก ๆ ซึ่งผู้ป่วยในกลุ่มนี้จัดเป็นกลุ่มที่วินิจฉัยโรคได้ค่อนข้างยาก ต้องอาศัยรังสีแพทย์ผู้เชี่ยวชาญในการวินิจฉัยแยกโรคได้

5.3.2 ปริมาณข้อมูลที่ใช้ในการวิจัย เนื่องจากงานวิจัยนี้เลือกใช้ภาพถ่ายการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) ในระยะ T1 เท่านั้น ซึ่งประกอบไปด้วย 6 ระยะย่อย (six time points: TPs) คือ ระยะก่อนฉีดสารทึบแสง (Pre-contrast phase: TP1) และระยะหลังฉีดสารทึบแสง 5 ระยะย่อย (Post-contrast: TP2 - TP6) และนอกจากนี้แหล่งข้อมูลที่ได้มามีจำนวนผู้ป่วยค่อนข้างน้อย หากสามารถเพิ่มจำนวนผู้ป่วยที่มีผลการตรวจคลื่นแม่เหล็กไฟฟ้า (MRI) ได้ แบบจำลองน่าจะมีประสิทธิภาพที่ดีมากขึ้น

บรรณานุกรม

- (CPTAC), N. C. I. C. P. T. A. C. (2018). *The Clinical Proteomic Tumor Analysis Consortium Pancreatic Ductal Adenocarcinoma Collection (CPTAC-PDA) (Version 14) [Data set]*. Retrieved from: <https://doi.org/10.7937/K9/TCIA.2018.SC20FO18>
- Castaldo, A., De Lucia, D. R., Pontillo, G., Gatti, M., Cocozza, S., Ugga, L., และ Cuocolo, R. (2021). State of the Art in Artificial Intelligence and Radiomics in Hepatocellular Carcinoma. *Diagnostics*, 11(7), 1194.
- Davila, A., Colan, J., และ Hasegawa, Y. (2024). Comparison of fine-tuning strategies for transfer learning in medical image classification. *Image and Vision Computing*, 146, 105012.
- Erickson, B. J., Kirk, S., Lee, Y., Bathe, O., Kearns, M., Gerdes, C., Rieger-Christ, K., & Lemmerman, J. (2016). *The Cancer Genome Atlas Liver Hepatocellular Carcinoma Collection (TCGA-LIHC) (Version 5) [Data set]*. Retrieved from: <https://doi.org/10.7937/K9/TCIA.2016.IMMQW8UQ>
- Linehan, M., Gautam, R., Kirk, S., Lee, Y., Roche, C., Bonaccio, E., Filippini, J., Rieger-Christ, K., Lemmerman, J., & Jarosz, R. . (2016). *The Cancer Genome Atlas Cervical Kidney Renal Papillary Cell Carcinoma Collection (TCGA-KIRP) (Version 4) [Data set]*. Retrieved from: <https://doi.org/10.7937/K9/TCIA.2016.ACWOGBEF>
- Maithra, R., และ Chiyuan, Z. (2019). Understanding Transfer Learning for Medical Imaging.
- Manders, E., Hoebe, R., Strackee, J., Vossepoel, A., และ Aten, J. A. (1996). Largest contour segmentation: A tool for the localization of spots in confocal images. *Cytometry*, 23, 15-21.
- Taouli, B., Losada, M., Holland, A., และ Krinsky, G. (2004). Magnetic resonance imaging of hepatocellular carcinoma. *Gastroenterology*, 127(5), S144-S152.
- WHO. (2021). Cancer. <https://www.who.int/news-room/fact-sheets/detail/cancer>
- Wu, Y., White, G. M., Cornelius, T., Gowdar, I., Ansari, M. H., Supanich, M. P., และ Deng, J. (2020). Deep learning LI-RADS grading system based on contrast enhanced

- multiphase MRI for differentiation between LR-3 and LR-4/LR-5 liver tumors. *Annals of Translational Medicine*, 8(11).
- Yamashita, R., Mittendorf, A., Zhu, Z., Fowler, K. J., Santillan, C. S., Sirlin, C. B., . . . Do, R. K. (2020). Deep convolutional neural network applied to the liver imaging reporting and data system (LI-RADS) version 2014 category classification: a pilot study. *Abdominal Radiology*, 45(1), 24-35.
- Yan, Y., Chen, M., Shyu, M.-L., และ Chen, S.-C. (2015). *Deep learning for imbalanced multimedia data classification*. Paper presented at the 2015 IEEE international symposium on multimedia (ISM).
- Yasaka, K., Akai, H., Abe, O., และ Kiryu, S. (2018). Deep Learning with Convolutional Neural Network for Differentiation of Liver Masses at Dynamic Contrast-enhanced CT: A Preliminary Study. *Radiology*, 286(3), 887-896.
- พิเชษ, ว., พัฒนพงษ์, ว., พนิดา, ท., พัฒนพงษ์, ช., และ นัฐธริยา, เ. (2013). การวิเคราะห์ประสิทธิภาพการเพิ่มคุณภาพภาพโครโมโซมชนิดจีแบนด์ แบบอัตโนมัติด้วยวิธีฮิสโตแกรมอีควอไลเซชัน.
- มหาวิทยาลัยมหิดล. (2018). การตรวจคลื่นแม่เหล็กไฟฟ้า.
<https://www.rama.mahidol.ac.th/aimc/th/content/09122017-1414-th>
- สุริยะ, ช. (2020). การจำแนกผลึกน้ำตาลด้วยวิธีการเรียนรู้เชิงลึก. 13.

ประวัติผู้เขียน

ชื่อ-สกุล

กานต์พิชชา พรมมา

วุฒิการศึกษา

พ.ศ. 2560

วิทยาศาสตร์บัณฑิต สาขาเวชระเบียน

จากมหาวิทยาลัยมหิดล

