



การวินิจฉัยมะเร็งปอดจากภาพถ่ายรังสีคอมพิวเตอร์ผ่านการเรียนรู้เชิงลึก
Lung Cancer Diagnosis from Computed Tomography Images via Deep
Learning

ปริญญานิพนธ์

ของ

นางสาวปิยรัตน์ สิ้นทรัพย์

โครงการปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม

ปีการศึกษา 1/2569



การวินิจฉัยมะเร็งปอดจากภาพถ่ายรังสีคอมพิวเตอร์ผ่านการเรียนรู้เชิงลึก

Lung Cancer Diagnosis from Computed Tomography Images via Deep Learning

ปริญญาานิพนธ์

ของ

นางสาวปิยรัตน์ สิ้นทรัพย์

โครงการปริญญาานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม

ปีการศึกษา 1/2569

ชื่อโครงการปริญญานิพนธ์	: การวินิจฉัยมะเร็งปอดจากภาพซีทีสแกนผ่านการเรียนรู้เชิงลึก
ผู้จัดทำปริญญานิพนธ์	: นางสาวปิยรัตน์ สิ้นทรัพย์
อาจารย์ที่ปรึกษาปริญญานิพนธ์	: อาจารย์ ดร.อาทิตย์พร โรจรัตน์

บทคัดย่อ

วัตถุประสงค์ของปริญญานิพนธ์ฉบับนี้มุ่งเน้นการศึกษาและเปรียบเทียบประสิทธิภาพของโมเดลการเรียนรู้เชิงลึก (Deep Learning) ในการวินิจฉัยและจำแนกประเภทมะเร็งปอดจากภาพถ่ายรังสีคอมพิวเตอร์ (CT Scan) โดยมุ่งเน้นการแก้ปัญหาความไม่สมดุลของข้อมูลและการขาดแคลนรายละเอียดพยาธิสภาพผ่านเทคโนโลยีปัญญาประดิษฐ์เชิงสร้างสรรค์ (Generative AI) ข้อมูลที่ใช้ในการทดลองมาจากชุดข้อมูลมาตรฐาน IQ-OTH/NCCD ประกอบด้วยคลาสปกติ เนื้อดี และเนื้อร้าย การดำเนินงานวิจัยได้บูรณาการโครงข่ายประสาทเทียมแบบปฏิปักษ์ (GAN) 2 รูปแบบ ได้แก่ SRGAN เพื่อเพิ่มความคมชัดของรายละเอียดเนื้อเยื่อ และ DCGAN เพื่อสังเคราะห์ข้อมูลสำหรับปรับสมดุลกลุ่มตัวอย่าง ร่วมกับการเพิ่มขยายข้อมูลแบบออนไลน์ด้วยเทคนิค Random Pixel Swap (RPS) โมเดลที่ถูกนำมาประเมินประกอบด้วย DenseNet121, MobileNetV3 และ ResNet50V2 ผลการทดลองแสดงให้เห็นว่า ResNet50V2 มีประสิทธิภาพสูงสุด โดยเมื่อทำงานร่วมกับเทคนิค SRGAN และการเพิ่มขยายข้อมูลแบบออนไลน์ สามารถบรรลุค่า Accuracy สูงสุดที่ 98.20% และค่า AUC ที่ 0.9994 ในขณะที่โมเดลขนาดเล็กอย่าง MobileNetV3 แสดงการพัฒนาอย่างก้าวกระโดดเมื่อใช้ข้อมูลสังเคราะห์ นอกจากนี้ งานวิจัยยังได้ประยุกต์ใช้เทคนิค Grad-CAM เพื่อสร้างแผนภูมิความร้อนระบุตำแหน่งรอยโรค ช่วยเพิ่มความโปร่งใสและสร้างความเชื่อมั่นในเชิงรังสีวินิจฉัย บทสรุปจากงานวิจัยนี้ชี้ให้เห็นว่าการบูรณาการ GAN ร่วมกับสถาปัตยกรรมโครงข่ายประสาทเทียมเชิงลึก เป็นกลยุทธ์ที่มีประสิทธิภาพสูงในการยกระดับความแม่นยำและความทนทานของระบบแพทย์อัจฉริยะ เพื่อสนับสนุนการตัดสินใจรักษาที่รวดเร็วและแม่นยำยิ่งขึ้น

Thesis Title : Lung Cancer Diagnosis from Computed Tomography Images
via Deep Learning

Author : Piyarat Sinsap

Supervisor : Artitayaporn Rojarath

Abstract

The objective of this thesis is to investigate and compare the performance of deep learning models for diagnosing and classifying lung cancer from Computed Tomography (CT) scans. The study focuses on addressing class imbalance and the scarcity of pathological details through Generative Artificial Intelligence (Generative AI). Utilizing the IQ-OTH/NCCD standard dataset, which includes normal, benign, and malignant classes, the research integrates two Generative Adversarial Network (GAN) frameworks: SRGAN for enhancing tissue texture clarity and DCGAN for synthetic data balancing. These are combined with online data augmentation using the Random Pixel Swap (RPS) technique. The evaluated architectures consist of DenseNet121, MobileNetV3, and ResNet50V2. Experimental results demonstrate that ResNet50V2 achieved the highest efficiency; when integrated with SRGAN and online augmentation, it reached a peak accuracy of 98.20% and an AUC of 0.9994. Furthermore, the Grad-CAM technique was employed to generate heatmaps for clinical interpretability, identifying regions of interest (ROI) used in model decision-making. This research concludes that the synergy between GAN-based enhancement and deep learning architectures provides a robust and highly accurate strategy for medical AI systems. These findings serve as a foundation for supporting radiologists in making faster and more precise clinical decisions, ultimately improving patient survival rates.

กิตติกรรมประกาศ

ปริญญานิพนธ์ฉบับนี้สำเร็จลุล่วงไปด้วยดี เนื่องจากได้รับความอนุเคราะห์ ความช่วยเหลือ และคำแนะนำในการแก้ไขข้อบกพร่องต่าง ๆ จากอาจารย์ที่ปรึกษา คณาจารย์ ตลอดจนผู้มีพระคุณทุกท่าน ผู้จัดทำรู้สึกซาบซึ้งในความอนุเคราะห์อันยิ่งใหญ่นี้ จึงใคร่ขอขอบพระคุณทุกท่านไว้ ณ โอกาสนี้ ขอขอบพระคุณ อาจารย์ ดร.อาทิตยาพร โรจรัตน์ อาจารย์ที่ปรึกษาโครงงานปริญญานิพนธ์ ที่ได้กรุณาให้ความรู้ คำปรึกษา ตรวจทาน ตลอดจนให้แนวทางในการแก้ไขปัญหาดังต่าง ๆ ด้วยความเอาใจใส่เป็นอย่างดี จนทำให้ปริญญานิพนธ์ฉบับนี้สำเร็จได้อย่างสมบูรณ์

ขอขอบพระคุณคณาจารย์ในภาควิชาเทคโนโลยีสารสนเทศ คณะวิทยาการสารสนเทศทุกท่าน ที่ได้ประสิทธิ์ประสาทวิชาความรู้ ทั้งในด้านวิชาการ คุณธรรม และจริยธรรม ซึ่งเป็นประโยชน์อย่างยิ่งในการนำมาประยุกต์ใช้ในการทำโครงงานและการดำเนินชีวิต ขอกราบขอบพระคุณบิดา มารดา และครอบครัว ที่ให้การอบรมเลี้ยงดู คอยสนับสนุน ให้ความช่วยเหลือ และเป็นกำลังใจที่สำคัญยิ่งมาโดยตลอด ขอขอบคุณเพื่อน ๆ นิสิตสาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาการสารสนเทศ ทุกคน ที่ให้คำปรึกษา แลกเปลี่ยนความรู้ และคอยช่วยเหลือซัพพอร์ตกันตลอดระยะเวลาการศึกษาคุณค่าและประโยชน์อันพึงมีจากปริญญานิพนธ์ฉบับนี้ ผู้จัดทำขออัญมฤตเวทิตาและขอมอบให้แก่ผู้มีพระคุณทุกท่าน ที่มีส่วนช่วยให้ปริญญานิพนธ์ฉบับนี้สำเร็จลุล่วงไปด้วยดี

ปิยรัตน์ สิ้นทรัพย์

สารบัญ

หน้า

บทคัดย่อ	ก
Abstract	ข
กิตติกรรมประกาศ	ค
สารบัญ	ง
สารบัญตาราง	ช
สารบัญภาพประกอบ	ซ
บทที่ 1 บทนำ	1
1.1 หลักการและเหตุผล	1
1.2 วัตถุประสงค์	2
1.3 ขอบเขตโครงการ	2
1.3.1 ขอบเขตด้านชุดข้อมูล	2
1.3.2 ขอบเขตด้านเทคนิคและวิธีการ	2
1.4 ประโยชน์ที่คาดว่าจะได้รับ	4
1.5 อุปกรณ์และเครื่องมือที่ใช้	5
1.6 แผนการดำเนินงานตลอดโครงการ	7
บทที่ 2 ทฤษฎีและผลงานที่เกี่ยวข้อง	8
2.1 ทฤษฎีที่เกี่ยวข้อง	8
2.1.1 ความรู้พื้นฐานเกี่ยวกับโรคมะเร็งปอดและภาพถ่าย CT Scan	8
2.1.2 การประมวลผลภาพดิจิทัล (Digital Image Processing)	8
2.1.3 ปัญญาประดิษฐ์และการเรียนรู้เชิงลึก (AI and Deep Learning)	9
2.1.4 สถาปัตยกรรมแบบจำลองการเรียนรู้เชิงลึกที่ใช้เปรียบเทียบ	9
2.1.5 โครงข่ายประสาทเทียมแบบสร้างสรรค์ (GAN) Generative Adversarial Networks	10
2.1.6 เทคนิคการถ่ายโอนการเรียนรู้ (Transfer Learning)	10
2.1.7 การวัดประสิทธิภาพของแบบจำลอง (Performance Evaluation Metrics)	10
2.1.8 ขั้นตอนการดำเนินงานตามมาตรฐาน CRISP-DM (Research Methodology)	12
2.2 วิจัยที่เกี่ยวข้อง	13

2.2.1 งานวิจัยด้านระบบการวินิจฉัยแบบไฮบริดและโครงสร้างแบบครบวงจร	13
2.2.2 งานวิจัยด้านการเพิ่มคุณภาพภาพและความละเอียดสูง.....	15
2.2.3 งานวิจัยด้านเทคนิคการเพิ่มขยายข้อมูลและการจัดการข้อมูลไม่สมดุล	19
2.2.4 งานวิจัยด้านการเปรียบเทียบสถาปัตยกรรมและประสิทธิภาพของโมเดล.....	22
2.2.5 งานวิจัยด้านการสกัดคุณลักษณะและการแสดงผลเพื่อการอธิบายผล	25
บทที่ 3 ขั้นตอนการดำเนินโครงการ.....	29
3.1 ภาพรวมกรอบการวิจัย.....	29
3.2 ชุดข้อมูล	30
3.3 การเตรียมและแบ่งชุดข้อมูล	31
3.3.1 โปรโตคอลการแบ่งข้อมูล (Split Protocol).....	31
3.3.2 การล็อกชุดข้อมูลทดสอบ (Locking the Test Set).....	32
3.4 การประมวลผลภาพเบื้องต้น	33
3.4.1 การปรับขนาดภาพ (Image Resizing).....	33
3.4.2 การปรับความต่างสีแบบปรับเหมาะจำกัดเกณฑ์ (CLAHE).....	33
3.4.3 การปรับบรรทัดฐานค่าพารามิเตอร์พิกเซล (Normalization) และการแปลงช่องสัญญาณ.....	34
3.5 การเพิ่มข้อมูลด้วย GAN (GAN-based Data Augmentation).....	35
3.5.1 กลยุทธ์กฎทวิคูณ (Double Rule Strategy).....	35
3.5.2 SRGAN (Super-Resolution GAN)	36
3.5.3 DCGAN (Deep Convolutional GAN).....	38
3.6 การออกแบบกระบวนการทดลอง (Experimental Design).....	39
3.6.1 ภาพรวม 12 กระบวนการทดลอง (12 Experimental Processes).....	40
3.6.2 จำนวนภาพที่ใช้จริงต่อ Process (Actual Dataset Distribution per Process).....	40
3.6.3 การปรับสมดุลชั้นเรียนข้อมูล (Class Balancing Strategy).....	42
3.6.4 เทคนิคการเพิ่มขยายข้อมูลแบบออนไลน์ (Online Augmentation Techniques)	43
3.7 สถาปัตยกรรมตัวจำแนกประเภท (CNN Classification Architectures).....	44
3.7.1 การเลือกโมเดล (Model Selection)	44
3.7.2 DenseNet121	45
3.7.3 MobileNetV3-Large.....	45

3.7.4 ResNet50V2	45
3.7.5 กลยุทธ์การเรียนรู้แบบย้ายฝากและการฝึกสอน	46
3.7.6 โครงสร้างส่วนหัวการจำแนกประเภทและกระบวนการฝึกสอน	48
3.8 สรุปจำนวนการทดลองทั้งหมด (Summary of Experimental Matrix).....	49
บทที่ 4 ผลการวิจัย.....	50
4.1 ผลการทดลอง.....	50
4.1.1 ผลการทดลองบน Split V1 — Abe (70:30).....	51
4.1.2 ผลการทดลองบน Split V2 — Jozi (90:10)	53
4.3 บทวิเคราะห์และอภิปรายผลเชิงลึก (In-Depth Discussion & Medical Insights)	54
บทที่ 5 สรุป อภิปรายผล และข้อเสนอแนะ.....	56
5.1 สรุปผลการวิจัย	56
5.2 อภิปรายผลการวิจัย (Discussion).....	57
5.3 ข้อเสนอแนะสำหรับการวิจัยในอนาคต.....	59
เอกสารอ้างอิง	61

สารบัญตาราง

หน้า

ตารางที่ 1.1	รายละเอียดจำนวนข้อมูลทั้ง 3 class.....	2
ตารางที่ 1.2	ระยะเวลาการดำเนินโครงการ	7
ตารางที่ 2.1	ตารางเปรียบเทียบงานวิจัยที่ใช้ชุดข้อมูล IQ-OTH/NCCD	28
ตารางที่ 3.1	กระบวนการเตรียมและประมวลผลภาพเบื้องต้น (Preprocessing Pipeline).....	34
ตารางที่ 3.2	จำนวนภาพถ่ายในชุดข้อมูลฝึกฝนก่อนและหลังการใช้นวัตกรรม Double Rule (2x).....	35
ตารางที่ 3.3	ค่าพารามิเตอร์ของแบบจำลอง SRGAN ที่ใช้ในการทดลองจริง.....	37
ตารางที่ 3.4	ค่าพารามิเตอร์ของแบบจำลอง DCGAN ที่ใช้ในการทดลอง	39
ตารางที่ 3.5	รายละเอียดและวัตถุประสงค์ของ 12 กระบวนการทดลอง (Experimental Processes)	40
ตารางที่ 3.6	จำนวนภาพถ่าย CT-Scan ในชุดฝึกสอน	41
ตารางที่ 3.7	จำนวนภาพถ่ายในชุดฝึกสอนก่อนและหลังการปรับสมดุลข้อมูล (Class Balancing).....	42
ตารางที่ 3.8	เปรียบเทียบคุณลักษณะทางเทคนิคและสถาปัตยกรรมของตัวจำแนกประเภททั้ง 3 โมเดล ...	46
ตารางที่ 3.9	เมทริกซ์สรุปจำนวนสภาวะการทดลองทั้งหมด (Experimental Matrix).....	49
ตารางที่ 4.1	Split V1 (Abe 70:30) กลุ่ม SRGAN เปรียบเทียบ 3 สถาปัตยกรรม CNN.....	51
ตารางที่ 4.2	Split V1 (Abe 70:30) กลุ่ม DCGAN เปรียบเทียบ 3 สถาปัตยกรรม CNN.....	52
ตารางที่ 4.3	Split V1 (Abe 70:30) กลุ่ม Augmentation-only.....	52
ตารางที่ 4.4	Split V2 (Jozi 90:10) กลุ่ม SRGAN เปรียบเทียบ 3 สถาปัตยกรรม CNN.....	53
ตารางที่ 4.5	Split V2 (Jozi 90:10) กลุ่ม Augmentation-only	54

สารบัญภาพประกอบ

หน้า

ภาพประกอบที่ 2.1	ตัวอย่างภาพ CT SCAN	8
ภาพประกอบที่ 2.2	ตัวอย่างกระบวนการทำ GAN.....	10
ภาพประกอบที่ 2.3	ตาราง Confusion Matrix	10
ภาพประกอบที่ 2.4	วงจรการทำงาน CRISP-DM (Research Methodology).....	12
ภาพประกอบที่ 3.1	Framework of preprocessing technique.....	30
ภาพประกอบที่ 3.2	ภาพct-scans ตัวอย่างแต่ละคลาส	30
ภาพประกอบที่ 3.3	กราฟแสดงจำนวนข้อมูลทั้ง 3 คลาส	31
ภาพประกอบที่ 3.4	จำนวนภาพในชุดข้อมูลฝึกสอนและทดสอบภายใต้ 2 โปรโตคอลมาตรฐาน	32

บทที่ 1

บทนำ

1.1 หลักการและเหตุผล

โรคมะเร็งปอดเป็นสาเหตุอันดับต้นของการเสียชีวิตจากโรคมะเร็งทั่วโลก โดยมีสถิติผู้เสียชีวิตสูงถึง 1.8 ล้านคนต่อปี หรือคิดเป็นร้อยละ 18 ของยอดผู้เสียชีวิตจากมะเร็งทั้งหมด ปัจจัยวิกฤตที่ส่งผลต่ออัตราการรอดชีวิตคือมิติด้านเวลาและความแม่นยำในการตรวจพบโรคในระยะเริ่มต้น[1] ซึ่งการคัดกรองด้วยภาพถ่ายรังสีคอมพิวเตอร์ปริมาณรังสีต่ำ (LDCT)[2] ได้รับการพิสูจน์ว่าช่วยลดอัตราการตายลงได้ถึงร้อยละ 20 อย่างไรก็ตาม ด้วยปริมาณภาพถ่ายจำนวนมากมหาศาลทำให้รังสีแพทย์ต้องเผชิญกับภาระงานที่หนักอึ้ง นำไปสู่ความผิดพลาดจากการเหนื่อยล้า (Fatigue-related errors)[3] ซึ่งอาจก่อให้เกิดความผิดพลาดจากการรายงานผลลบลงได้แม้จะมีการนำโครงข่ายประสาทเทียมแบบคอนโวลูชัน (CNNs) มาเป็นระบบช่วยวินิจฉัย แต่การพัฒนาที่ยังคงติดอุปสรรคสำคัญด้าน "ความไม่สมดุลของข้อมูลเชิงกลุ่ม" (Class Imbalance) โดยเฉพาะในชุดข้อมูลมาตรฐาน IQ-OTH/NCCD ที่กลุ่มรอยโรคชนิดไม่ร้ายแรง (Benign) มีปริมาณน้อยมากเพียงร้อยละ 11 ของข้อมูลทั้งหมด ส่งผลให้โมเดลเกิดอคติในการเรียนรู้และขาดความไวในการตรวจจברอยโรคที่เกิดขึ้นจริง[4]

การศึกษานี้จึงมีแนวคิดที่จะเพื่อพัฒนาตัวแบบการเรียนรู้เชิงลึกในการจำแนกประเภทมะเร็งปอดแบบ 3 กลุ่มดังนี้ โดยบูรณาการเทคโนโลยี GAN เพื่อยกระดับคุณภาพและปรับสมดุลข้อมูล พร้อมทั้งเปรียบเทียบสมรรถนะของสถาปัตยกรรม CNN 3 รูปแบบ และประยุกต์ใช้เทคนิค Grad-CAM เพื่อสร้างภาพโปร่งใสในการวินิจฉัยเพื่อก้าวข้ามขีดจำกัดของการเพิ่มข้อมูลแบบดั้งเดิมที่มักสูญเสียรายละเอียดพยาธิสภาพ ผู้วิจัยจึงได้นำเสนอการใช้ ปัญญาประดิษฐ์เชิงสร้างสรรค์ (Generative AI) ผ่านสถาปัตยกรรมโครงข่ายประสาทเทียมแบบปฏิปักษ์ 2 รูปแบบหลัก ได้แก่ SRGAN เพื่อกู้คืนรายละเอียดพื้นผิวเนื้อเยื่อที่ซับซ้อน (Texture Fidelity) และรายละเอียดรอยโรคขนาดเล็ก (Micro-nodules) ที่มักสูญหายจากการบีบอัดไฟล์ภาพและ DCGAN สำหรับการสังเคราะห์ภาพรอยโรคชุดใหม่เพื่อชดเชยรอยร้าวทางสถิติในกลุ่มข้อมูลส่วนน้อยให้เกิดความสมดุลควบคู่ไปกับการใช้นวัตกรรมการสลับพิกเซลแบบสุ่ม (Random Pixel Swap: RPS)[5] ซึ่งเป็นกลยุทธ์การเพิ่มข้อมูลแบบออนไลน์ที่ช่วยรักษาโครงสร้างทางกายวิภาคและความสมบูรณ์ทางพยาธิสภาพไว้ได้อย่างแม่นยำข้อมูลที่ผ่านการยกระดับคุณภาพแล้วจะถูกนำไปฝึกสอนผ่านเทคนิคการเรียนรู้แบบถ่ายโอน (Transfer Learning) กับสถาปัตยกรรม CNN ขั้นนำ 3 รูปแบบที่มีจุดเด่นต่างกัน ได้แก่ DenseNet121 ที่โดดเด่นด้านการใช้คุณลักษณะซ้ำ, MobileNetV3 สำหรับการประมวลผลที่รวดเร็ว และ ResNet50V2 ที่มีความเสถียรสูงในการเรียนรู้โครงสร้างรอยโรคที่ซ้อนทับลึก

งานวิจัยนี้มุ่งเน้นการศึกษาและเปรียบเทียบสมรรถนะของสถาปัตยกรรมโครงข่ายประสาทเทียมรูปแบบต่าง ๆ เพื่อค้นหาแบบจำลองที่มีประสิทธิภาพสูงสุดในการจำแนกประเภทมะเร็งปอด พร้อมทั้งประยุกต์ใช้เทคนิค Grad-CAM (Explainable AI) เพื่ออธิบายเหตุผลในการทำนายผลของโมเดลผ่านแผนภูมิความร้อน ซึ่งจะช่วยลดความคลุมเครือและเพิ่มความโปร่งใสในกระบวนการตรวจสอบตำแหน่งรอยโรค อย่างไรก็ตาม แบบจำลองที่ศึกษาขึ้นมีวัตถุประสงค์เพื่อเป็นเครื่องมือสนับสนุนการทำนายเบื้องต้นเท่านั้น ไม่สามารถทดแทนการวินิจฉัยยืนยันผลโดยแพทย์ผู้เชี่ยวชาญได้ เนื่องด้วยขีดจำกัดของข้อมูลและความซับซ้อนของโรค โดยการตัดสินใจขั้นสุดท้ายยังคงขึ้นอยู่กับดุลยพินิจของแพทย์เป็นสำคัญ

1.2 วัตถุประสงค์

1.2.1 เพื่อพัฒนาและสร้างโมเดลการเรียนรู้เชิงลึกที่สามารถวิเคราะห์และทำนายผลโรคมะเร็งปอดจากภาพถ่ายเอกซเรย์คอมพิวเตอร์

1.2.2 เพื่อประยุกต์ใช้เทคนิคสังเคราะห์ข้อมูลด้วย DCGAN และเพิ่มความคมชัดของภาพด้วย SRGAN ร่วมกับเทคนิค Data Augmentation ในการแก้ปัญหาความไม่สมดุลของข้อมูลและเพิ่มความหลากหลายของชุดข้อมูล

1.2.3 เพื่อพัฒนาสถาปัตยกรรมหลอมรวมคุณลักษณะภาพ (CNN Feature Fusion) ในการรวมจุดเด่นจากโมเดลฐานเพื่อยกระดับความแม่นยำในการทำนายผล

1.2.4 เพื่อประยุกต์ใช้เทคนิคปัญญาประดิษฐ์แบบอธิบายได้ (Explainable AI: Grad-CAM) ในการระบุตำแหน่งรอยโรคและสร้างภาพอธิบายบริเวณความผิดปกติบนภาพถ่ายเอกซเรย์คอมพิวเตอร์

1.3 ขอบเขตโครงการ

1.3.1 ขอบเขตด้านชุดข้อมูล

ใช้ชุดข้อมูลภาพถ่ายเอกซเรย์คอมพิวเตอร์ (CT Scan) จากแหล่งข้อมูลมาตรฐานเปิดเผย (Kaggle) ชื่อชุดข้อมูล The IQ-OTH/NCCD Lung Cancer Dataset รวบรวมจากโรงพยาบาลเฉพาะทาง 2 แห่งในประเทศอิรัก ได้แก่ Iraq-Oncology Teaching Hospital และ National Center for Cancer Diseases (NCCD) โดยเก็บข้อมูลในช่วงฤดูใบไม้ร่วงปี 2019 และได้รับการรับรองจากคณะกรรมการพิจารณาจริยธรรมของศูนย์การแพทย์ที่เกี่ยวข้อง ข้อมูลแบ่งออกเป็น 3 กลุ่มอาการ ได้แก่ กลุ่มปกติ (Normal), กลุ่มเนื้องอกไม่อันตราย (Benign) และกลุ่มเนื้องอกร้าย/มะเร็ง (Malignant)

ตารางที่ 1.1 รายละเอียดจำนวนข้อมูลทั้ง 3 class

คลาส (Class)	จำนวนเคส (Cases)	จำนวนภาพ (Slices)	ลักษณะของก้อนเนื้อ
Normal (ปกติ)	55	416	ไม่พบก้อนเนื้อผิดปกติ
Benign (เนื้อดี)	15	120	ก้อนเนื้อขนาดเล็ก ขอบชัด ไม่แพร่กระจาย
Malignant (เนื้อร้าย)	40	561	ก้อนเนื้อขอบไม่ชัด มีลักษณะรุกราน
รวม	110	1097	

1.3.2 ขอบเขตด้านเทคนิคและวิธีการ

1.3.2.1 เทคนิคการเตรียมข้อมูลและการแบ่งโปรโตคอล (Data Preparation & Protocol Splitting)

1) ชุดข้อมูลวิจัย: ใช้ชุดข้อมูลภาพถ่ายรังสีคอมพิวเตอร์ปอด IQ-OTH/NCCD Lung Cancer Dataset จำนวนรวม 1,097 ภาพ ประกอบด้วย 3 คลาสการวินิจฉัย ได้แก่ ภาพปอดปกติ (Normal) จำนวน 416 ภาพ, ภาพก้อนเนื้อชนิดไม่ร้ายแรง (Benign) จำนวน 120 ภาพ และภาพก้อนเนื้อชนิดร้ายแรง/มะเร็ง (Malignant) จำนวน 561 ภาพ

2) การแบ่งชุดข้อมูล (Dataset Splitting Protocol): ดำเนินการแบ่งข้อมูลแบบคงสัดส่วนคลาส (Stratified Split) ออกเป็น 2 โปรโตคอลหลักตามมาตรฐานงานวิจัยอ้างอิง และทำการล็อกชุดข้อมูลทดสอบ (Locked Test Set) ไว้ตั้งแต่นิยามเริ่มต้น ไม่ผ่านการประมวลผลเพิ่มขยายใดๆ เพื่อใช้ประเมินผลประสิทธิภาพที่แท้จริง A) Split V1 (Abe Protocol 70:30): กำหนดชุดข้อมูลทดสอบ (Locked Test Set) สัดส่วนร้อยละ 30 (จำนวน 330 ภาพ) B) Split V2 (Jozi Protocol 90:10): กำหนดชุดข้อมูลทดสอบ (Locked Test Set) สัดส่วนร้อยละ 10 (จำนวน 111 ภาพ)

1.3.2.2 เทคนิคการประมวลผลและปรับแต่งภาพก่อนนำเข้า (Image Preprocessing)

1) Contrast Enhancement: ประยุกต์ใช้เทคนิค Contrast Limited Adaptive Histogram Equalization (CLAHE) โดยกำหนดค่า clipLimit = 2.0 และขนาด tileGridSize = 8x8 เพื่อเพิ่มความเปรียบต่างเฉพาะที่ (Local Contrast) ของเนื้อเยื่อปอดในภาพ CT-Scan

2) Image Resizing & Normalization: ปรับขนาดภาพเป็น 128x128 พิกเซล สำหรับการฝึกสอนแบบจำลองสังเคราะห์ภาพ (GAN) และปรับขนาดเป็น 224x224 พิกเซล สำหรับแบบจำลองจำแนกประเภท (Classifier) พร้อมทำการปรับระดับค่าพิกเซลให้อยู่ในช่วง [-1, 1] สำหรับ GAN และช่วง [0, 1] สำหรับ Classifier

1.3.2.3 เทคนิคการเพิ่มปริมาณข้อมูลด้วย Generative AI และ Pixel Augmentation (GAN & Data Augmentation)

ประยุกต์ใช้เทคนิค Generative AI แบบคู่ขนาน (Double Rule: ภาพต้นฉบับ 1 ภาพ ต่อภาพสังเคราะห์ 1 ภาพ) ร่วมกับเทคนิคการแปลงพิกเซลแบบออนไลน์ (Online Pixel-level Augmentation) ครอบคลุม 12 กระบวนการทดลอง (Process P1-P12) ได้แก่:

1) Super-Resolution GAN (SRGAN): สถาปัตยกรรม Generator แบบ 16 Residual Blocks ร่วมกับ 2x PixelShuffle Upsampling และ Perceptual Loss จาก VGG19 เพื่อยกระดับความคมชัดของขอบเขตรอยโรค (Nodule Boundaries) โดยมีคุณภาพสัญญาณผ่านเกณฑ์มาตรฐานทางการแพทย์ ($PSNR > 20$ dB, $SSIM > 0.80$)

2) Deep Convolutional GAN (DCGAN): ฝึกสอนแยกตามคลาสการวินิจฉัย (Class-specific Training) โดย Generator รับสัญญาณรบกวน (Noise Vector) ขนาด 128 มิติ ร่วมกับการใช้ Label Smoothing เพื่อสร้างภาพรังสีสังเคราะห์จำลองพยาธิวิทยา

3) Pixel-level Augmentation Strategies: ประยุกต์ใช้เทคนิคการดัดแปลงระดับพิกเซล ได้แก่ Random Pixel Swap, Random Erasing และ Cutout เพื่อสร้างสภาวะรบกวน และป้องกันการเกิด Overfitting ของตัวแบบ

1.3.2.4 สถาปัตยกรรมแบบจำลองการจำแนกประเภทและการฝึกสอน (Model Architectures & Training Strategy)

1) สถาปัตยกรรมที่ใช้เปรียบเทียบ: ประยุกต์ใช้การเรียนรู้เชิงลึกแบบ Transfer Learning จาก ImageNet บน 3 สถาปัตยกรรมหลัก ได้แก่ DenseNet121, MobileNetV3-Large และ ResNet50V2 (รวมการทดลองทั้งสิ้น 72 สภาวะ: 12 Processes $\times$ 3 Models $\times$ 2 Splits)

2) กลยุทธ์การฝึกสอน 2 ระยะ (Two-Phase Fine-Tuning Strategy): A) Phase 1 (Frozen Backbone): แช่แข็งค่าน้ำหนักส่วนสกัดคุณลักษณะ (Feature Extractor) และฝึกสอนส่วนจำแนกประเภทจำนวน 10 Epochs B) Phase 2 (Fine-Tuning): ปลดล็อกการเรียนรู้ในระดับ Top 30 Layers ด้วยอัตราการเรียนรู้คงที่ (Learning Rate = 1×10^{-5}) ฝึกสอนต่อเนื่องจนครบ 100 Epochs โดยไม่ใช้ Early Stopping เพื่อให้สามารถเปรียบเทียบเสถียรภาพการรู้เข้าของทุกสภาวะทดสอบได้อย่างเป็นธรรมและเที่ยงตรง

1.3.2.5 เทคนิคการอธิบายผลของแบบจำลอง (Explainable AI: XAI)

Grad-CAM (Gradient-weighted Class Activation Mapping): คำนวณค่าความชันและน้ำหนักบน Feature Maps ของ Convolutional Layer ชั้นสุดท้าย เพื่อสร้างแผนภาพความร้อน (Heatmap) แสดงตำแหน่งพื้นที่บนภาพ CT-Scan ที่ตัวแบบใช้เป็นน้ำหนักสำคัญในการตัดสินใจจำแนกคลาสมะเร็ง ยืนยันความสอดคล้องทางการแพทย์ (Clinical Validity)

1.3.2.6 ตัวชี้วัดและการประเมินผลประสิทธิภาพ (Performance Evaluation Metrics)

1) การประเมินประสิทธิภาพการจำแนกประเภท (Classification Performance): ประเมินผลบน Locked Test Set ด้วยตารางเมทริกซ์ความสับสน (Confusion Matrix) ผ่านค่าความถูกต้องรวม (Accuracy), ค่าเอฟวัน (F1-Score), ค่าพื้นที่ใต้กราฟคุณลักษณะการทำงานของตัวรับ (AUROC), ค่าความไว (Sensitivity / Recall), ค่าความจำเพาะ (Specificity) และค่าความแม่นยำ (Precision) โดยมุ่งเน้น Malignant Class Sensitivity เป็นสำคัญ

2) การประเมินคุณภาพภาพสังเคราะห์ (GAN Quality Metrics): ประเมินคุณภาพสัญญาณและความเหมือนเชิงโครงสร้างเนื้อเยื่อระหว่างภาพจริงกับภาพสังเคราะห์ด้วยค่าอัตราส่วนสัญญาณสูงสุดต่อสัญญาณรบกวน (PSNR) และดัชนีความเหมือนเชิงโครงสร้าง (SSIM)

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1.4.1 ได้แบบจำลองการเรียนรู้เชิงลึกที่สามารถจำแนกและทำนายผลโรคมะเร็งปอดจากภาพซีทีสแกนได้อย่างแม่นยำ

1.4.2 ได้แนวทางการแก้ปัญหาข้อมูลไม่สมดุลและเพิ่มความหลากหลายของภาพฝึกสอนด้วยเทคนิค DCGAN, SRGAN และ Data Augmentation

1.4.3 ได้สถาปัตยกรรมหลอมรวมคุณลักษณะภาพ (CNN Feature Fusion) ที่ช่วยยกระดับประสิทธิภาพและความแม่นยำในการทำนายผล

1.4.4 ได้ระบบอธิบายผลด้วยแผนภาพความร้อน (Grad-CAM) ที่ระบุตำแหน่งรอยโรคได้อย่างชัดเจน ช่วยเพิ่มความน่าเชื่อถือในการตัดสินใจทางรังสีวินิจฉัย

1.5 อุปกรณ์และเครื่องมือที่ใช้

1.5.1 อุปกรณ์ที่ใช้

1.5.1.1 เครื่องคอมพิวเตอร์พกพา (Laptop): สำหรับการเขียนโปรแกรมและจัดการเอกสาร

- 1) Processor: AMD Ryzen 5 7530U with Radeon Graphics (2.00 GHz)
- 2) RAM: 16.0 GB
- 3) System type: 64-bit operating system

1.5.1.2 ระบบประมวลผลบนคลาวด์ (Cloud Computing)

- 1) Google Colab สำหรับการฝึกสอนโมเดล (Model Training)

ที่ต้องใช้ทรัพยากรการคำนวณสูง โครงการได้เลือกใช้ Google Colab ที่ให้บริการ GPU T4 หรือสูงกว่า เนื่องจากข้อจำกัดของเครื่องคอมพิวเตอร์พกพาหลักที่ไม่มีหน่วยประมวลผลกราฟิกแยก (Discrete GPU) ประสิทธิภาพสูงเพียงพอสำหรับรองรับภาระงานด้าน Deep Learning ที่มีความซับซ้อน

1.5.2 เครื่องมือที่ใช้

1.5.2.1 ภาษา Python

เป็นภาษาโปรแกรมระดับสูงที่ถูกเลือกใช้สำหรับการพัฒนาโมเดล Deep Learning เนื่องจากเป็นภาษาที่ได้รับความนิยมสูงสุดในงานด้านวิทยาศาสตร์ข้อมูลและปัญญาประดิษฐ์ ด้วยลักษณะเด่นที่มีโครงสร้างภาษาอ่านและเขียนง่าย จึงเอื้อต่อการทำความเข้าใจและการพัฒนา อีกทั้งยังมีคลังคำสั่งหรือไลบรารีสนับสนุนงานด้าน AI ที่ครบถ้วนและหลากหลาย เช่น PyTorch, TensorFlow หรือ OpenCV พร้อมทั้งมีชุมชนนักพัฒนาขนาดใหญ่ที่ช่วยให้การสืบค้นวิธีแก้ไขปัญหาทำได้โดยง่ายและรองรับการทำงานข้ามแพลตฟอร์ม แม้ว่าในเชิงเทคนิค Python อาจมีการประมวลผลช้ากว่าภาษา C++ และมีการจัดการหน่วยความจำที่ไม่เทียบเท่าภาษาระดับต่ำ แต่ปัญหานี้สามารถจัดการได้ด้วยการเลือกใช้ไลบรารีที่พัฒนาด้วยภาษา C++ เป็นเบื้องหลัง เช่น NumPy เพื่อเพิ่มประสิทธิภาพในการคำนวณ

1.5.2.2 Google Colab (Google Colaboratory)

เป็นสภาพแวดล้อมสำหรับการพัฒนาโปรแกรมภาษา Python บนระบบคลาวด์ที่ให้บริการโดย Google ซึ่งช่วยให้ผู้ใช้งานสามารถเขียนและประมวลผลโค้ดผ่านเว็บเบราว์เซอร์ได้ทันทีโดยไม่ต้องติดตั้งซอฟต์แวร์ลงบนเครื่อง ความสามารถหลักที่สำคัญคือการเปิดให้เข้าถึงทรัพยากร GPU ประสิทธิภาพสูงได้โดยไม่มีค่าใช้จ่าย ซึ่งจำเป็นอย่างยิ่งสำหรับการฝึกสอนโมเดล Deep Learning ที่คอมพิวเตอร์ทั่วไปไม่สามารถทำได้ อีกทั้งยังลดความยุ่งยากในการตั้งค่าสภาพแวดล้อมต่าง ๆ และสามารถทำงานร่วมกับ Google Drive เพื่อบันทึกหรือเรียกใช้ข้อมูลได้อย่างสะดวกรวดเร็ว อย่างไรก็ตาม การใช้งาน Google Colab มีข้อพิจารณาเรื่องข้อจำกัดด้านเวลา หากใช้งานต่อเนื่องนานเกินไปหรือไม่มีการตอบสนองระบบอาจเกิดการเชื่อมต่อและจำเป็นต้องอาศัยสัญญาณอินเทอร์เน็ตที่เสถียรตลอดระยะเวลาการทำงาน

1.5.2.3 PyTorch

เป็นไลบรารีโอเพนซอร์สสำหรับการเรียนรู้เชิงลึกที่พัฒนาโดย Meta AI ซึ่งมีชื่อเสียงในด้านการคำนวณแบบพลวัต (Dynamic Computation Graph) และเป็นที่นิยมอย่างมากในงานวิจัยระดับสูง โดย PyTorch มีความยืดหยุ่นสูงซึ่งเหมาะอย่างยิ่งสำหรับการสร้างและปรับแต่งโมเดลที่มีโครงสร้างซับซ้อน เช่น WGAN-GP หรือโมเดลรูปแบบใหม่ ๆ อีกทั้งรูปแบบการเขียนโค้ดยังมีความคล้ายคลึงกับภาษา Python ปกติ ทำให้ง่ายต่อการตรวจสอบข้อผิดพลาด (Debug) และทำความเข้าใจขั้นตอนการทำงาน นอกจากนี้ยังได้รับการสนับสนุนจากชุมชนนักวิจัยขนาดใหญ่ ทำให้มีตัวอย่างโค้ดและงานวิจัยรองรับจำนวนมาก โดยเฉพาะใน

งานวิจัยทางการแพทย์ แม้ว่าสำหรับผู้เริ่มต้นอาจจะต้องใช้เวลาเรียนรู้โครงสร้างคำสั่งที่ละเอียดกว่าไลบรารีสำเร็จรูปอื่น ๆ ในการจัดการกระบวนการฝึกสอนโมเดลก็ตาม

1.5.2.4 TensorFlow และ Keras

เป็นไลบรารีสำหรับงานปัญญาประดิษฐ์ที่พัฒนาโดย Google โดยใช้งานร่วมกับ Keras ซึ่งเป็น API ระดับสูงที่ช่วยให้การเขียนโค้ดสร้างโครงข่ายประสาทเทียมทำได้กระชับและง่ายขึ้น เครื่องมือนี้ได้รับการออกแบบมาให้ใช้งานง่าย เหมาะสำหรับผู้ที่ต้องการขึ้นรูปโมเดลมาตรฐานได้อย่างรวดเร็ว พร้อมทั้งมีเอกสารคู่มือที่ครบถ้วนและเครื่องมือสนับสนุนการนำโมเดลไปใช้งานจริงบนแพลตฟอร์มต่าง ๆ ได้เป็นอย่างดี อย่างไรก็ตาม หากต้องการปรับแต่งการทำงานในส่วนลึกของโมเดลหรือสร้างกระบวนการฝึกสอนที่ซับซ้อนเป็นพิเศษ อาจทำได้ยากและซับซ้อนกว่าการใช้ PyTorch รวมถึงการตรวจสอบข้อผิดพลาดในบางกรณีอาจทำความเข้าใจได้ยากสำหรับผู้เริ่มต้น

1.5.2.5 กลุ่มไลบรารีเพื่อการจัดการข้อมูลและประมวลผลภาพ (NumPy, Pandas และ OpenCV) เป็นชุดเครื่องมือเสริมที่มีความสำคัญอย่างยิ่งในกระบวนการเตรียมข้อมูลภาพ CT Scan โดย NumPy และ Pandas ถูกนำมาใช้จัดการข้อมูลเชิงตัวเลขและอาร์เรย์มัลติมิติเพื่อให้การคำนวณทางคณิตศาสตร์หลังบ้านเป็นไปอย่างรวดเร็ว ในขณะที่ OpenCV และ Pillow ทำหน้าที่ในการปรับแต่งรูปภาพ (Image Preprocessing) เช่น การปรับขนาดภาพ (Resizing) และการทำ Normalization เพื่อให้อินพุตมีความพร้อมก่อนเข้าสู่แบบจำลอง แม้ไลบรารีเหล่านี้จะช่วยลดภาระการเขียนโค้ดพื้นฐานได้มหาศาล แต่ผู้พัฒนาจำเป็นต้องบริหารจัดการหน่วยความจำให้ดีเนื่องจากการประมวลผลภาพจำนวนมากพร้อมกันอาจส่งผลให้เกิดปัญหาทรัพยากรระบบไม่เพียงพอได้

1.5.2.6 กลุ่มไลบรารีเพื่อการวัดผลและแสดงผลสถิติ (Scikit-learn และ Matplotlib) เป็นเครื่องมือหลักที่ใช้ในการประเมินประสิทธิภาพของสถาปัตยกรรมทั้ง 5 รูปแบบ โดย Scikit-learn ช่วยในการคำนวณค่าตัวบ่งชี้ทางสถิติที่สำคัญ ได้แก่ Accuracy, Precision, Recall และ F1-Score รวมถึงการสร้าง Confusion Matrix เพื่อวิเคราะห์ความแม่นยำในการแยกแยะชนิดของมะเร็งปอด ส่วน Matplotlib และ Seaborn จะถูกนำมาใช้สร้างทัศนภาพข้อมูลผ่านกราฟแสดงแนวโน้มค่าความสูญเสีย (Loss) และความแม่นยำในแต่ละรอบการเทรน ช่วยให้สามารถวิเคราะห์ปัญหา Overfitting ได้อย่างชัดเจน แม้เครื่องมือกลุ่มนี้จะใช้งานง่ายและให้ผลลัพธ์ที่ละเอียด แต่ผู้ใช้ต้องมีความเข้าใจพื้นฐานทางสถิติและการแปลความหมายข้อมูลเพื่อให้สามารถสรุปผลการวิจัยได้อย่างถูกต้องและน่าเชื่อถือ

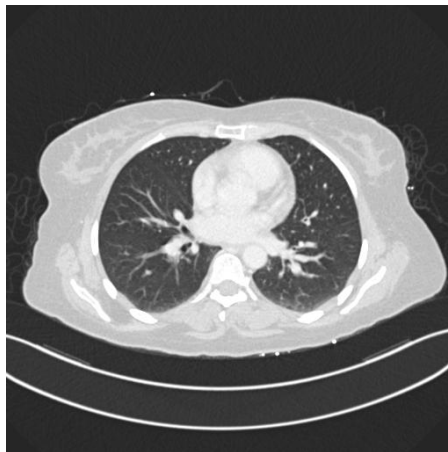
บทที่ 2

ทฤษฎีและผลงานที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 ความรู้พื้นฐานเกี่ยวกับโรคมะเร็งปอดและภาพถ่าย CT Scan

มะเร็งปอด (Lung Cancer) เกิดจากการเจริญเติบโตของเซลล์ที่ผิดปกติในปอด ซึ่งมักปรากฏในรูปแบบของจุดหรือก้อนเนื้อ (Nodules) โดยรอยโรคเหล่านี้สามารถจำแนกออกเป็น 3 ประเภทหลัก ได้แก่ เนื้อปกติ (Normal), เนื้อไม่ร้าย (Benign) และเนื้อร้าย (Malignant) การคัดกรองที่มีประสิทธิภาพที่สุดคือการใช้ ภาพถ่ายเอกซเรย์คอมพิวเตอร์ (Computed Tomography: CT Scan) ซึ่งเป็นการใช้รังสีเอกซ์หมุนรอบตัวผู้ป่วยเพื่อสร้างภาพภาคตัดขวางที่มีรายละเอียดสูง ช่วยให้แพทย์เห็นลักษณะทางกายภาพของรอยโรคได้ชัดเจนกว่าการเอกซเรย์ทั่วไป สำหรับชุดข้อมูล IQ-OTH/NCCD ที่ใช้ในงานวิจัยนี้ เป็นชุดข้อมูลสาธารณะที่รวบรวมภาพถ่าย CT Scan จากผู้ป่วยจริงในประเทศอิตาลี ประกอบด้วยเคสทั้งหมด 110 เคส แบ่งเป็นคลาส Malignant 40 เคส, Benign 15 เคส และ Normal 55 เคส ซึ่งเป็นข้อมูลมาตรฐานที่นักวิจัยทั่วโลกนิยมใช้ในการทดสอบอัลกอริทึมการตรวจจับมะเร็งปอด (alyasriy et al., n.d.)



ภาพประกอบที่ 2.1 ตัวอย่างภาพ CT SCAN

2.1.2 การประมวลผลภาพดิจิทัล (Digital Image Processing)

ก่อนนำภาพเข้าสู่โมเดลปัญญาประดิษฐ์ จำเป็นต้องผ่านกระบวนการเตรียมข้อมูล (Preprocessing) เพื่อเพิ่มประสิทธิภาพการเรียนรู้ ประกอบด้วย

- 1) การปรับขนาดภาพ (Image Resizing): เนื่องจากการออกแบบสถาปัตยกรรม CNN แต่ละรุ่นต้องการขนาด Input ที่แน่นอน เช่น 224x224 หรือ 299x299 พิกเซล เพื่อให้โครงสร้างเลเยอร์ภายในประมวลผลได้อย่างถูกต้อง

2) การปรับค่าสีและคอนทราสต์ (Normalization & Enhancement): การทำ Normalization จะปรับสเกลของค่าพิกเซลให้อยู่ในช่วง 0 ถึง 1 เพื่อช่วยให้โมเดลลู่เข้า (Converge) ได้เร็วขึ้น และการเพิ่มคอนทราสต์ช่วยให้รอยโรคในปอดมีความชัดเจนขึ้นจากเนื้อเยื่อโดยรอบ

3) การทำ Data Augmentation: เป็นเทคนิคการสร้างภาพจำลองเพิ่มเติมจากข้อมูลเดิม เช่น การหมุน (Rotation), การพลิก (Flipping) หรือการซูม (Zooming) เพื่อช่วยแก้ปัญหาการขาดแคลนข้อมูลและป้องกันการเกิด Overfitting (Goceri, 2023) สำหรับภาพทางการแพทย์ การทำ Data Augmentation ต้องระมัดระวังไม่ให้เปลี่ยนแปลงลักษณะทางคลินิกที่สำคัญของรอยโรค

2.1.3 ปัญญาประดิษฐ์และการเรียนรู้เชิงลึก (AI and Deep Learning)

การเรียนรู้เชิงลึก (Deep Learning) (ชิตพงษ์ กิตตินราทร, 2020) เป็นส่วนหนึ่งของ Machine Learning ที่พัฒนามาจากการแนวคิดโครงข่ายประสาทเทียม (Artificial Neural Networks) โดยเลียนแบบการทำงานของเซลล์ประสาทในสมองมนุษย์ โดยมีโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Networks: CNN) เป็นหัวใจหลักในการประมวลผลภาพ ซึ่งประกอบด้วยเลเยอร์สำคัญ 3 ส่วน ได้แก่

1. Convolution Layer: ทำหน้าที่สกัดลักษณะเด่น (Feature Extraction) โดยการใช้ตัวกรอง (Filters) วิ่งผ่านภาพ
2. Pooling Layer: ทำหน้าที่ลดขนาดของข้อมูล (Downsampling) เพื่อลดภาระการคำนวณแต่ยังคงรักษาฟีเจอร์สำคัญไว้
3. Fully Connected Layer: ทำหน้าที่รวบรวมฟีเจอร์ที่สกัดได้มาเพื่อตัดสินใจจำแนกประเภทภาพ (Classification)

2.1.4 สถาปัตยกรรมแบบจำลองการเรียนรู้เชิงลึกที่ใช้เปรียบเทียบ

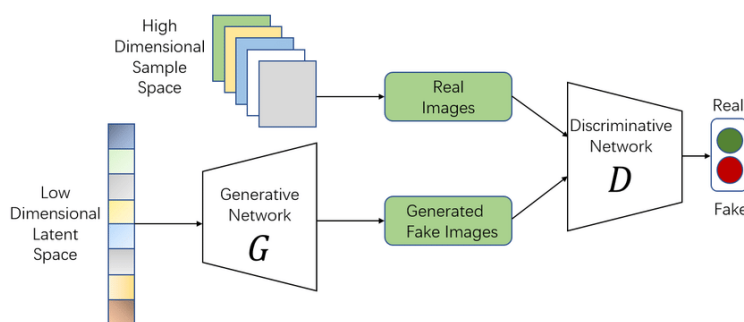
งานวิจัยนี้คัดเลือก 5 โมเดลชั้นนำที่มีโครงสร้างแตกต่างกันเพื่อเปรียบเทียบประสิทธิภาพ (Afroz, n.d.)

- 1) VGG16: โดดเด่นที่ความเรียบง่าย โดยใช้การวางซ้อนเลเยอร์ Convolution ขนาดเล็ก 3x3 อย่างต่อเนื่องเพื่อให้โมเดลมีความลึกสูง
- 2) ResNet50: นำเสนอเทคนิค Residual Learning (Skip Connection) เพื่อแก้ปัญหา Vanishing Gradient ทำให้สามารถเทรนโครงข่ายที่ลึกมากถึง 50 เลเยอร์ได้โดยไม่สูญเสียความแม่นยำ
- 3) DenseNet201: เน้นการเชื่อมต่อกันอย่างหนาแน่น โดยทุกเลเยอร์จะเชื่อมต่อกับทุกเลเยอร์ถัดไป ช่วยให้เกิดการใช้ Feature ซ้ำ (Feature Reuse) อย่างมีประสิทธิภาพ
- 4) Inception-V3: ใช้การประมวลผลแบบขนานในโมดูลเดียว (Parallel Paths) ทำให้สามารถสกัดฟีเจอร์ในหลายระดับขนาด (Multi-scale) ได้ในคราวเดียว
- 5) EfficientNet-CBAM: ใช้เทคนิค Compound Scaling เพื่อปรับขนาดของความกว้าง ความลึก และ

ความละเอียดภาพให้สมดุลกัน ทำให้โมเดลมีขนาดเล็กแต่ประสิทธิภาพสูง

2.1.5 โครงข่ายประสาทเทียมแบบสร้างสรรค์ (GAN) Generative Adversarial Networks

ประกอบด้วย 2 เครือข่ายที่แข่งกัน คือ Generator (สร้างภาพเสมือน) และ Discriminator (ตรวจสอบว่าเป็นภาพจริงหรือภาพปลอม) ในงานวิจัยนี้เลือกใช้ WGAN-GP (Wasserstein GAN with Gradient Penalty) ซึ่งใช้ Wasserstein Distance เป็นฟังก์ชันการสูญเสีย (Loss Function) และเพิ่มบทลงโทษการเปลี่ยนแปลงของค่า Gradient (Gradient Penalty) เพื่อช่วยให้การฝึกสอน GAN มีความเสถียรมากขึ้น ลดปัญหาการเกิด Mode Collapse และสร้างภาพถ่ายทางการแพทย์ที่มีคุณภาพสูงเพื่อแก้ปัญหา Data Imbalance



ภาพประกอบที่ 2.2 ตัวอย่างกระบวนการทำ GAN

2.1.6 เทคนิคการถ่ายโอนการเรียนรู้ (Transfer Learning)

เป็นกระบวนการนำองค์ความรู้จากแบบจำลองที่ได้รับการฝึกสอนมาก่อนแล้ว (Pre-trained Models) บนชุดข้อมูลขนาดใหญ่ เช่น ImageNet มาปรับปรุงให้เข้ากับงานเฉพาะทาง (Fine-tuning) เทคนิคนี้ช่วยให้โมเดลสามารถทำงานได้ดีแม้จะมีข้อมูลภาพทางการแพทย์จำนวนจำกัด โดยการแช่แข็ง (Freeze) เลเยอร์ส่วนแรกๆ ที่เรียนรู้รูปร่างพื้นฐาน และฝึกสอนใหม่เฉพาะเลเยอร์ส่วนปลายเพื่อให้จดจำลักษณะรอยโรคเฉพาะโรค (Shubham Saxena, 2024)

2.1.7 การวัดประสิทธิภาพของแบบจำลอง (Performance Evaluation Metrics)

Confusion Matrix (Nisha Arya Ahmed, 2024) คือตารางสำคัญในการวัดความสามารถของ machine learning ในการแก้ปัญหา classification การที่จะนำ Model ไปใช้งานจริงได้นั้น จำเป็นต้องมีการวัดประสิทธิภาพ Model ก่อนว่า Model นั้นมีประสิทธิภาพเพียงพอที่จะนำมาพัฒนา หรือนำไปใช้งานด้านต่างๆหรือไม่ ซึ่งการวัดประสิทธิภาพนั้นส่วนใหญ่จะวัดค่าจากใน Table ข้อมูลที่มี (Confusion Matrix)

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

ภาพประกอบที่ 2.3 ตาราง Confusion Matrix

True Positive (TP) คือ สิ่งที่โปรแกรมทำนายว่า “จริง” และมีค่าเป็น “จริง ”

True Negative (TN) คือ สิ่งที่โปรแกรมทำนายว่า “ไม่จริง” และมีค่า “ไม่จริง ”

False Positive (FP) คือ สิ่งที่โปรแกรมทำนายว่า “จริง” แต่ มีค่าเป็น “ไม่จริง”

False Negative (FN) คือ สิ่งที่โปรแกรมทำนายว่า “ไม่จริง” แต่ มีค่าเป็น “จริง”

การประเมินค่าความแม่นยำ

- 1) Accuracy: ความแม่นยำโดยรวมของโมเดล

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

สมการที่ 2.1 สมการ Accuracy

- 2) Precision: ความถูกต้องในการทำนายคลาสบวก (ทำนายว่าป่วยแล้วป่วยจริงเท่าไร)

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

สมการที่ 2.2 สมการ Precision

- 3) Recall (Sensitivity): ความสามารถในการตรวจหาผู้ป่วยทั้งหมด (สำคัญมากในทางการแพทย์ เพื่อไม่ให้หลุดเคสผู้ป่วยจริง)

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

สมการที่ 2.3 สมการ Recall (Sensitivity)

- 4) F1-score: ค่าเฉลี่ยแบบ Harmonic ระหว่าง Precision และ Recall เพื่อดูความสมดุลของโมเดล

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

สมการที่ 2.4 สมการ F1-score

TP + TN: คือจำนวนครั้งที่ทั้งหมดที่โมเดล "ทายถูก" ไม่ว่าจะเป็นคลาสบวกหรือลบ

FP + FN: คือจำนวนครั้งที่ทั้งหมดที่โมเดล "ทายผิด" (Total Errors) ซึ่งส่งผลต่อค่าความเชื่อมั่นของโมเดล

TP + FN: คือจำนวน "ผู้ป่วยจริงทั้งหมด" ที่มีอยู่ในชุดข้อมูล (Actual Positive)

TP + TN + FP + FN: คือจำนวน "ข้อมูลทั้งหมด" (Total Samples) ที่นำมาทดสอบ

2.1.8 ขั้นตอนการดำเนินงานตามมาตรฐาน CRISP-DM (Research Methodology)

วงจร CRISP-DM เป็นกรอบการทำงานมาตรฐานสากลสำหรับโครงการ Data Mining และ Machine Learning ประกอบด้วย 6 ขั้นตอนหลักที่เชื่อมโยงกันเป็นวงกลม ซึ่งหมายความว่าเราสามารถย้อนกลับไปปรับปรุงขั้นตอนก่อนหน้าได้เสมอ

1) ความเข้าใจในเป้าหมาย (Business/Research Understanding)

ขั้นตอนนี้คือการกำหนดวัตถุประสงค์ของโครงการ คือการตั้งเป้าหมายว่าเราจะสร้างระบบตรวจจับมะเร็งปอดที่มีความแม่นยำสูง โดยใช้ชุดข้อมูลที่มีจำกัดและไม่สมดุล (IQ-OTH/NCCD)

2) ความเข้าใจในข้อมูล (Data Understanding)

การรวบรวมและสำรวจข้อมูลเบื้องต้น คือการวิเคราะห์ชุดข้อมูล IQ-OTH/NCCD ว่ามีกี่คลาส (Malignant, Benign, Normal) มีสัดส่วนเท่าไร และลักษณะของภาพ CT Scan เป็นอย่างไร เพื่อหาปัญหา (เช่น พบว่าคลาส Benign น้อยเกินไป)

3) การเตรียมข้อมูล (Data Preparation)

เป็นขั้นตอนที่สำคัญที่สุดและใช้เวลามากที่สุด (ประมาณ 70-80% ของงาน) การจัดการภาพ: Image Resizing, Normalization และ CLAHE การสังเคราะห์ข้อมูล แก้ปัญหา Data Imbalance และการทำ Data Augmentation แบบพื้นฐาน

4) การสร้างแบบจำลอง (Modeling)

การเลือกและฝึกสอนอัลกอริทึม คือการนำสถาปัตยกรรมทั้ง 5 รุ่น (VGG16, ResNet50, DenseNet201, Inception-V3, EfficientNet) มาติดตั้ง CBAM Module และทำการฝึกสอน (Training) โดยใช้เทคนิค Transfer Learning

5) การประเมินผล (Evaluation)

การตรวจสอบว่าโมเดลที่สร้างขึ้นตอบโจทย์ในขั้นตอนที่ 1 หรือไม่ การวัดค่าด้วย Accuracy, Precision, Recall, และ F1-score รวมถึงการวิเคราะห์ Confusion Matrix เพื่อดูว่าโมเดลทำนายผิดพลาดที่จุดไหน

6) การนำไปใช้งาน/สรุปผล (Deployment/Reporting)

การนำโมเดลที่ได้ไปประยุกต์ใช้หรือเขียนสรุปผลงานวิจัย คือการสรุปว่าเทคนิค EfficientNet-CBAM ร่วมกับ GAN มีประสิทธิภาพเพียงพอที่จะนำไปช่วยรังสีแพทย์ในการวินิจฉัยมะเร็งปอดได้จริง



ภาพประกอบที่ 2.4 วงจรการทำงาน CRISP-DM (Research Methodology)

2.2 วิจัยที่เกี่ยวข้อง

2.2.1 งานวิจัยด้านระบบการวินิจฉัยแบบไฮบริดและโครงสร้างแบบครบวงจร

Samiul Based Shuvo และ Tasnia Binte (2026) (Shuvo & Mamun, 2025) นำเสนอเฟรมเวิร์กการเรียนรู้เชิงลึกแบบอัตโนมัติครบวงจร (End-to-end) เพื่อยกระดับความแม่นยำในการวินิจฉัยมะเร็งปอดระยะเริ่มต้น โดยมีวัตถุประสงค์หลักเพื่อแก้ปัญหาข้อจำกัดในพื้นที่ที่มีทรัพยากรทางการแพทย์และรังสีแพทย์จำกัด ซึ่งการวินิจฉัยแบบเดิมมักประสบปัญหาความล่าช้าและความผิดพลาดจากการเหนื่อยล้าของผู้วินิจฉัย รวมถึงข้อจำกัดของระบบคอมพิวเตอร์ช่วยวินิจฉัย (CAD) แบบดั้งเดิมที่ยังเผชิญกับปัญหาความไม่สมดุลของข้อมูล (Class imbalance) และความแปรปรวนของลักษณะรอยโรค ผู้วิจัยได้ดำเนินการพัฒนาโมเดลผ่านชุดข้อมูลมาตรฐาน LUNA16 (ซึ่งเป็นส่วนหนึ่งของ LIDC-IDRI) ประกอบด้วยภาพถ่าย CT Scan จำนวน 888 ราย พร้อมการระบุตำแหน่งและระดับความรุนแรงของรอยโรคโดยรังสีแพทย์ผู้เชี่ยวชาญ ในส่วนของกระบวนการเตรียมข้อมูลภาพ (Image Processing) ซึ่งถือเป็นหัวใจสำคัญของงานวิจัยนี้ ได้มีการดำเนินการอย่างเป็นระบบเริ่มจากการปรับค่าความหนาแน่นของเนื้อเยื่อ (Hounsfield Unit: HU) ให้อยู่ในช่วง [-1200, 600] เพื่อให้เหมาะสมกับการวิเคราะห์เนื้อเยื่อปอดโดยเฉพาะ ตามด้วยการทำ Min-Max Normalization เพื่อปรับสเกลค่าพิกเซลให้อยู่ในช่วง สำหรับความเสถียรในการฝึกสอนโมเดล นอกจากนี้ยังมีการทำ Binarization ของ Lung Mask และการปรับขนาดภาพ (Resizing) รวมถึงการปรับระยะห่างระหว่างพิกเซล (Resampling) ให้เป็น 1 มม. ในทุกมิติ ($256 \times 256 \times 256$) เพื่อสร้างมาตรฐานเดียวกันของข้อมูล กระบวนการสร้างโมเดลแบ่งออกเป็น 3 ขั้นตอนหลัก (Three-stage pipeline) เริ่มจากการใช้ 3D Res-U-Net ในการแยกส่วนปอด (Segmentation) เพื่อลดสัญญาณรบกวนจากอวัยวะข้างเคียง จากนั้นจึงใช้ YOLOv5s ในการตรวจจับก้อนเนื้อ (Nodule Detection) บนภาพ axial slices และปิดท้ายด้วยการใช้ Vision Transformer (ViT) สำหรับการจำแนกประเภท (Classification) ว่าเป็นเนื้อดีหรือเนื้อร้ายโดยอาศัยกลไก Self-attention ในการดึงคุณลักษณะเด่น ผลการทดลองพบว่าระบบสามารถทำประสิทธิภาพได้อย่างยอดเยี่ยม โดยมีค่า Dice score ในการแยกส่วนปอดสูงถึง 98.82% และโมเดล ViT สามารถทำความเข้าใจ (Accuracy) ได้ถึง 96.29% ซึ่งสูงกว่าโมเดลมาตรฐานเดิมถึง 4.25% นอกจากนี้ยังมีการใช้ตัววัดประสิทธิภาพอื่นๆ อย่างครอบคลุม อาทิ Precision, Recall, F1-score และค่า ROC-AUC ที่สูงถึง 0.989 รวมถึงการใช้เทคนิค Attention-rollout heatmaps เพื่ออธิบายผลการตัดสินใจของโมเดล (Explainability) ช่วยเพิ่มความเชื่อถือในเชิงการแพทย์ อย่างไรก็ตาม งานวิจัยนี้ยังมีช่องว่าง (Gap) ที่สำคัญคือความต้องการทรัพยากรคำนวณที่สูง (High computational resources) และยังขาดการทดสอบในสภาพแวดล้อมทางคลินิกจริง (Prospective clinical validation) รวมถึงความหลากหลายของชุดข้อมูลที่ยังจำกัดเฉพาะบางแหล่ง ซึ่งประเด็นเหล่านี้สามารถนำมาประยุกต์ใช้เป็นแนวทางในการต่อยอดงานวิจัย โดยเฉพาะการพัฒนาโมเดลให้มีความเบาบางลง แต่ยังคงความแม่นยำ หรือการใช้เทคนิค Generative AI เพื่อสร้างข้อมูลจำลองที่หลากหลายขึ้นเพื่ออุดช่องว่างความแตกต่างของเครื่องสแกนในแต่ละสถาบันได้

Shenson Joseph และคณะ (2025) (Joseph et al., 2025) นำเสนอสถาปัตยกรรมไฮบริดขั้นสูง เพื่อเพิ่มประสิทธิภาพในการตรวจจับมะเร็งปอด ซึ่งเป็นหนึ่งในสาเหตุหลักของการเสียชีวิตทั่วโลกเนื่องจากความซับซ้อนในการระบุโรคในระยะลุกลาม โดยผู้วิจัยได้ระบุถึงปัญหาสำคัญของงานวิจัยในอดีตที่มักเผชิญกับข้อจำกัดด้านขนาดของชุดข้อมูลที่เล็กเกินไป (Small-scale datasets) ความสามารถในการสรุปผลไปยังข้อมูลที่ไม่เคยเห็นมาก่อน (Generalizability) ที่ยังไม่ดีพอ รวมถึงปัญหาภาพที่มีความเปรียบต่างต่ำ (Low contrast) และสัญญาณรบกวนในภาพถ่าย CT Scan ในส่วนของกระบวนการเตรียมข้อมูลภาพ (Image Processing) ผู้วิจัยให้ความสำคัญอย่างยิ่งกับการปรับปรุงคุณภาพภาพถ่ายเริ่มต้นผ่านการปรับขนาดภาพ (Image resizing) ให้เป็นรูปสี่เหลี่ยมจัตุรัสเพื่อความเหมาะสมของโครงสร้างเครือข่ายประสาทเทียม และการทำ Normalization เพื่อขยายช่วงความเข้มพิกเซลให้ครอบคลุมและเหมาะสมต่อการประมวลผล นอกจากนี้ยังได้บูรณาการเทคนิคการเพิ่มขยายข้อมูลเชิงสังเคราะห์ (Data augmentation) ด้วยแบบจำลอง GAN ที่ผ่านการฝึกสอนด้วยอัลกอริทึม Elk Herd Optimizer (EHO) เพื่อสร้างภาพ CT Scan จำลองที่มีความสมจริงสูง ช่วยแก้ปัญหการขาดแคลนข้อมูลที่มีการติดป้ายกำกับและสร้างสมดุลให้กับชุดข้อมูล กระบวนการสร้างแบบจำลองดำเนินการผ่านชุดข้อมูลมาตรฐาน "Chest CT-Scan images Dataset" ซึ่งรวบรวมภาพจำแนกออกเป็น 4 ประเภท ได้แก่ Adenocarcinoma, Large cell carcinoma, Squamous cell carcinoma และกลุ่มปกติโดยขั้นตอนการดำเนินงานประกอบด้วยการนำภาพที่ผ่านการ Pre-processing และ Augmentation เข้าสู่โมเดลไฮบริด GA-AGN ซึ่งเป็นการผสมผสานระหว่าง GAN และ Attention Gated Network (AGN) โดย AGN จะทำหน้าที่เพิ่มพลังในการแสดงผลข้อมูลผ่านกลไกการให้น้ำหนักความสำคัญ (Compatibility score) เพื่อเฟ้นงัดไปยังภูมิภาคพยาธิสภาพที่สำคัญและกีดสัญญาณรบกวนจากพื้นหลัง จากนั้นผลลัพธ์จากทั้งสองส่วนจะถูกนำมาเชื่อมต่อกัน (Concatenate) และผ่านชั้น Fully Connected เพื่อทำนายผลลัพธ์สุดท้ายผ่านฟังก์ชัน Softmax ผลการทดลองพบว่าแบบจำลอง GA-AGN ให้ประสิทธิภาพที่ยอดเยี่ยมเหนือกว่าแบบจำลองมาตรฐานอื่นอย่างชัดเจน โดยใช้ค่า ความแม่นยำ (Accuracy) เป็นตัววัดหลักซึ่งทำได้สูงถึง 0.938 ร่วมกับการรายงานผลตัววัดอื่นอย่างละเอียด ได้แก่ ค่าความไว (Sensitivity) 0.948, ค่าความจำเพาะ (Specificity) 0.927 และค่าความเที่ยงตรง (Precision) 0.944 นอกจากนี้ยังพบว่าใช้เวลาในการประมวลผลเพียง 808.52 วินาที ซึ่งรวดเร็วกว่าวิธีเดิมอย่างมีนัยสำคัญ อย่างไรก็ตาม งานวิจัยนี้ยังมีช่องว่าง (Gap) ที่สำคัญคือแบบจำลองยังไม่สามารถแยกแยะความแตกต่างระหว่างก้อนเนื้อชนิดไม่ร้ายแรงในระยะเริ่มต้น (Early benign nodules) กับรอยโรคเนื้อร้ายขนาดเล็ก (Small malignant lesions) ได้อย่างชัดเจนรวมถึงการวิจัยยังจำกัดอยู่เพียงการใช้ข้อมูลภาพถ่ายเพียงอย่างเดียวโดยไม่ได้บูรณาการข้อมูลประวัติคนไข้หรือรายงานทางพันธุกรรมเข้ามาร่วมด้วย และยังมี ความกังวลด้านการนำไปใช้จริงในสภาพแวดล้อมทางคลินิกที่มีความแปรปรวนของเครื่องสแกนที่ต่างกัน ประเด็นเหล่านี้จึงเป็นโอกาสสำคัญในการประยุกต์ใช้เพื่อต่อยอดงานวิจัย โดยเฉพาะการพัฒนาโมเดลให้สามารถจำแนกกระตบความรุนแรงของรอยโรคขนาดเล็กได้แม่นยำยิ่งขึ้นผ่านเทคนิคการเพิ่มคุณภาพภาพด้วย SRGAN และการใช้กลไก CBAM Attention เพื่อโฟกัสคุณลักษณะทางพยาธิสภาพที่ซับซ้อนซึ่งสอดคล้องกับข้อเสนอแนะในการพัฒนาความสามารถในการวินิจฉัยระยะเริ่มต้นให้มีประสิทธิภาพสูงสุดตามที่ผู้วิจัยได้ระบุไว้

Mobarak Abumohsen และคณะ (2025) (Abumohsen et al., 2025) ได้เสนอการใช้ระบบ Hard-voting Ensemble โดยรวบรวมจุดเด่นของหลายโมเดลเข้าด้วยกันเพื่อให้ได้ผลลัพธ์ที่ดีกว่าการใช้โมเดลเดี่ยว โดยเฉพาะอย่างยิ่งการผสมผสานระหว่าง Inception-V3 ซึ่งเป็นโมเดลที่มีความสามารถพิเศษในการสกัดฟีเจอร์หลายสเกล (multi-scale feature extraction) ผ่านการใช้ inception modules ที่ประกอบด้วย convolutional filters หลายขนาด (1x1, 3x3, 5x5) ทำงานแบบ parallel และรวมผลลัพธ์เข้าด้วยกัน ทำให้สามารถจับฟีเจอร์ที่มีขนาดต่างกันได้พร้อมกันในชั้นเดียว ซึ่งเป็นประโยชน์อย่างมากในการวิเคราะห์ภาพ CT Scan ที่มีรอยโรคขนาดต่างๆ ตั้งแต่เล็กมากไปจนถึงขนาดใหญ่ และ DenseNet201 ซึ่งเป็นโมเดลที่มีจุดเด่นในการเชื่อมต่อข้อมูลระหว่างเลเยอร์อย่างมีประสิทธิภาพผ่านการใช้ dense connections ที่ทำให้ทุก layer รับ input จากทุก layers ที่อยู่ก่อนหน้า ส่งผลให้มีการไหลของข้อมูลและ gradient ที่ดีมาก ลด parameter redundancy และช่วยให้โมเดลเรียนรู้ได้ดีแม้จะมีข้อมูลจำกัด การใช้ hard-voting ensemble หมายความว่าแต่ละโมเดลจะทำการทำนายแยกกันและผลลัพธ์สุดท้ายจะถูกตัดสินจากการลงคะแนนแบบเสียงข้างมากของทั้งสองโมเดล ซึ่งวิธีนี้ช่วยลดความผิดพลาดที่อาจเกิดจากข้อจำกัดของโมเดลใดโมเดลหนึ่ง เพราะแต่ละโมเดลอาจมีจุดแข็งและจุดอ่อนที่ต่างกัน ผลลัพธ์ของงานวิจัยแสดงให้เห็นว่าแนวทางแบบผสมผสานนี้ช่วยลดความผิดพลาดในการระบุก้อนเนื้อขนาดเล็ก (small nodules) ได้อย่างมีประสิทธิภาพสูงกว่าการใช้โมเดลเดี่ยวอย่างเห็นได้ชัด โดยเฉพาะในกรณีของ false negative rate ซึ่งเป็นปัญหาสำคัญในการวินิจฉัยมะเร็งปอด เพราะการพลาดการตรวจพบมะเร็งอาจส่งผลกระทบร้ายแรงต่อชีวิตผู้ป่วย งานวิจัยนี้ยังได้ทำการวิเคราะห์เปรียบเทียบ computational cost ระหว่าง ensemble และ single model และพบว่าแม้ ensemble จะใช้เวลาในการประมวลผลมากกว่าเล็กน้อย แต่ความแม่นยำที่เพิ่มขึ้นคุ้มค่างับต้นทุนที่เพิ่มขึ้น โดยเฉพาะในบริบทของการวินิจฉัยทางการแพทย์ที่ความแม่นยำมีความสำคัญสูงสุด

2.2.2 งานวิจัยด้านการเพิ่มคุณภาพภาพและความละเอียดสูง

Nahed Tawfik และคณะ (2024) (Tawfik et al., 2024) นำเสนอแนวทางการยกระดับประสิทธิภาพของระบบคอมพิวเตอร์ช่วยวินิจฉัย (CAD) เพื่อการตรวจจับมะเร็งปอดในระยะเริ่มต้นจากภาพถ่ายเอกซเรย์คอมพิวเตอร์ (CT Scans) ซึ่งมีจุดมุ่งหมายหลักในการลดอัตราการเสียชีวิตที่สูงเป็นอันดับต้นๆ ของโลก โดยงานวิจัยนี้มุ่งเน้นการแก้ปัญหาความซับซ้อนของเนื้อเยื่อปอดและสัญญาณที่บ่งบอกถึงโรคในระยะแรกที่มีภูมิความคลุมเครือ รวมถึงข้อจำกัดของโมเดลการเรียนรู้เชิงลึกแบบเดิมที่มักเผชิญกับอัตราการเกิดผลบวกปลอม (False Positives) ที่สูงและความยากลำบากในการคัดเลือกสถาปัตยกรรมโมเดลที่เหมาะสม

ในส่วนของการเตรียมข้อมูลภาพ (Image Processing) ซึ่งเป็นจุดเน้นสำคัญของงานวิจัยนี้ ผู้วิจัยได้เลือกใช้เทคนิค Contrast Limited Adaptive Histogram Equalization (CLAHE) เพื่อปรับปรุงคุณภาพของภาพถ่าย CT Scan ให้มีความคมชัดสูงขึ้นโดยขั้นตอนการทำงานของ CLAHE เริ่มจากการแบ่งภาพต้นฉบับออกเป็นส่วนย่อยๆ ที่ไม่ซ้อนทับกัน (Tiles) จากนั้นจึงคำนวณฮิสโตแกรมของแต่ละส่วนเพื่อวิเคราะห์การกระจายตัวของระดับความเข้มแสง จุดเด่นสำคัญคือการใช้กลไกการจำกัดความแปรปรวน (Clipping

Mechanism) เพื่อป้องกันไม่ให้เกิดการขยายสัญญาณรบกวน (Noise) ในบริเวณที่ภาพมีความสม่ำเสมอ ซึ่งผู้วิจัยได้นำนวัตกรรมการหาค่าที่เหมาะสมด้วยอัลกอริทึม Modified CMA-ES (Covariance Matrix Adaptation Evolution Strategy) มาใช้ในการกำหนดค่า Clip Limit และ Tile Grid Size ที่ดีที่สุด เพื่อให้ได้ภาพที่มีความสมดุลระหว่างความคมชัดและการรักษาคุณลักษณะดั้งเดิมของรอยโรคไว้ได้อย่างครบถ้วน ก่อนจะปิดท้ายด้วยการใช้กระบวนการ Bilinear Interpolation เพื่อเชื่อมต่อรอยต่อระหว่างส่วนย่อยให้มีความต่อเนื่องและเรียบเนียน กระบวนการสร้างแบบจำลองได้ดำเนินการผ่านชุดข้อมูลมาตรฐาน (Dataset) ที่รวบรวมโดย Mohammad Hany ซึ่งประกอบด้วยภาพถ่าย CT Scan จำนวน 1,000 ภาพ ครอบคลุมมะเร็ง 3 ชนิด ได้แก่ Adenocarcinoma, Large cell carcinoma, และ Squamous cell carcinoma รวมถึงกลุ่มเซลล์ปกติ

ผู้วิจัยได้แบ่งข้อมูลออกเป็นส่วนการฝึกสอน 70% การทดสอบ 20% และการตรวจสอบผล 10% โดยมีการบูรณาการเทคนิคการเพิ่มขยายข้อมูลแบบ CutMix Data Augmentation ซึ่งใช้การตัดบางส่วนของภาพหนึ่งไปวางทับอีกภาพหนึ่งพร้อมกับการผสมผสานป้ายกำกับ (Labels) ตามสัดส่วนพื้นที่ เพื่อลดปัญหาการเกิด Overfitting และช่วยให้โมเดลมีความทนทานต่อความแปรปรวนของข้อมูลมากขึ้น โดยมีการเปรียบเทียบสถาปัตยกรรมระดับโลกถึง 4 รูปแบบ ได้แก่ Inception-V3, ResNet101, EfficientNet และ Xception ผลการทดสอบประสิทธิภาพ พบว่าโมเดล Xception ที่ผ่านการประมวลผลด้วย Optimized CLAHE ให้ผลลัพธ์ที่ยอดเยี่ยมที่สุด โดยมีตัววัดประสิทธิภาพหลักคือ ความแม่นยำ (Accuracy) สูงถึง 99.03% และความไว (Sensitivity) ที่ 99.03% นอกจากนี้ผู้วิจัยยังได้ประเมินผลอย่างละเอียดผ่านตัววัดที่หลากหลาย ประกอบด้วย ความจำเพาะ (Specificity) ที่ 99.68%, ความเที่ยงตรง (Precision) 99.04%, อัตราการเกิดผลบวกปลอม (FPR) ที่ต่ำเพียง 0.0032, ค่าสัมประสิทธิ์สหสัมพันธ์แมตทิวส์ (MCC) 98.71%, ค่า F1_score 99.02%, และค่า Cohen's Kappa ที่ 97.41% ซึ่งบ่งบอกถึงความสอดคล้องของโมเดลในระดับเกือบสมบูรณ์ อีกทั้งยังยืนยันผลด้วยการวิเคราะห์พื้นที่ใต้กราฟ (AUC-ROC) เพื่อแสดงความสามารถในการแยกแยะคลาสต่างๆ ได้อย่างเที่ยงตรง อย่างไรก็ตาม ช่องว่าง (Gap) ที่พบจากงานวิจัยนี้คือผู้วิจัยได้ระบุไว้ว่ามีโอกาสในการพัฒนาความน่าเชื่อถือของระบบเพื่อใช้งานในสถานการณ์ฉุกเฉินทางการแพทย์ (Emergency situations) และการขยายผลไปสู่การวินิจฉัยอวัยวะอื่นๆ รวมถึงยังคงมีพื้นที่สำหรับการพัฒนาความแม่นยำผ่านเทคนิคการเตรียมข้อมูลภาพในรูปแบบที่หลากหลายกว่าเดิม

ประเด็นนี้จึงสอดคล้องกับแนวทางที่เรานำมาประยุกต์ใช้ โดยการเลือกใช้เทคนิค SRGAN เพื่อเพิ่มความละเอียดของภาพในระดับที่สูงขึ้นกว่า CLAHE แบบปกติ และการใช้กลไก CBAM Attention เข้ามาช่วยในการคัดเลือกคุณลักษณะ (Feature Selection) อย่างเจาะจง ซึ่งจะเป็นการอุดช่องว่างในเรื่องของการระบุพิกัดรอยโรคขนาดเล็กที่มีความซับซ้อนสูงได้ดียิ่งขึ้นตามที่งานวิจัยอ้างอิงได้เปิดประเด็นไว้สำหรับการศึกษาในอนาคต

Hongbo Zhu และคณะ (2021) (Zhu et al., 2021) นำเสนอนวัตกรรมเพื่อแก้ปัญหาความท้าทายสำคัญในการตรวจจับก้อนเนื้อในปอดระยะเริ่มต้นผ่านระบบคอมพิวเตอร์ช่วยวินิจฉัย (CADs) ซึ่งมักเผชิญกับข้อจำกัดด้านปริมาณพิกเซลที่น้อยเกินไปในภาพถ่ายเอกซเรย์คอมพิวเตอร์ความละเอียดต่ำ (Low-resolution

CT) ส่งผลให้ก้อนเนื้อขนาดเล็กเสี่ยงต่อการถูกมองข้ามในการวินิจฉัย, ผู้วิจัยได้วิเคราะห์ว่าแม้เทคนิคการเพิ่มความละเอียดของภาพ (Super-resolution: SR) แบบดั้งเดิมรวมถึง SRGAN จะสามารถสร้างภาพที่มีความสมจริงในเชิงทัศนวิสัย (Photo-realism) ได้ดี แต่ภาพเหล่านั้นยังขาดความแม่นยำในเชิงโครงสร้างทางพยาธิสภาพหรือความสมจริงเชิงหน้าที่ทางการแพทย์ (Functional-realism) ซึ่งมีความจำเป็นอย่างยิ่งต่อการวินิจฉัยโรค, ผู้วิจัยจึงเสนอเฟรมเวิร์ก FRGAN (Functional-Realistic GAN) ซึ่งเป็นสถาปัตยกรรม Deep GAN ที่ทำงานร่วมกับโมเดลไวยากรณ์เชิงลึก (Deep grammar model) เพื่อสร้างภาพที่มีรายละเอียดทางพยาธิสภาพที่ถูกต้องตามความรู้ของแพทย์

ในส่วนของการเตรียมข้อมูลภาพ (Image Processing) ผู้วิจัยได้ให้ความสำคัญกับการสกัดคุณลักษณะทางอรรถศาสตร์ (Semantic features) โดยใช้เครือข่าย Region Proposal Network (RPN) ที่มีพื้นฐานมาจาก VGG-16Net ทำหน้าที่ระบุตำแหน่งของคำศัพท์ทางอรรถศาสตร์ 9 ประการ เช่น ความขรุขระ (Spiculation), การเป็นแฉก (Lobulation), และความเป็นเนื้อร้าย (Malignancy) เพื่อสร้างหน่วยพื้นฐานของโครงสร้างเชิงหน้าที่, กระบวนการนี้มีการใช้เทคนิค Deconvolutional layer เพื่อเพิ่มรายละเอียดให้กับการบรรยายลักษณะภาพ และกำหนดสมอภาพ (Anchors) หลายขนาดเพื่อตรวจจับคำศัพท์ที่มีขนาดต่างกัน นอกจากนี้ยังได้ใช้เทคนิคการเพิ่มขยายข้อมูล (Data Augmentation) เพื่อแก้ปัญหาความไม่สมดุลของกลุ่มตัวอย่าง (Sample imbalance) ระหว่างภาพปกติและภาพที่มีรอยโรค โดยการใช้ภาพสังเคราะห์ที่มีความละเอียดสูงมาปรับสมดุลสัดส่วนให้เป็น 1:1, การดำเนินงานวิจัยนี้ได้ทดสอบบนชุดข้อมูล (Dataset) ที่รวบรวมขึ้นใหม่ร่วมกับชุดข้อมูลมาตรฐาน LUNA16 รวมทั้งสิ้น 1,034 สแกน ซึ่งครอบคลุมก้อนเนื้อที่มีขนาดและระดับความรุนแรงที่หลากหลาย

สำหรับขั้นตอนการสร้างโมเดล ผู้วิจัยได้ประยุกต์ใช้แนวคิดจาก TreeGAN เพื่อสร้างชุดกฎการสร้าง (Production rules) ที่มีพื้นฐานมาจากเวชระเบียนอิเล็กทรอนิกส์ (EMR) โดยใช้หน่วย LSTM แบบวนลูปเป็นตัวสร้างต้นไม้ (Tree generator), โมเดล FRGAN ทำหน้าที่เป็น Generator แบบ image-to-image ที่ทำงานควบคู่ไปกับตัวสร้างโครงสร้างอรรถศาสตร์ โดยใช้ And-Or Graphs (AOGs) ในการจัดเก็บความสัมพันธ์เชิงพื้นที่และบริบทของพยาธิสภาพ, ในส่วนของตัวจำแนก (Discriminator) ได้ใช้สถาปัตยกรรม Child-sum Tree-LSTM เพื่อตรวจสอบความถูกต้องของรูปแบบไวยากรณ์ที่สร้างขึ้น พร้อมทั้งคำนวณค่า Functional loss เพื่อให้มั่นใจว่าภาพที่สร้างขึ้นใหม่นั้นรักษาข้อมูลทางการแพทย์ที่สำคัญไว้ได้อย่างครบถ้วน, ผลการทดลองพบว่า FRGAN สามารถสร้างภาพความละเอียดสูง (SR images) ที่มีความสมจริงเชิงหน้าที่และมีลักษณะทางพยาธิสภาพที่ชัดเจนกว่าโมเดล MCGAN และ SRGAN แบบดั้งเดิมอย่างเห็นได้ชัด, ผู้วิจัยได้ใช้ตัววัดประสิทธิภาพที่ครอบคลุมหลายมิติ ได้แก่ ค่าพื้นที่ใต้กราฟ FROC (Free Receiver Operating Characteristic), ความแม่นยำในการจำแนกความรุนแรง (Accuracy), ค่า PSNR, ค่า SSIM และค่าความสูญเสียเชิงหน้าที่ (Functional loss) โดยให้ความสำคัญกับค่า CPM (Competition Performance Metric) score เป็นตัวชี้วัดหลัก ซึ่งผลการทดสอบบนชุดข้อมูลรวมสามารถทำคะแนน CPM ได้สูงถึง 0.915 ซึ่งเหนือกว่าวิธีการมาตรฐานอื่นๆ อย่าง Multi-view CNN และ CUMedVis

อย่างไรก็ตาม ช่องว่าง (Gap) ของงานวิจัยนี้คือความซับซ้อนของสถาปัตยกรรมที่ส่งผลให้ต้นทุนในการคำนวณ และพื้นที่จัดเก็บข้อมูลสูง (High computational and storage costs) รวมถึงความต้องการโมเดลที่เบาบางลงเพื่อการใช้งานจริงที่รวดเร็วยิ่งขึ้น ประเด็นนี้จึงสอดคล้องกับแนวทางที่สามารถนำมาประยุกต์ใช้ในการต่อยอด โดยการนำเทคนิคการเพิ่มคุณภาพภาพด้วย SRGAN มาบูรณาการร่วมกับกลไก CBAM Attention เพื่อเพิ่มความแม่นยำในการระบุตำแหน่งรอยโรคขนาดเล็กในภาพ CT Scan ในขณะเดียวกันก็อาจพิจารณาการใช้สถาปัตยกรรมที่มีประสิทธิภาพ (Lightweight architectures) เพื่ออุดช่องว่างด้านการใช้ทรัพยากรตามที่ผู้วิจัยได้เสนอแนะไว้

Hongqiang Xie และคณะ (2021) (Xie et al., 2021) นำเสนอแบบจำลองการเพิ่มความละเอียดของภาพ (Super-resolution: SR) สำหรับภาพถ่ายเอกซเรย์คอมพิวเตอร์ (CT) ของโรคปอดอักเสบจากเชื้อ *Pneumocystis carinii* (PCP) ซึ่งเป็นโรคที่มีความรุนแรงและมีอัตราการเสียชีวิตสูงหากไม่ได้รับการวินิจฉัยอย่างรวดเร็ว, โดยผู้วิจัยได้ระบุถึงปัญหาสำคัญของแบบจำลอง SRGAN พื้นฐานและอัลกอริทึมการเรียนรู้เชิงลึกในอดีตที่มักประสบปัญหาการเกิดสิ่งแปลกปลอมที่พร่ามัว (Blurry artifacts) ในการสร้างรายละเอียดพื้นผิว และการใช้ค่า Perceptual loss จากชั้นเปิดใช้งาน (Activation layer) ของ VGG Network ที่มีความเบาบางเกินไปจนไม่สามารถดักจับรายละเอียดความถี่สูงที่จำเป็นต่อการวินิจฉัยทางการแพทย์ได้, ผู้วิจัยจึงเสนอการแก้ปัญหาด้วยสถาปัตยกรรม Self-Attention Generative Adversarial Network (SAGAN) เพื่อกู้คืนรายละเอียดพื้นผิวของภาพ CT โดยใช้ชุดข้อมูลมาตรฐานสำหรับการฝึกสอน ได้แก่ DIV2K (800 ภาพ) และ RAISE (8,156 ภาพ) และทดสอบประสิทธิภาพบนชุดข้อมูล Benchmark เช่น Set5, Set14 และ BSD100 ร่วมกับชุดข้อมูลภาพถ่ายปอดอักเสบจริง

ในส่วนของกระบวนการเตรียมข้อมูลและประมวลผลภาพ (Image Processing) ซึ่งเป็นจุดเน้นสำคัญ งานวิจัยนี้ได้กำหนดขั้นตอนอย่างซับซ้อน เริ่มจากการใช้โมดูล Down-sampling ที่ประกอบด้วยชุดตัวกรองขนาดต่างกัน (Multi-convolution kernels) ได้แก่ 3×3 , 5×5 และ 7×7 เพื่อสกัดคุณลักษณะของภาพจากมุมมองที่หลากหลาย, จุดที่โดดเด่นคือการเปลี่ยนจากชั้น Batch Normalization แบบดั้งเดิมมาเป็น Instance Normalization (IN) ภายในตัวสร้างภาพ (Generator) เพื่อเพิ่มประสิทธิภาพในการสร้างตัวอย่างภาพเดี่ยวให้มีความแม่นยำสูงขึ้น, นอกจากนี้ยังมีการใช้กลไก Self-attention mechanism ภายในโมดูล Residual เพื่อคำนวณหาความสัมพันธ์ของคุณลักษณะในระดับโลก (Global feature information) โดยการสร้างพื้นที่คุณลักษณะ $f(x)$ สำหรับสกัดพิกเซล และ $g(x)$ สำหรับสกัดข้อมูลระดับโลก ช่วยให้โมเดลสามารถประสานรายละเอียดทั้งระยะใกล้และไกลได้อย่างสม่ำเสมอ, ในส่วนของกระบวนการสร้างแบบจำลอง ผู้วิจัยได้ใช้ Charbonnier content loss แทนที่ L2 loss แบบเดิมเพื่อเพิ่มความทนทานต่อค่าผิดปกติ (Outliers) และเพิ่มความแม่นยำในการสร้างภาพ, ร่วมกับการคำนวณ Perceptual loss จากค่า Feature value ก่อนการเปิดใช้งาน (Before activation) ของ VGG-19 เพื่อรักษาสถาปัตยกรรมพื้นผิวให้ใกล้เคียงกับภาพจริงมากที่สุด และใช้ Wasserstein distance ในการเพิ่มความเสถียรของการฝึกสอนแบบสังกั

ผลการทดลองพบว่า SAGAN ให้ประสิทธิภาพเหนือกว่าอัลกอริทึมระดับรัฐ (State-of-the-art) อื่นๆ เช่น SRCNN, FSRCNN และ LapSRN อย่างมีนัยสำคัญ โดยให้รายละเอียดพื้นผิวที่คมชัดและดูสมจริงมากกว่า

ผู้วิจัยได้ใช้ตัววัดประสิทธิภาพอย่างละเอียด ประกอบด้วยค่า PSNR (Peak Signal-to-Noise Ratio) เพื่อวัดความผิดเพี้ยนของพิกเซล และ SSIM (Structural Similarity Score) เพื่อวัดความคล้ายคลึงของโครงสร้าง ความสว่าง และความต่างสี โดยมีค่า PSNR และ SSIM เป็นตัวชี้วัดหลัก, ซึ่งผลการทดสอบบนชุดข้อมูล Set5 พบว่า SAGAN สามารถทำค่า PSNR ได้สูงถึง 31.94 dB ซึ่งดีกว่า SRGAN ถึง 1.53 dB อย่างไรก็ตาม ช่องว่าง (Gap) ของงานวิจัยนี้คือการที่ SRGAN แบบเดิมยังคงทิ้งร่องรอยความไม่ชัดเจนในรายละเอียดพื้นผิวเฉพาะจุด และมีสิ่งแปลกปลอมทางภาพหลงเหลืออยู่ ประเด็นนี้จึงสอดคล้องกับแนวทางที่สามารถนำมาประยุกต์ใช้ในการต่อยอดงานวิจัย โดยการนำกลไก Self-attention และการคำนวณ Perceptual loss แบบก่อนการเปิดใช้งานมาบูรณาการร่วมกับกลไก CBAM Attention เพื่อเพิ่มความแม่นยำในการระบุตำแหน่งรอยโรคละเอียดขนาดเล็ก ซึ่งจะช่วยลดข้อผิดพลาดจากการแปลงผลภาพที่พร่ามัวและเพิ่มความน่าเชื่อถือในการวินิจฉัยเชิงคลินิกได้ดียิ่งขึ้นตามที่งานวิจัยอ้างอิงได้พิสูจน์ศักยภาพไว้

2.2.3 งานวิจัยด้านเทคนิคการเพิ่มขยายข้อมูลและการจัดการข้อมูลไม่สมดุล

Ayomide Adeyemi Abe และ Mpumelelo Nyathi (2025) (Abe & Nyathi, 2025) นำเสนอแนวทางการแก้ไขปัญหาคือการจำกัดของการวินิจฉัยมะเร็งปอดโดยอัตโนมัติ ซึ่งมักเผชิญกับอุปสรรคสำคัญจากการขาดแคลนข้อมูลทางคลินิกที่มีการติดป้ายกำกับ (Annotated data) และข้อจำกัดของเทคนิคการเพิ่มขยายข้อมูล (Data Augmentation) แบบเดิม เช่น Cutout หรือ MixUp ที่อาจทำให้ข้อมูลสำคัญทาง การแพทย์สูญหายหรือเกิดการบิดเบือนของความหมายในป้ายกำกับภาพ ผู้วิจัยจึงได้พัฒนานวัตกรรมที่เรียกว่า Random Pixel Swap (RPS) ซึ่งเป็นอัลกอริทึมการเพิ่มขยายข้อมูลแบบไร้พารามิเตอร์ที่ใช้วิธีการสลับพิกเซลภายในภาพถ่าย CT Scan ของผู้ป่วยรายเดียวกัน เพื่อสร้างความหลากหลายของข้อมูลในขณะที่ยังคงรักษา ความสมบูรณ์ของลักษณะทางกายวิภาคและพยาธิสภาพไว้ได้อย่างแม่นยำ

ในการศึกษานี้ได้ดำเนินการทดสอบบนชุดข้อมูลมาตรฐาน 2 ชุด ได้แก่ Iraq-Oncology Teaching Hospital/National Center for Cancer Diseases (IQ-OTH/NCCD) ซึ่งมีภาพจำนวน 1,097 ภาพ และชุดข้อมูล Chest CT scan images จำนวน 1,000 ภาพ ในส่วนของกระบวนการเตรียมข้อมูลภาพ (Image Processing) ผู้วิจัยให้ความสำคัญกับการประมวลผลอย่างละเอียดเพื่อเพิ่มประสิทธิภาพของโมเดล เริ่มจากการทำ Normalization ด้วยค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานที่ 0.5 เพื่อสร้างความเสถียรและเร่งการลู่เข้าของแบบจำลอง ที่สำคัญคือขั้นตอนการแยกส่วนภูมิภาคที่น่าสนใจของปอด (Lung ROI Segmentation) ซึ่งประกอบด้วยกระบวนการใช้อัลกอริทึมหาค่าขีดจำกัด (Thresholding) เพื่อสร้างหน้ากากปอด (Lung mask) ตามด้วยกระบวนการ Dilation และการเติมเต็มช่องว่าง (Hole-filling) เพื่อให้ครอบคลุมโครงสร้างปอดทั้งหมดอย่างครบถ้วน ก่อนจะทำการตัดภาพ (Cropping) ให้เหลือเฉพาะบริเวณที่เป็นจุดสนใจเพื่อลดสัญญาณรบกวนจากพื้นหลัง

กระบวนการสร้างแบบจำลองได้ทำการเปรียบเทียบสถาปัตยกรรมสมัยใหม่ทั้งในกลุ่ม Convolutional Neural Networks (CNN) และ Transformer ได้แก่ ResNet-34, MobileNetV3, Vision Transformer (ViT) และ Swin Transformer โดยใช้ AdamW optimizer และ Cross-entropy loss ในการฝึกสอน

ผลการทดลองพบว่าเทคนิค RPS ให้ประสิทธิภาพเหนือกว่าเทคนิคการเพิ่มขยายข้อมูลมาตรฐานอื่นๆ อย่างมีนัยสำคัญ โดยสามารถทำความเข้าใจ (Accuracy) ได้สูงสุดถึง 97.56% บนชุดข้อมูล IQ-OTH/NCCD และ 97.78% บนชุดข้อมูล Chest CT scan โดยมีค่าพื้นที่ใต้กราฟ ROC (AUROC) สูงถึง 98.61% และ 99.46% ตามลำดับ ซึ่งถือเป็นการยกระดับมาตรฐานการวินิจฉัยให้สูงกว่างานวิจัยในอดีต

อย่างไรก็ตาม งานวิจัยนี้ยังมีช่องว่าง (Gap) ที่สำคัญคือการทดสอบส่วนใหญ่อ้างอิงอยู่ที่การประมวลผลภาพแบบ 2 มิติ และยังต้องการการยืนยันผลในสภาพแวดล้อมทางคลินิกที่มีความหลากหลายของอุปกรณ์สแกนมากขึ้นประเด็นเหล่านี้สามารถนำมาประยุกต์ใช้เป็นแนวทางในการต่อยอดงานวิจัย โดยเฉพาะการขยายขอบเขตของเทคนิค RPS ไปสู่การประมวลผลข้อมูลภาพแบบ 3 มิติ (3D Applications) รวมถึงการทดสอบกับภาพถ่ายทางการแพทย์ประเภทอื่นๆ เช่น MRI, Ultrasound หรือ X-ray เพื่อสร้างระบบการวินิจฉัยที่ครอบคลุมและแม่นยำยิ่งขึ้นในอนาคต

Evgin Gocer (2023) (Gocer, 2023) ศึกษาเปรียบเทียบเชิงระบบเกี่ยวกับประสิทธิภาพของเทคนิคการเพิ่มขยายข้อมูลภาพ (Data Augmentation) สำหรับการวินิจฉัยทางการแพทย์ผ่านการเรียนรู้เชิงลึก, โดยมีวัตถุประสงค์หลักเพื่อแก้ปัญหาคาดแคลนข้อมูลภาพทางการแพทย์ (Data scarcity) ซึ่งมักเกิดขึ้นจากข้อจำกัดด้านความเป็นส่วนตัวของผู้ป่วย การขาดแคลนอุปกรณ์ หรือความยากลำบากในการติดป้ายกำกับโดยผู้เชี่ยวชาญ ซึ่งปัญหาเหล่านี้นำไปสู่การเกิดความลำเอียงในชุดข้อมูล (Bias) และการเกิด Overfitting ของโมเดล, ผู้วิจัยจึงได้ดำเนินการทดสอบบนชุดข้อมูลมาตรฐานถึง 4 ประเภท ได้แก่ ข้อมูลภาพ MR สมองจาก BRATS, ภาพ CT ปอดจาก LUNA16, ภาพแมมโมแกรมเต้านมจาก INbreast และภาพถ่ายจอประสาทตาจาก Messidor, ในส่วนของกระบวนการเตรียมข้อมูลภาพ (Image Processing) ซึ่งเป็นจุดเน้นสำคัญ งานวิจัยนี้ได้ออกแบบและทดสอบเทคนิคการเพิ่มขยายข้อมูลถึง 11 รูปแบบ โดยเน้นที่การแปลงทางเรขาคณิต (Geometric Transformations) และการปรับแต่งค่าความเข้มพิกเซล (Intensity Modifications) เช่น การเฉือนภาพ (Shearing) การเลื่อนตำแหน่ง (Translation) การหมุน (Rotation) และการเพิ่มสัญญาณรบกวนทั้งแบบ Gaussian และ Salt-and-pepper, นอกจากนี้สำหรับภาพถ่ายจอประสาทตา ได้มีการใช้เทคนิคการสลับค่าสี (Color shifting) การเพิ่มความคมชัด (Sharpening) และการปรับค่าความต่างสี (Contrast changing) ร่วมด้วย

สำหรับขั้นตอนการสร้างโมเดล ผู้วิจัยได้เลือกใช้สถาปัตยกรรม ResNet101 เป็นสถาปัตยกรรมหลักเนื่องจากความสามารถในการแก้ปัญหา Gradient vanishing และประสิทธิภาพในการเรียนรู้ที่สูง โดยประยุกต์ใช้เทคนิคการเรียนรู้แบบส่งผ่าน (Transfer Learning) จากโมเดลที่ผ่านการฝึกสอนด้วย ImageNet, กระบวนการฝึกสอนถูกกำหนดอย่างละเอียดด้วยตัวปรับปรุงประสิทธิภาพ Adam (Learning rate 0.0003), ฟังก์ชันกระตุ้น ReLU และคำนวณ Cross-entropy loss เพื่อปรับปรุงค่าน้ำหนัก, ผลการทดลองระบุว่าประสิทธิภาพของเทคนิคแต่ละรูปแบบขึ้นอยู่กับลักษณะของภาพเป็นสำคัญ โดยสำหรับภาพ CT ปอด เทคนิคที่ให้ผลลัพธ์ที่ดีที่สุดคือการผสมผสานระหว่างการเลื่อนตำแหน่งและการเฉือนภาพ (Method 9: Translation + Shearing) ซึ่งให้ค่าความแม่นยำ (Accuracy) สูงถึง 85.74%, ผู้วิจัยได้ใช้ตัววัดประสิทธิภาพอย่างครอบคลุม

ถึง 5 ตัว ได้แก่ Accuracy, Sensitivity, Specificity, F1-score และ Matthew Correlation Coefficient (MCC) โดยให้ความสำคัญกับค่า Accuracy เป็นตัวชี้วัดหลักในการเปรียบเทียบศักยภาพของแต่ละเทคนิค อย่างไรก็ตาม งานวิจัยนี้ยังพบช่องว่าง (Gap) ที่สำคัญคือ เทคนิคการแปลงทางเรขาคณิตพื้นฐานแม้จะเพิ่มจำนวนข้อมูลได้แต่ยังขาดความหลากหลายทางพยาธิสภาพของรอยโรค และการใช้ GAN (Generative Adversarial Networks) แม้จะเพิ่มความหลากหลายได้แต่ยังเผชิญกับปัญหาเรื่องความไม่เสถียรในการฝึกสอน (Mode collapse) และต้องการทรัพยากรคำนวณสูง, ประเด็นเหล่านี้จึงเป็นโอกาสในการประยุกต์ใช้เพื่อต่อยอดงานวิจัย โดยเฉพาะการนำเทคนิคการเพิ่มขยายข้อมูลที่เหมาะสมที่สุดจากงานวิจัยนี้ (เช่น การใช้ Translation และ Shearing สำหรับ CT ปอด) มาบูรณาการร่วมกับการปรับปรุงคุณภาพภาพด้วย SRGAN เพื่อเพิ่มความละเอียดของรอยโรค และการใช้กลไก CBAM Attention เพื่อเพิ่มความแม่นยำในการระบุตำแหน่งรอยโรคที่ซับซ้อน ซึ่งจะช่วยอุดช่องว่างเรื่องความเสถียรของภาพสังเคราะห์และความเฉพาะเจาะจงในการตรวจจับตามที่ได้ผู้วิจัยได้เสนอแนะไว้

Noor S. Jozi และ Ghaida A. Al-Suhail นำเสนอกรอบการทำงานของโครงข่ายประสาทเทียมแบบคอนโวลูชัน (CNN) รูปแบบใหม่ที่มีความสามารถในการอธิบายผลได้ (Explainable AI) เพื่อยกระดับการตรวจจับมะเร็งปอดในระยะเริ่มต้น งานวิจัยนี้มุ่งเน้นการแก้ปัญหาความซับซ้อนของข้อมูลภาพถ่ายเอกซเรย์คอมพิวเตอร์ (CT Scans) ที่มีปริมาณมหาศาลซึ่งสร้างภาระงานให้กับรังสีแพทย์ และปัญหาด้านความโปร่งใสของโมเดลการเรียนรู้เชิงลึกที่มักถูกมองว่าเป็นระบบ "กล่องดำ" (Black-box) ซึ่งขาดความน่าเชื่อถือในเชิงคลินิก โดยผู้วิจัยได้ดำเนินการพัฒนาและประเมินผลผ่านชุดข้อมูลมาตรฐาน IQ-OTH/NCCD ซึ่งประกอบด้วยภาพถ่าย CT Scan ที่จำแนกออกเป็น 3 คลาส ได้แก่ เนื้อดี (Benign), เนื้อร้าย (Malignant) และปอดปกติ ในส่วนของกระบวนการเตรียมข้อมูลภาพ (Image Processing) ซึ่งถือเป็นขั้นตอนที่มีความสำคัญอย่างยิ่ง ผู้วิจัยได้กำหนด Pipeline อย่างเป็นระบบเริ่มจากการปรับขนาดภาพ (Resizing) ให้เป็น 220×220 พิกเซลเพื่อความเป็นเอกภาพ ตามด้วยการประยุกต์ใช้ Gaussian filter เพื่อลดสัญญาณรบกวนและเพิ่มความคมชัดของคุณลักษณะที่สำคัญ และการแปลงภาพเป็นสเกลสีเทา (Grayscale conversion) เพื่อลดความซับซ้อนในการวิเคราะห์ นอกจากนี้ยังได้ใช้เทคนิค CLAHE (Contrast Limited Adaptive Histogram Equalization) เพื่อปรับปรุงค่าความต่างสีเฉพาะจุด (Local contrast) และเน้นรายละเอียดของภาพรอยโรคให้ชัดเจนยิ่งขึ้น ก่อนจะปิดท้ายด้วยการทำ Data scaling เพื่อปรับมาตรฐานค่าข้อมูลให้อยู่ในช่วง

กระบวนการสร้างแบบจำลอง LCxNet เริ่มต้นจากการจัดการปัญหาความไม่สมดุลของข้อมูล (Data Imbalance) ด้วยเทคนิค SMOTE (Synthetic Minority Oversampling Technique) เพื่อสร้างข้อมูลสังเคราะห์สำหรับคลาสที่มีน้อย ช่วยให้โมเดลเรียนรู้ขอบเขตการตัดสินใจได้ดียิ่งขึ้น สถาปัตยกรรมของโมเดลประกอบด้วย 4 บล็อกหลักที่ประกอบด้วยชั้น Conv2D, Batch Normalization และ Max Pooling ก่อนจะเข้าสู่ส่วนการจำแนกประเภทที่มีชั้น Dense พร้อมการทำ L2 Regularization และ Dropout เพื่อป้องกันการเกิด Overfitting งานวิจัยนี้ยังมีความโดดเด่นในการเปรียบเทียบพฤติกรรมการเรียนรู้ผ่านตัวปรับปรุงประสิทธิภาพ (Optimizers) ถึง 5 รูปแบบ ได้แก่ SGD, RMSProp, Adam, AdamW และ NAdam ผลการทดลองพบว่าโมเดล LCxNet ให้ประสิทธิภาพสูงสุดเมื่อใช้ตัวปรับปรุง AdamW โดยสามารถทำค่าความ

แม่นยำ (Accuracy) ได้สูงถึง 99.39% ค่าความจำเพาะ (Specificity) 99.45% และค่าพื้นที่ใต้กราฟ (AUC) สมบูรณ์แบบที่ 100% ทั้งนี้ผู้วิจัยได้ใช้ตัววัดประสิทธิภาพที่ครอบคลุมอย่างละเอียด ได้แก่ Accuracy, Weighted Precision, Weighted Sensitivity (Recall), Weighted Specificity, Weighted F1-Score, Cohen's Kappa, Loss และ Misclassification Rate (MR) โดยให้ความสำคัญกับ Accuracy และ Sensitivity เป็นหลัก เนื่องจากในทางคลินิกการลดอัตราผลบปลอม (False Negative) หรือการไม่พลาดเคสที่เป็นเนื้อร้ายถือเป็นภารกิจสำคัญสูงสุด นอกจากนี้ยังมีการใช้เทคนิค Grad-CAM เพื่อสร้าง Heatmap แสดงบริเวณที่โมเดลใช้ในการตัดสินใจ และ t-SNE เพื่อวิเคราะห์การแยกตัวของฟีเจอร์ ซึ่งช่วยสร้างความเชื่อมั่นให้กับบุคลากรทางการแพทย์ได้เป็นอย่างดี

อย่างไรก็ตาม ช่องว่าง (Gap) ที่สำคัญจากงานวิจัยนี้คือการที่ Heatmap จาก Grad-CAM ในบางครั้งยังมีความคลุมเครือหรือกระจายตัวกว้างเกินไปจนขาดความเฉพาะเจาะจงในบางตำแหน่ง และโมเดลยังต้องการการทดสอบกับชุดข้อมูลที่มีความหลากหลายของอุปกรณ์สแกนมากกว่านี้ อีกทั้งผู้วิจัยยังได้เสนอแนะถึงความจำเป็นในการเพิ่มขั้นตอนการแยกส่วนปอด (Lung segmentation) เพื่อกำจัดสัญญาณรบกวนจากพื้นหลัง และการใช้ Generative Adversarial Networks (GANs) เพื่อสร้างภาพจำลองที่มีความสมจริงสูง (High-fidelity) สำหรับคลาสที่ขาดแคลนข้อมูล ประเด็นเหล่านี้จึงเป็นจุดที่งานวิจัยของเรานำมาประยุกต์ใช้และต่อยอด โดยการนำเทคนิค SRGAN มาช่วยยกระดับความละเอียดของภาพ และการใช้กลไก CBAM Attention ร่วมกับโมเดลอย่าง EfficientNet เพื่อให้การเพ่งเล็ง (Attention) ของโมเดลไปยังก้อนเนื้อขนาดเล็กมีความแม่นยำและเฉพาะเจาะจงมากกว่าสถาปัตยกรรม CNN แบบดั้งเดิมที่งานวิจัยอ้างอิงนำเสนอไว้

2.2.4 งานวิจัยด้านการเปรียบเทียบสถาปัตยกรรมและประสิทธิภาพของโมเดล

Nitipon Pongphaw และ Prommin Buaphan (2026) (Pongphaw & Buaphan, 2026) นำเสนอการศึกษาเปรียบเทียบเชิงลึกเกี่ยวกับประสิทธิภาพของสถาปัตยกรรมโครงข่ายประสาทเทียมแบบคอนโวลูชัน (CNN) สมัยใหม่ เพื่อแก้ปัญหาความท้าทายในการจำแนกประเภทเนื้องอกในสมองจากภาพถ่าย MRI ที่ต้องการทั้งความแม่นยำสูงและความสามารถในการอธิบายผลได้ (Interpretability) โดยผู้วิจัยได้ระบุถึงปัญหาสำคัญคือโมเดลการเรียนรู้เชิงลึกส่วนใหญ่มักมีลักษณะเป็น "กล่องดำ" (Black-box) ซึ่งขาดความโปร่งใสในการตัดสินใจ ทำให้บุคลากรทางการแพทย์ขาดความเชื่อมั่นในการนำไปใช้งานจริง ผู้วิจัยจึงเสนอการใช้สถาปัตยกรรม ConvNeXt-Tiny เป็นโมเดลพื้นฐาน (Baseline) เพื่อเปรียบเทียบกับโมเดลมาตรฐานอื่น ๆ ได้แก่ DenseNet169, Xception, MobileNetV3-Large และ ResNet50 ผ่านเทคนิคการเรียนรู้แบบส่งผ่าน (Transfer Learning)

ในการศึกษานี้ได้ดำเนินการพัฒนาโมเดลผ่านชุดข้อมูลมาตรฐาน (Dataset) สองชุดหลัก โดยชุดที่ 1 ประกอบด้วยภาพถ่าย MRI จำนวน 7,023 ภาพ ซึ่งรวบรวมจากแหล่งข้อมูลสาธารณะที่หลากหลาย (เช่น figshare และ Br35H) เพื่อใช้ในการฝึกสอนและประเมินผลเบื้องต้น และชุดที่ 2 ซึ่งเป็นชุดข้อมูลภายนอก (External validation) จำนวน 2,443 ภาพ เพื่อทดสอบความสามารถในการสรุปอ้างอิงของโมเดล (Generalization) ในส่วนของกระบวนการเตรียมข้อมูลภาพ (Image Processing) ซึ่งเป็นหัวใจสำคัญ ได้มี

การกำหนดขั้นตอนอย่างเป็นระบบเริ่มจากการทำ Data Normalization เพื่อปรับสมดุลของค่าข้อมูล ตามด้วยการปรับขนาดภาพ (Resizing) ให้เป็น 224x224 พิกเซล ซึ่งเป็นขนาดมาตรฐานสำหรับโครงข่าย CNN เพื่อให้เกิดความสอดคล้องกันในทุกสถาปัตยกรรม ก่อนจะทำการแปลงข้อมูลเป็น PyTorch tensors และใช้การแบ่งข้อมูลแบบ Stratified split ในสัดส่วน 90:10 เพื่อรักษาความสมดุลของแต่ละคลาสในระหว่างการฝึกสอนและตรวจสอบผล

กระบวนการสร้างแบบจำลองได้กำหนดให้ทุกโมเดลผ่านการฝึกสอนภายใต้เงื่อนไขการทดลองที่เหมือนกันเป็นเวลา 20 รอบ (Epochs) โดยใช้ตัวปรับปรุงประสิทธิภาพ Adam ด้วยอัตราการเรียนรู้ 1×10^{-4} และใช้ฟังก์ชันการสูญเสียแบบ CrossEntropyLoss, ผลการทดลองพบว่า โมเดลส่วนใหญ่สามารถทำประสิทธิภาพได้ในระดับสูงมาก โดย DenseNet169 ให้ค่าความแม่นยำสูงสุดที่ 0.9969 ขณะที่ ConvNeXt-Tiny ซึ่งเป็นโมเดลหลักทำคะแนนได้อย่างโดดเด่นและมีเสถียรภาพด้วยค่าความแม่นยำ (Accuracy) 0.9924 และค่า F1-score 0.9918 โดยมีความได้เปรียบในแง่ของต้นทุนการคำนวณ (Computational cost) ที่ต่ำกว่าโมเดลขนาดใหญ่ ผู้วิจัยได้ใช้ตัววัดประสิทธิภาพที่ครอบคลุมและละเอียดประกอบด้วย Accuracy, Precision, Recall และ F1-score โดยใช้แนวทาง Macro-average เพื่อให้ความสำคัญกับทุกคลาสอย่างเท่าเทียมกัน แม้ในสถานะที่ข้อมูลอาจไม่สมดุล นอกจากนี้ยังมีการวัดประสิทธิภาพเชิงคำนวณผ่าน GPU memory usage และ Training time ต่อรอบ เพื่อวิเคราะห์ความคุ้มค่าในการใช้งาน, โดยผู้วิจัยให้ความสำคัญกับค่า Accuracy และ F1-score เป็นหลัก ในการยืนยันความน่าเชื่อถือของแบบจำลอง, ทั้งยังได้ประยุกต์ใช้เทคนิค Grad-CAM เพื่อสร้าง Heatmap แสดงตำแหน่งที่โมเดลใช้ในการตัดสินใจ ซึ่งผลลัพธ์ยืนยันว่าโมเดลมีการเพ่งเล็งไปยังบริเวณรอยโรคที่เกี่ยวข้องทางคลินิกได้อย่างแม่นยำ

อย่างไรก็ตาม งานวิจัยนี้ยังมีช่องว่าง (Gap) ที่สำคัญคือการพึ่งพาชุดข้อมูลสาธารณะซึ่งอาจไม่สะท้อนถึงความแปรปรวนของข้อมูลในหน่วยงานจริงที่มาจากเครื่องสแกนหรือสถาบันที่ต่างกัน รวมถึงการศึกษายังจำกัดอยู่กับการจำแนกประเภทในระดับภาพถ่าย (Image-level classification) โดยปราศจากการรวมขั้นตอนการแยกส่วนเนื้องอก (Tumor segmentation) และเทคนิค Grad-CAM ที่ใช้ยังให้ผลการระบุตำแหน่งที่ค่อนข้างหยาบ (Coarse localization) ประเด็นเหล่านี้จึงเป็นจุดที่สามารถนำมาประยุกต์ใช้เพื่อต่อยอดงานวิจัย โดยเฉพาะการเพิ่มประสิทธิภาพการระบุตำแหน่งรอยโรคให้มีความเฉพาะเจาะจงมากขึ้นผ่านเทคนิคการอธิบายผลขั้นสูง และการประยุกต์ใช้ร่วมกับการปรับปรุงคุณภาพภาพด้วย SRGAN หรือกลไก Attention ที่ซับซ้อนขึ้นเพื่ออุดช่องว่างในเรื่องความแม่นยำของการตรวจจับพยาธิสภาพขนาดเล็กในภาพถ่ายรังสีวินิจฉัย

Rehan Raza และคณะ (2023) (Raza et al., 2023) นำเสนอนวัตกรรมการจำแนกประเภทมะเร็งปอดที่เน้นความคุ้มค่าของทรัพยากรการคำนวณเพื่อการใช้งานในสถานพยาบาลที่มีข้อจำกัดด้านอุปกรณ์ ปัญหาหลักคือโมเดลการเรียนรู้เชิงลึกแบบเดิมมักต้องการหน่วยประมวลผลประสิทธิภาพสูงและใช้เวลานาน ผู้วิจัยจึงพัฒนาโมเดลโดยใช้สถาปัตยกรรม EfficientNet-B1 เป็นแกนหลักผ่านเทคนิค Compound Scaling ที่ปรับสมดุลระหว่างความลึก ความกว้าง และความละเอียดของภาพ ดำเนินการฝึกสอนบนชุดข้อมูล IQ-OTH/NCCD ในขั้นตอนการเตรียมข้อมูลภาพ (Image Processing) ผู้วิจัยได้ใช้เทคนิคการปรับสเกลภาพให้มีความละเอียดที่เหมาะสมร่วมกับการเพิ่มปริมาณข้อมูลเพื่อลดความไม่สมดุล ผลการทดลองระบุว่าโมเดล

สามารถบรรลุความแม่นยำ (Accuracy) สูงถึง 99.10% และมีค่า Precision สูงถึง 100% โดยใช้พารามิเตอร์เพียง 6.5 ล้านตัว ซึ่งน้อยกว่าโมเดลแบบเดิมอย่างเห็นได้ชัด อย่างไรก็ตาม ช่องว่าง (Gap) ของงานวิจัยนี้คือการมุ่งเน้นที่ความเร็วและประสิทธิภาพของสถาปัตยกรรมเป็นหลัก แต่ยังขาดการใช้กลไกการให้ความสนใจ (Attention Mechanism) ที่เจาะจงไปยังก้อนเนื้อขนาดเล็ก ผู้วิจัยจึงนำจุดแข็งเรื่องความเบาบางของสถาปัตยกรรม EfficientNet มาประยุกต์ใช้ แต่ได้เพิ่มกลไก CBAM Attention และขั้นตอน SRGAN Enhancement เพื่อยกระดับความชัดเจนของขอบรอยโรคให้เหนือกว่างานวิจัยต้นแบบ

Parul Nasra (2024) (Nasra, 2024) นำเสนอการแก้ปัญหาความซับซ้อนของรอยโรคในปอดที่มีลักษณะใกล้เคียงกับเนื้อเยื่อปกติ ซึ่งมักทำให้เกิดความผิดพลาดในการวินิจฉัย ปัญหาสำคัญที่งานวิจัยนี้เผชิญคือปรากฏการณ์เกรเดียนต์ถดถอย (Vanishing Gradient Problem) ในโครงข่ายที่มีความลึกมาก ผู้วิจัยจึงเสนอการใช้สถาปัตยกรรม ResNet50 ที่มีจุดเด่นในเรื่อง Skip Connection เพื่อรักษาประสิทธิภาพการเรียนรู้ในเลเยอร์ที่ลึกขึ้น ในส่วนของกระบวนการเตรียมข้อมูลภาพ (Image Processing) งานวิจัยนี้ให้ความสำคัญกับการทำ Data Augmentation อย่างเข้มข้นเพื่อขยายขอบเขตข้อมูลจากหลักร้อยละภาพไปสู่อีก 15,000 ภาพ ผ่านเทคนิคการหมุนภาพ, การกลับด้าน, การปรับความสว่าง (Brightness) และการปรับค่าความต่างสี (Contrast) เพื่อสร้างความทนทานต่อความแปรปรวนของข้อมูลทางการแพทย์ ผลการทดสอบพบว่าโมเดลสามารถทำประสิทธิภาพได้ดีเยี่ยมโดยใช้ค่าความแม่นยำ (Accuracy) เป็นเกณฑ์ตัดสินหลักที่ร้อยละ 98.1 พร้อมทั้งวิเคราะห์ผ่าน Confusion Matrix เพื่อยืนยันความแม่นยำในการแยกแยะระหว่างเนื้อดีและเนื้อร้าย อย่างไรก็ตาม ช่องว่าง (Gap) ที่พบคือการพึ่งพาเพียงเทคนิคการแปลงภาพเชิงเรขาคณิตแบบดั้งเดิมในการเพิ่มจำนวนข้อมูล ซึ่งอาจยังขาดความหลากหลายทางกายวิภาคที่สมจริงเหมือนการใช้ Generative AI ผู้วิจัยจึงนำแนวคิดโครงข่าย Residual มาใช้เป็นหนึ่งในโมเดลเปรียบเทียบ แต่ได้เสริมความแข็งแกร่งด้วยการใช้ GAN Synthesis เพื่ออุดช่องว่างด้านความหลากหลายของข้อมูลในกลุ่มคลาสที่มีน้อย (Minority class)

Rapeepat Klangbunrueang และคณะ (2025) (Klangbunrueang et al., 2025) นำเสนอกระบวนการจำแนกประเภทมะเร็งปอดเพื่อลดภาระงานของรังสีแพทย์และลดความผิดพลาดจากการแปลผลภาพถ่ายเอกซเรย์คอมพิวเตอร์ ปัญหาหลักที่ผู้วิจัยมุ่งเน้นคือความท้าทายในการคัดกรองโรคในระยะเริ่มต้นที่ต้องการความละเอียดสูงในการวิเคราะห์รอยโรค ผู้วิจัยได้ดำเนินการทดลองผ่านชุดข้อมูลมาตรฐาน IQ-OTH/NCCD โดยมีขั้นตอนการเตรียมข้อมูลภาพ (Image Processing) ที่สำคัญคือการปรับขนาดภาพ (Resizing) ให้เป็นมาตรฐาน 224x224 พิกเซล และการทำ Normalization เพื่อปรับช่วงพิกเซลให้เหมาะสมต่อความเสถียรของโมเดล ร่วมกับการทำ Data Augmentation ผ่านเทคนิคการหมุน (Rotation), การปรับสเกล (Scaling) และการกลับด้านภาพ (Flipping) เพื่อเพิ่มความหลากหลายของข้อมูลและป้องกันการเกิด Overfitting ในส่วนของกระบวนการสร้างโมเดล ผู้วิจัยได้ประยุกต์ใช้เทคนิค Transfer Learning กับสถาปัตยกรรม VGG16 โดยมีการปรับจูนพารามิเตอร์อย่างละเอียดและใช้เทคนิค Dropout ที่อัตรา 0.5 เพื่อเพิ่มความทนทาน ผลการทดลองพบว่าสถาปัตยกรรม VGG16 สามารถสกัดคุณลักษณะเชิงลึกได้อย่างมีประสิทธิภาพ โดยมีค่าความแม่นยำ (Accuracy) เป็นตัววัดหลักซึ่งทำได้สูงถึง 98.67% ร่วมกับการวัดผลผ่าน Precision, Recall และ F1-Score อย่างไรก็ตาม ช่องว่าง (Gap) ที่สำคัญคือข้อจำกัดด้านทรัพยากรการ

คำนวณที่สูงเนื่องจากจำนวนพารามิเตอร์มหาศาลของ VGG16 รวมถึงการใช้เพียงภาพ 2 มิติที่อาจจำกัดการดึงคุณลักษณะในเชิงมิติสัมพันธ์ ผู้วิจัยจึงนำแนวทางการใช้สถาปัตยกรรมที่มีความลึกสูงมาเป็นบรรทัดฐาน แต่ได้ทำการต่อยอดโดยการเลือกใช้สถาปัตยกรรมที่เบาบางกว่าอย่าง EfficientNet บูรณาการร่วมกับกลไก CBAM และเทคนิค GAN เพื่อเพิ่มความทนทานต่อข้อมูลที่ไม่สมดุล

Nurdina Gita Pratiwi และคณะ (2021) (Pratiwi et al., 2021) นำเสนอการศึกษาเชิงเปรียบเทียบเกี่ยวกับผลกระทบของกระบวนการเตรียมข้อมูลภาพ (Pre-processing) ต่อความแม่นยำในการจำแนกโรคโควิด-19 จากภาพถ่าย CT-Scan เพื่อแก้ปัญหาข้อจำกัดของวิธี RT-PCR ที่มีอัตราการเกิดผลลบปลอม (False Negative Rate) สูงและเข้าถึงได้ยากในพื้นที่ห่างไกล. ในการศึกษาได้ใช้ชุดข้อมูลภาพถ่าย CT-Scan 2 มิติจำนวน 2,481 ภาพ จากเว็บไซต์ Kaggle โดยแบ่งสัดส่วนการฝึกสอนและการตรวจสอบที่ 80:20. กระบวนการเตรียมข้อมูลภาพ (Image Processing) ซึ่งเป็นส่วนที่ผู้วิจัยให้ความสำคัญสูงสุด ประกอบด้วยการทดสอบเทคนิค 3 ประเภทหลัก ได้แก่ การปรับขนาดภาพ (Resizing) ทั้งแบบ Deform ที่บังคับสัดส่วนภาพตามขนาดอินพุต และแบบ Non-deform ที่รักษาอัตราส่วนเดิมและเพิ่มพื้นที่ว่างสีเทา, การเพิ่มคุณภาพภาพ (Enhancement) ด้วยเทคนิค Histogram Equalization (HE), Contrast Limited Histogram (CLAHE) และ Exposure Fusion Framework (EFF), และการกำหนดช่วงการทำ Normalization ระหว่าง $[-1, 1]$. ในส่วนของกระบวนการสร้างแบบจำลอง ผู้วิจัยได้ใช้สถาปัตยกรรม VGG-16 ร่วมกับเทคนิค Transfer Learning และ Fine-tuning เพื่อดึงคุณลักษณะเด่นจากภาพถ่ายปอดที่มีความซับซ้อน. ผลการทดลองระบุว่าเทคนิค Pre-processing ส่งผลโดยตรงต่อประสิทธิภาพของแบบจำลอง โดยใช้ค่าความแม่นยำ (Accuracy) เป็นตัววัดประสิทธิภาพหลัก ซึ่งผลลัพธ์ที่เหมาะสมที่สุด (Optimum Result) อยู่ที่ 88.54% จากการใช้เทคนิค Resizing แบบ Deform ร่วมกับการเพิ่มคุณภาพด้วย CLAHE และการทำ Normalization ในช่วง หรือ $[-1, 1]$. อย่างไรก็ตาม งานวิจัยนี้มีช่องว่าง (Gap) ที่สำคัญคือ เทคนิคการปรับขนาดแบบ Deform แม้จะให้ความแม่นยำสูงกว่า แต่การปรับขนาดแบบ Non-deform อาจมีความเหมาะสมในการวินิจฉัยทางคลินิกมากกว่าเนื่องจากไม่ทำลายสัดส่วนและขอบเขตดั้งเดิมของรอยโรคในปอด. นอกจากนี้ ผู้วิจัยยังระบุถึงความจำเป็นในการทำ Segmentation เพื่อแยกส่วนปอดออกมาก่อนการปรับขนาดเพื่อเพิ่มประสิทธิภาพในการจดจำคุณลักษณะของโรคให้ดียิ่งขึ้น. ผู้วิจัยจึงนำแนวทางการเลือกเทคนิค Pre-processing ที่เหมาะสมที่สุดจากงานวิจัยนี้มาประยุกต์ใช้เพื่อเป็นบรรทัดฐานในการออกแบบเฟรมเวิร์กสำหรับการวินิจฉัยรอยโรคในปอดที่ต้องการความแม่นยำและการรักษาความสมบูรณ์ทางกายวิภาคของภาพถ่ายรังสีวินิจฉัย

2.2.5 งานวิจัยด้านการสกัดคุณลักษณะและการแสดงผลเพื่อการอธิบายผล

Muhammad Hassan และคณะ (2025) (Hassan et al., 2025) นำเสนอแนวทางการวินิจฉัยมะเร็งปอดที่ทรงประสิทธิภาพโดยมุ่งเน้นการดึงขีดความสามารถของสถาปัตยกรรม DenseNet-201 มาเป็นแกนหลักในการสกัดคุณลักษณะเชิงลึกแบบอัตโนมัติ (Auto-feature Extraction) ซึ่งโมเดล DenseNet-201 โดดเด่นด้วยโครงสร้างการเชื่อมต่อแบบหนาแน่น (Dense Connections) ที่ช่วยเพิ่มประสิทธิภาพการไหลของค่า

เกรเดียนต์และเกิดการหลอมรวมคุณลักษณะ (Feature Fusion) ที่เข้มข้นในทุกเลเยอร์ ส่งผลให้สามารถสกัดข้อมูลที่มีความซับซ้อนสูงจากภาพถ่าย CT Scan ได้อย่างละเอียดและแม่นยำกว่าโครงข่ายในรูปแบบเดิม

นวัตกรรมที่น่าสนใจของงานวิจัยนี้คือการใช้แนวคิดแบบไฮบริด (Hybrid Approach) โดยการนำฟีเจอร์เชิงลึกที่สกัดได้จาก DenseNet-201 ไปรวมพลัง (Concatenation) กับคุณลักษณะเด่นจากสถาปัตยกรรมอื่นและฟีเจอร์แบบดั้งเดิม (Handcrafted Features) เพื่อให้โมเดลได้รับข้อมูลที่ครอบคลุมทุกมิติ จากนั้นจึงส่งต่อข้อมูลไปยัง Support Vector Machine (SVM) ซึ่งเป็นตัวจำแนกประเภท (Classifier) ที่มีความสามารถสูงในการสร้างเส้นแบ่งการตัดสินใจ (Decision Boundary) ที่ดีเยี่ยมแม้ในพื้นที่ข้อมูลที่มีมิติสูงและมีจำนวนตัวอย่างจำกัด นอกจากนี้ยังมีการเพิ่มประสิทธิภาพในขั้นตอนการเตรียมข้อมูล (Preprocessing) ด้วยการใช้ Gaussian Filter เพื่อลดสัญญาณรบกวน (Noise) และรักษาความคมชัดของขอบรอยโรค ช่วยให้โมเดลสามารถเรียนรู้ลักษณะเด่นที่เป็นประโยชน์ได้ดียิ่งขึ้น ผลการทดลองเชิงประจักษ์ยืนยันว่าการผสมผสานเทคโนโลยี DenseNet-201 เป็นรากฐานสำคัญนี้ ช่วยให้แบบจำลองบรรลุค่าความแม่นยำสูงสุดถึงร้อยละ 99.3 พร้อมทั้งมีค่า Precision และ Recall ที่สมดุลกันอย่างน่าทึ่งในทุกระดับความรุนแรงของโรค ซึ่งสะท้อนให้เห็นว่าแนวทางแบบไฮบริดที่ใช้โมเดล DenseNet-201 ร่วมกับเทคนิคดั้งเดิม สามารถยกระดับมาตรฐานการตรวจคัดกรองมะเร็งปอดให้มีความน่าเชื่อถือในระดับสูงสุดทางคลินิก

Xuhai Xu และคณะ (2025) (Xu et al., n.d.) นำเสนอการศึกษาเชิงลึกเกี่ยวกับผลกระทบของปัญญาประดิษฐ์แบบอธิบายได้ (Explainable AI หรือ XAI) ต่อการตัดสินใจวินิจฉัยโรคผิวหนัง โดยเปรียบเทียบระหว่างกลุ่มประชาชนทั่วไป (Laypeople) และบุคลากรทางการแพทย์ระดับปฐมภูมิ (PCPs) ปัญหาหลักที่งานวิจัยนี้มุ่งแก้ไขคือความคลุมเครือของอัลกอริทึม AI ที่ส่งผลต่อความน่าเชื่อถือและการนำไปใช้จริงในทางการแพทย์ รวมถึงปรากฏการณ์ที่ XAI อาจกลายเป็น "ดาบสองคม" ที่กระตุ้นให้เกิดความลำเอียงจากการพึ่งพาเทคโนโลยีมากเกินไป (Automation bias) หรือการตัดสินใจที่ผิดพลาดเมื่อ AI ให้ข้อมูลที่ไม่ถูกต้อง ผู้วิจัยได้ดำเนินการทดสอบผ่านกระบวนการสร้างแบบจำลองโดยใช้สถาปัตยกรรม ViT-B/32 สำหรับการจำแนกประเภทแบบไบนารี (Melanoma vs. Nevus) และ DenseNet-121 สำหรับการวินิจฉัยแยกโรคแบบเปิด (Differential Diagnosis) โดยใช้ชุดข้อมูลขนาดใหญ่ที่รวบรวมจากแหล่งข้อมูลสาธารณะและส่วนตัวรวม 16 ชุดข้อมูล มีจำนวนภาพรวมทั้งสิ้น 108,585 ภาพ ครอบคลุมความหลากหลายของสีผิวตามเกณฑ์ Fitzpatrick

ในส่วนของการเตรียมข้อมูลภาพ (Image Processing) งานวิจัยนี้ให้ความสำคัญอย่างยิ่งกับการจัดการความหลากหลายของข้อมูล โดยมีการปรับมาตรฐานภาพถ่ายคลินิก (Standardization) และการทำความสะอาดข้อมูลเพื่อขจัดภาพซ้ำ ขั้นตอนที่น่าสนใจเป็นพิเศษคือการระบุค่าสีผิว (Skin Tone Labeling) โดยใช้อัลกอริทึม YCbCr เพื่อแยกส่วนผิวหนังและคำนวณค่า Individual Typology Angle (ITA) เพื่อกำหนดเลเวล Fitzpatrick ให้กับภาพที่ไม่มีข้อมูลดั้งเดิม เพื่อนำไปใช้ในการฝึกสอนโมเดลด้วยอัลกอริทึม CDANN (Conditional Domain Adversarial Neural Network) เพื่อลดความเหลื่อมล้ำในการวินิจฉัยระหว่างสีผิวนอกจากนี้ยังมีการทำ Normalization และการปรับขนาดภาพให้มีคุณภาพสูง (ความละเอียดมากกว่า 400x400 พิกเซล) ก่อนเข้าสู่กระบวนการสกัดคุณลักษณะ กระบวนการสร้างคำอธิบาย (XAI Generation)

ประกอบด้วย 3 เทคนิคหลัก ได้แก่ GradCAM สำหรับระบุพื้นที่สำคัญในระดับพิกเซล CBIR (Content-based Image Retrieval) เพื่อดึงภาพที่มีลักษณะคล้ายคลึงกันด้วยค่า Cosine Distance และการใช้ Multimodal LLM (GPT-4V) เพื่อสร้างคำอธิบายเชิงภาษาที่เป็นเหตุเป็นผล

ผลการทดลองพบว่า AI สามารถยกระดับความแม่นยำในการวินิจฉัยได้ในทุกกลุ่มตัวอย่าง โดยในกลุ่มประชาชนทั่วไปความแม่นยำเพิ่มขึ้นจากร้อยละ 69.7 เป็น 75.8 ส่วนในกลุ่มแพทย์ PCPs ความแม่นยำ Top-1 เพิ่มขึ้นร้อยละ 21.5 งานวิจัยนี้ใช้ตัววัดประสิทธิภาพที่ครอบคลุมและละเอียดมาก โดยตัววัดหลักสำหรับประสิทธิภาพโมเดลคือ Weighted AUROC (ทำได้ 0.930 ใน Study 1 และ 0.772 ใน Study 2) และ Balanced Accuracy เพื่อวัดความเป็นธรรม (Fairness) สำหรับประสิทธิภาพการตัดสินใจของมนุษย์ใช้ค่า Top-1 และ Top-3 Accuracy เป็นเกณฑ์หลัก ร่วมกับการวัดระดับความเชื่อมั่น (Confidence levels) ความสอดคล้องระหว่างความแม่นยำและความมั่นใจ (Accuracy-confidence calibration) และคะแนนคุณภาพคำอธิบาย (Informativeness & Correctness scores) จากผู้เชี่ยวชาญ ผลลัพธ์ที่น่าสนใจคือกลุ่ม Laypeople มีแนวโน้มพึ่งพาคำอธิบายจาก LLM มากเกินไปจนถูกชักจูงได้ง่ายเมื่อ AI ทำนายผิด ในขณะที่แพทย์ผู้เชี่ยวชาญมีความยืดหยุ่น (Resilience) และใช้ XAI เพื่อตรวจสอบสมมติฐานของตนเองมากกว่า

อย่างไรก็ตาม ช่องว่าง (Gap) ที่สำคัญคือผู้วิจัยระบุว่าการออกแบบระบบ XAI แบบ "one-size-fits-all" อาจก่อให้เกิดอันตรายได้หากไม่คำนึงถึงระดับความเชี่ยวชาญของผู้ใช้ และงานวิจัยยังขาดการรวมข้อมูลบริบทของผู้ป่วย (เช่น อายุ เพศ ประวัติสุขภาพ) ซึ่งมีความสำคัญต่อการวินิจฉัยจริง ประเด็นเหล่านี้จึงเป็นจุดที่สามารถนำมาประยุกต์ใช้เพื่อต่อยอดงานวิจัย โดยเฉพาะการนำแนวคิด "Human-First Workflow" (ให้มนุษย์ตัดสินใจก่อน AI) มาใช้เพื่อลดปัญหา Anchoring bias และการพัฒนาโมเดลให้มีความแม่นยำเฉพาะเจาะจงมากขึ้นผ่านเทคนิค SRGAN เพื่อเพิ่มรายละเอียดในภาพ CT Scan และการใช้ CBAM Attention เพื่อระบุตำแหน่งพยาธิสภาพที่ซับซ้อน ซึ่งจะช่วยอุดช่องว่างในเรื่องความเฉพาะเจาะจงของการระบุตำแหน่งรอยโรค (Fine-grained localization) ที่เทคนิค GradCAM แบบดั้งเดิมในงานวิจัยนี้ยังทำได้ค่อนข้างจำกัด

ตารางที่ 2.1 ตารางเปรียบเทียบงานวิจัยที่ใช้ชุดข้อมูล IQ-OTH/NCCD

งานวิจัย (ปี)	โมเดลที่ใช้ (Model)	การเตรียมข้อมูลภาพ (Preprocessing)	ค่าความแม่นยำ (Accuracy)	ตัววัดประสิทธิภาพอื่นๆ (Metrics)
Jozi และ Al-Suhail (2025)	LCxNet (Custom CNN)	Resize (220x220), Gaussian filter, Grayscale, CLAHE, SMOTE	99.39%	Sensitivity 99.39%, Specificity 99.46%, Kappa 98.94%
Musthafa และคณะ (2024)	CNN (Double-Layered)	SMOTE, Hyperparameter Optimization	99.64%	-
Raza และคณะ (2023)	Lung-EffNet (EfficientNet-B1)	Compound Scaling, Data Augmentation	99.10%	-
Abumohsen และคณะ (2025),	DenseNet201	Regularization, Focal Loss, Data Augmentation	98.95%	Precision 99%, Recall 99%, F1-score 99%
Jozi และ Al-Suhail (2024)	ResNet50V2	Resize (224x224), Bilateral Filtering, Grayscale, CLAHE, SMOTE	98.18%	Sensitivity 95.46%, Specificity 99.09%
Klangbunrueang และคณะ (2025)	VGG16	Resize (224x224), Normalization, Augmentation (Rotation, Flip)	98.18%	-
Abdollahi (2023),	LeNet	Data Augmentation	97.88%	Sensitivity 93.14%, Specificity 95.91%
Yan และ Razmjoo (2024)	CNN + Snake Optimizer	Optimization Algorithm Integration	96.58%	F1-score 91.53%, Precision 84.16%
AL-Huseiny และ Sajit (2021)	GoogLeNet	Transfer Learning	94.38%	Sensitivity 95.08%, Specificity 93.7%
Al-Yasriy และคณะ (2020)	AlexNet	-	93.54%	Sensitivity 95.71%, Specificity 95%
งานวิจัยนี้ (Ours)	Densenet, MobileNetV3, Resnet50V2	SRGAN / GAN Synthesis / Geometric	(เป้าหมายสูง กว่า 98%)	ครบทุกตัว (ACC,P, R, F1, AUC-ROC,)

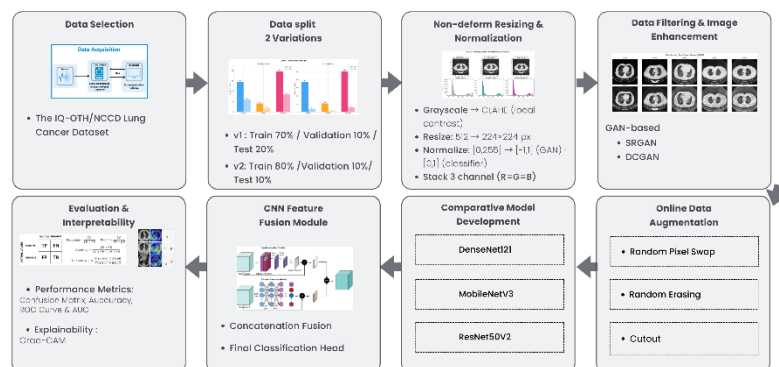
บทที่ 3

ขั้นตอนการดำเนินโครงการ

3.1 ภาพรวมกรอบการวิจัย

โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Networks: CNNs) ได้กลายเป็นรากฐานสำคัญและมาตรฐานที่เป็นสากลสำหรับการวิเคราะห์ภาพถ่ายทางการแพทย์อัตโนมัติในปัจจุบัน โดยเฉพาะในการกักการจำแนกประเภทมะเร็งปอดจากภาพถ่ายรังสีคอมพิวเตอร์ที่ต้องการความแม่นยำสูง เฟรมเวิร์กการดำเนินงานวิจัยนี้เริ่มต้นจากการคัดเลือกชุดข้อมูลมาตรฐาน IQ-OTH/NCCD ซึ่งครอบคลุมทั้งกลุ่มปกติ กลุ่มรอยโรคชนิดไม่ร้ายแรง และกลุ่มรอยโรคชนิดร้ายแรง โดยข้อมูลจะถูกจัดสรรผ่านระเบียบวิธีวิจัย 2 รูปแบบมาตรฐาน (Abe & Nyathi, 2025; Jozi & Al-Suhail, 2024) เพื่อทดสอบความทนทานของแบบจำลองภายใต้สัดส่วนการแบ่งข้อมูลที่แตกต่างกัน ก่อนจะเข้าสู่กระบวนการเตรียมข้อมูลภาพ (Image Preprocessing) ที่มีความละเอียดอ่อน ทั้งการแปลงภาพเป็นเกรayscale การใช้เทคนิค CLAHE เพื่อเพิ่มความคมชัดของลักษณะทางพยาธิสภาพ และการปรับขนาดภาพให้คงที่ที่ 224×224 พิกเซลเพื่อสร้างมาตรฐานเดียวกันในการประมวลผล นวัตกรรมหลักของเฟรมเวิร์กนี้อยู่ที่การบูรณาการระบบปัญญาประดิษฐ์เชิงสร้างสรรค์จำนวน 2 รูปแบบ ได้แก่ SRGAN เพื่อกู้คืนรายละเอียดพื้นผิวเนื้อเยื่อที่ซับซ้อน และ DCGAN เพื่อสังเคราะห์ภาพถ่ายชุดใหม่สำหรับปรับสมดุลปริมาณข้อมูลในคลาสส่วนน้อยให้เท่าเทียมกับกลุ่มประชากรส่วนใหญ่ ข้อมูลที่ผ่านการยกระดับคุณภาพแล้วจะถูกนำไปเสริมความแข็งแกร่งด้วยกลยุทธ์การเพิ่มขยายข้อมูลแบบออนไลน์ (Online Data Augmentation) ผ่านเทคนิค Random Pixel Swap (RPS) ร่วมกับเทคนิค Random Erasing และ Cutout (Abe & Nyathi, 2025) เพื่อจำลองความหลากหลายของสัญญาณรบกวนโดยไม่เสียลักษณะทางกายวิภาคที่สำคัญ ซึ่งข้อมูลทั้งหมดจะถูกนำไปฝึกฝนและเปรียบเทียบสมรรถนะผ่านสถาปัตยกรรม CNN ขั้นนำอันประกอบด้วย DenseNet121, MobileNetV3 และ ResNet50V2 ท้ายที่สุดระบบจะทำการประเมินผลผ่านดัชนีชี้วัดเชิงปริมาณที่ครอบคลุมควบคู่ไปกับการใช้เทคนิค Grad-CAM เพื่อสร้างแผนภูมิความร้อนระบุตำแหน่งรอยโรคปอดที่ตัวแบบใช้ตัดสินใจ ซึ่งช่วยสร้างหลักฐานเชิงการตีความที่โปร่งใสและสอดคล้องตามหลักการทางการแพทย์อย่างแท้จริง

Framework

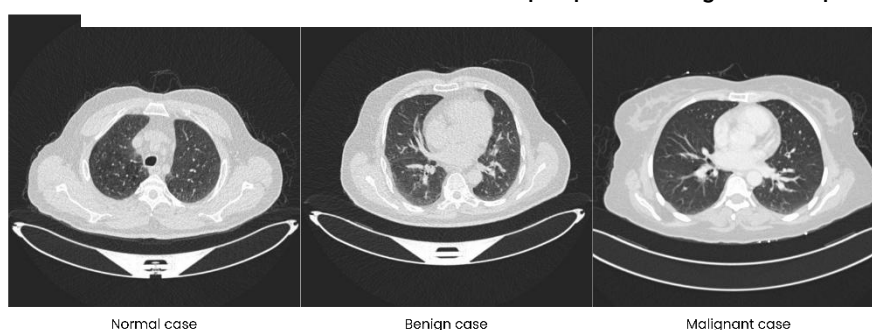


ดั่งภาพประกอบที่ 3.3.1

3.2 ชุดข้อมูล

การวินิจฉัยการวินิจฉัยมะเร็งปอดจากภาพถ่ายเอกซเรย์คอมพิวเตอร์ผ่านการเรียนรู้เชิงลึก ใช้ชุดข้อมูลภาพถ่ายเอกซเรย์คอมพิวเตอร์ (CT Scan) จากแหล่งข้อมูลมาตรฐานเปิดเผยจากเว็บไซต์ Kaggle (<https://www.kaggle.com/datasets/hamdallak/the-iqothnccd-lung-cancer-dataset/data>) ชื่อชุดข้อมูล The IQ-OTH/NCCD Lung Cancer Dataset รวบรวมจากโรงพยาบาลเฉพาะทาง 2 แห่งในประเทศอิรัก ได้แก่ Iraq-Oncology Teaching Hospital และ National Center for Cancer Diseases (NCCD) โดยเก็บข้อมูลในช่วงฤดูใบไม้ร่วงปี 2019 และได้รับการรับรองจากคณะกรรมการพิจารณาจริยธรรมของศูนย์

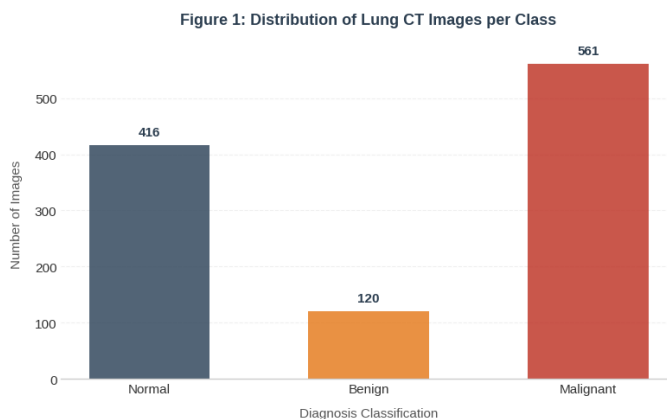
ภาพประกอบที่ 3.1 Framework of preprocessing technique



ภาพประกอบที่ 3.2 ภาพwct-scans ตัวอย่างแต่ละคลาส

การแพทย์ที่เกี่ยวข้อง ดั่งภาพประกอบที่2

ชุดข้อมูลมีทั้งหมด 1097 ภาพ ข้อมูลแบ่งออกเป็น 3 คลาส ได้แก่ กลุ่มปกติ (Normal) มีจำนวน 416 ภาพ คือไม่พบก้อนเนื้อผิดปกติ, กลุ่มเนื้องอกไม่อันตราย (Benign) มีจำนวน 120 ภาพ คือก้อนเนื้อขนาดเล็ก, ขอบชัด, ไม่แพร่กระจาย และกลุ่มเนื้องอกร้าย/มะเร็ง (Malignant) มีจำนวน 561 ภาพ คือก้อนเนื้อขอบไม่ชัด, มีลักษณะรุกราน โดยที่ชุดข้อมูลทั้งหมดยังไม่ผ่านการเตรียมข้อมูล ดังภาพที่ 3



ภาพประกอบที่ 3.3 กราฟแสดงจำนวนข้อมูลทั้ง 3 คลาส

3.3 การเตรียมและแบ่งชุดข้อมูล

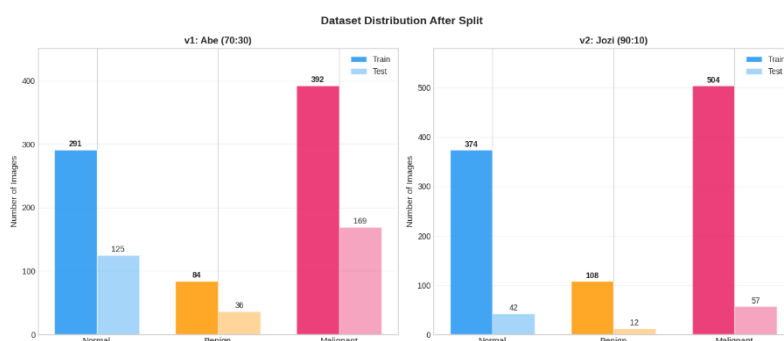
ในกระบวนการพัฒนาตัวแบบการเรียนรู้เชิงลึก การจัดสรรข้อมูลที่มีประสิทธิภาพถือเป็นขั้นตอนสำคัญในการป้องกันสถานะการเรียนรู้ที่จำเพาะเจาะจงเกินไป (Overfitting) และเพื่อให้มั่นใจว่าตัวแบบสามารถสรุปอ้างอิงองค์ความรู้ไปยังข้อมูลที่ไม่เคยพบเห็นมาก่อนได้อย่างแม่นยำ (Jozi & Al-Suhail, 2025; Tawfik et al., 2024) ผู้วิจัยจึงได้กำหนดระเบียบวิธีวิจัยในการจัดสรรชุดข้อมูลอย่างเป็นระบบ โดยครอบคลุมทั้งการรักษาความสมดุลของสัดส่วนคลาสและการทดสอบความคงเส้นคงวาของผลลัพธ์ผ่านเกณฑ์มาตรฐานสากล

3.3.1 โปรโตคอลการแบ่งข้อมูล (Split Protocol)

ในการทดลองนี้ได้เลือกใช้เทคนิคการแบ่งข้อมูลแบบสุ่มโดยรักษาสัดส่วนของคลาส (Stratified Random Split) ซึ่งเป็นวิธีการที่การันตีว่าสัดส่วนของภาพในแต่ละคลาส อันได้แก่ คลาสปกติ (Normal) คลาสถุงน้ำหรือก้อนเนื้อธรรมดา (Benign) และคลาสน้ำหรือมะเร็งร้ายแรง (Malignant) ในชุดข้อมูลฝึกฝน (Training Set) และชุดข้อมูลทดสอบ (Test Set) จะถูกจัดสรรในอัตราส่วนที่เท่ากับสัดส่วนจริงของชุดข้อมูลตั้งต้นทั้งหมด เพื่อป้องกันปัญหาการกระจายตัวของคลาสที่คลาดเคลื่อน ซึ่งอาจส่งผลให้โมเดลเกิดการเรียนรู้เอนเอียงไปยังคลาสที่มีจำนวนภาพมากกว่า นอกจากนี้เพื่อให้กระบวนการแบ่งข้อมูลมีความเสถียรและเพื่อให้การทดลองสามารถทำซ้ำได้ (Reproducibility) การสุ่มแบ่งข้อมูลทั้งหมดจึงถูกกำหนดค่าพารามิเตอร์ควบคุมการสุ่มคงที่ (random_state) หรือค่า Seed ไว้ที่ 42 สำหรับทุกชุดการทดลอง (Abe & Nyathi, 2025; Tabibian et al., 2025)

จุดเน้นสำคัญของเฟรมเวิร์กนี้คือการประยุกต์ใช้ โปรโตคอลการแบ่งข้อมูล 2 รูปแบบหลัก เพื่อตรวจสอบความเสถียร (Stability Validation) ของเทคนิค GAN และการเพิ่มขยายข้อมูลข้ามสัดส่วนปริมาณข้อมูลที่แตกต่างกัน โปรโตคอลแรก (v1_abe_7030) โดยแบ่งข้อมูลในสัดส่วน 70:30 สำหรับการฝึกสอนและทดสอบ ซึ่งมุ่งเน้นการวัดสมรรถนะภายใต้ปริมาณการเรียนรู้ระดับมาตรฐาน (Abe & Nyathi, 2025) ในขณะที่โปรโตคอลที่สอง (v2_jozi_9010) ในสัดส่วน 90:10 เพื่อทดสอบว่าหากเพิ่มปริมาณข้อมูลฝึกสอนให้สูงขึ้น

อย่างมีนัยสำคัญ ระบบจะยังคงรักษาความแม่นยำและระดับการพัฒนาสมรรถนะที่ได้จาก GAN ไว้ได้หรือไม่(Jozi & Al-Suhail, 2024) การวิเคราะห์เปรียบเทียบระหว่างสองโปรโตคอลนี้จะช่วยให้ผู้วิจัยทราบว่าสัดส่วนการแบ่งข้อมูลมีผลกระทบเชิงระบบต่อประสิทธิภาพโดยรวมของแบบจำลองหรือไม่ หรือเทคนิคที่นำเสนอสามารถมอบผลลัพธ์ที่เป็นอิสระต่อการจัดสรรข้อมูลได้อย่างแท้จริง โดยรายละเอียดจำนวนภาพที่ใช้ในแต่ละโปรโตคอลได้แสดงไว้ ดังภาพที่ 4



ภาพประกอบที่ 3.4 จำนวนภาพในชุดข้อมูลฝึกสอนและทดสอบภายใต้ 2 โปรโตคอลมาตรฐาน

3.3.2 การล็อกชุดข้อมูลทดสอบ (Locking the Test Set)

เพื่อให้การประเมินสมรรถนะของตัวแบบเป็นไปอย่างเที่ยงธรรมและสะท้อนถึงขีดความสามารถที่แท้จริง คณะผู้วิจัยได้กำหนดให้ ชุดข้อมูลทดสอบ (Test Set) ถูกแยกออกมาเป็นลำดับแรกสุดก่อนเริ่มกระบวนการประมวลผลใด ๆ โดยข้อมูลในส่วนนี้จะถูก "ล็อก" (Locked) และจัดเก็บแยกไว้อย่างอิสระโดยไม่ถูกนำมาผ่านกระบวนการยกระดับคุณภาพด้วย GAN (SRGAN/DCGAN) หรือเทคนิคการเพิ่มขยายข้อมูล (Data Augmentation) ใด ๆ ทั้งสิ้น ตลอดระยะเวลาการดำเนินงานวิจัย เหตุผลสำคัญของการแยกข้อมูลทดสอบไว้ในรูปแบบดั้งเดิม (Original/Raw images) คือเพื่อ จำลองสถานการณ์จริงทางคลินิก (Real-world clinical scenario) ที่ตัวแบบปัญญาประดิษฐ์จะต้องพบกับภาพถ่าย CT Scan ต้นฉบับจากผู้ป่วย ซึ่งเป็นภาพที่มีความแปรปรวนตามธรรมชาติและไม่ได้ผ่านการปรับแต่งฟิสิกส์โดยระบบคอมพิวเตอร์กระบวนการนี้ช่วยให้มั่นใจได้ว่าความแม่นยำที่วัดได้นั้นมาจากความสามารถในการจดจำลักษณะทางพยาธิวิทยาที่แท้จริง ไม่ใช่การจดจำรูปแบบฟิสิกส์ที่ถูกสร้างขึ้นใหม่ (Artifacts) ส่งผลให้การประเมินความสามารถในการสรุปอ้างอิงองค์ความรู้ (Generalizability) ของตัวแบบมีความน่าเชื่อถือสูงสุด นอกจากนี้การแยกชุดข้อมูลทดสอบอย่างเด็ดขาดตั้งแต่เริ่มต้นยังเป็นการป้องกัน ปัญหาการรั่วไหลของข้อมูลทดสอบ (Test set leakage) ซึ่งถือเป็นอุปสรรคสำคัญที่บั่นทอนความน่าเชื่อถือของงานวิจัย สอดคล้องกับข้อเสนอแนะที่ระบุว่าการรั่วไหลของข้อมูลระหว่างชุดฝึกสอนและชุดทดสอบ เป็นประเด็นวิกฤตในงานวิจัยด้านการวิเคราะห์ภาพถ่ายทางการแพทย์ (Medical imaging research) เนื่องจากจะทำให้ตัวแบบแสดงค่าความแม่นยำที่สูงเกินจริง (Overoptimistic results) และไม่สามารถนำไปใช้งานได้จริงในสถานการณ์ Clinical deployment(Bhatt et al., 2024)

3.4 การประมวลผลภาพเบื้องต้น

ในขั้นตอนการเตรียมข้อมูลภาพถ่ายเอกซเรย์คอมพิวเตอร์ (Computed Tomography: CT Scan) ปอด ได้มีการออกแบบกระบวนการประมวลผลล่วงหน้าเพื่อปรับปรุงคุณภาพของภาพให้มีความเหมาะสมสูงสุด ก่อนนำเข้าสู่สถาปัตยกรรมโครงข่ายประสาทเทียมเพื่อการขยายความละเอียดภาพ (Super-Resolution Generative Adversarial Network: SRGAN) เนื่องจากภาพถ่ายทางแพทยศาสตร์ต้นฉบับขนาด 512×512 พิกเซล มีระดับความต่างสี (Contrast) ในส่วนเนื้อปอดค่อนข้างต่ำและมีมิติช่องสัญญาณแบบระดับสีเทา (Grayscale) กระบวนการประมวลผลเบื้องต้นจึงประกอบด้วยขั้นตอนตามลำดับดังนี้

3.4.1 การปรับขนาดภาพ (Image Resizing)

เมื่อภาพต้นฉบับขนาด 512×512 เนื่องจากภาพถ่ายเอกซเรย์คอมพิวเตอร์ปอดต้นฉบับมีขนาดมิติที่หลากหลายและมีความละเอียดสูง การส่งภาพเข้าสู่แบบจำลองการเรียนรู้เชิงลึกโดยตรงจึงจำเป็นต้องมีการปรับมิติของภาพให้เป็นมาตรฐานเดียวกัน เพื่อให้สอดคล้องกับโครงสร้างและข้อกำหนดของสถาปัตยกรรมโมเดลแต่ละประเภท ในงานวิจัยนี้ได้กำหนดขนาดมิติเป้าหมายออกเป็น 2 รูปแบบหลัก ได้แก่

1. ภาพความละเอียดสูง (High Resolution: HR): ทำการย่อขนาดภาพจาก 512×512 เป็น 224×224 พิกเซล โดยใช้เทคนิคการสอดแทรกพิกเซลแบบ Bicubic Interpolation (INTER_CUBIC) เพื่อรักษาความคมชัดของขอบรอยโรค (Edges) และรายละเอียดเนื้อเยื่อปอดให้ใกล้เคียงต้นฉบับมากที่สุด โดยภาพส่วนนี้จะทำหน้าที่เป็นภาพเฉลยจริง (Ground Truth) สำหรับให้ Discriminator และ VGG Loss ใช้เปรียบเทียบวัดผล และเป็นการปรับมิติภาพให้ตรงตามมาตรฐานของโครงสร้างแบบจำลองการเรียนรู้เชิงลึกระดับฐาน (Pre-trained Models) ที่ได้รับการฝึกฝนมาก่อนหน้าบนชุดข้อมูล ImageNet ได้แก่ DenseNet121, MobileNetV3-Large และ ResNet50V2

2. ภาพความละเอียดต่ำ (Low Resolution: LR): ทำการย่อขนาดภาพจาก 512×512 เป็น 112×112 พิกเซล ด้วยเทคนิค Pixel Area Relation (INTER_AREA) เพื่อป้องกันการเกิดรอยต่าง (Moire patterns) หรือสิ่งแปลกปลอมในภาพย่อขนาดเล็ก โดยภาพส่วนนี้จะถูกใช้เป็นข้อมูลนำเข้า (Input) ให้แก่ Generator เพื่อความเหมาะสมในการประมวลผลของเครือข่ายสังเคราะห์ภาพ (Generative Adversarial Networks) แบบ SRGAN

3.4.2 การปรับความต่างสีแบบปรับเหมาะจำกัดเกณฑ์ (CLAHE)

ภาพถ่ายทั้งในชุดความละเอียด HR (224×224) และ LR (112×112) จะถูกนำมาปรับปรุงความคมชัดเฉพาะถิ่นด้วยเทคนิค Contrast Limited Adaptive Histogram Equalization (CLAHE) โดยกำหนดค่าจำกัดการตัดกราฟฮิสโตแกรม (Clip Limit) เท่ากับ 2.0 และขนาดพื้นที่ย่อย (Tile Grid Size) เท่ากับ 8×8 จุดภาพ เทคนิคดังกล่าวนำมาใช้เพื่อเพิ่มความเด่นชัดของรอยโรคขนาดเล็กและความแตกต่างของเนื้อเยื่อภายในปอด โดยไม่ทำให้เกิดสัญญาณรบกวน (Noise) เพิ่มขึ้นในบริเวณภาพ

3.4.3 การปรับบรรทัดฐานค่าพารามิเตอร์พิกเซล (Normalization) และการแปลงช่องสัญญาณ

ค่าความเข้มสีของพิกเซลในภาพถ่ายทางการแพทย์ดั้งเดิมจะอยู่ในช่วงสเกลความสว่าง $[0, 255]$ การนำค่าพิกเซลนี้เข้าสู่แบบจำลองโดยตรงอาจส่งผลให้ปัญหาสัญญาณความชันหาย (Exploding/Vanishing Gradients) ในระหว่างกระบวนการป้อนกลับวนซ้ำ (Backpropagation) งานวิจัยนี้จึงได้แบ่งกระบวนการปรับบรรทัดฐาน (Normalization Pipeline) ออกเป็น 2 รูปแบบตามสถาปัตยกรรมโมเดลที่รองรับ ดังนี้

1. การปรับบรรทัดฐานสำหรับสถาปัตยกรรม GAN: ปรับระดับค่าความเข้มพิกเซลให้อยู่ในช่วงช่วงความสว่าง $[-1, 1]$ ผ่านสูตรการคำนวณ $\text{Pixel}_{\text{normalized}} = (\text{Pixel}/127.5) - 1.0$ ซึ่งการกำหนดช่วงข้อมูลดังกล่าวมีความจำเป็นอย่างยิ่ง เนื่องจากฟังก์ชันกระตุ้น (Activation Function) ในชั้นส่งออก (Output Layer) ของแบบจำลอง Generator ใช้ฟังก์ชัน Hyperbolic Tangent (**tanh**) ซึ่งให้ช่วงผลลัพธ์ระหว่าง -1 ถึง 1 การแปลงข้อมูลให้อยู่ในช่วงเดียวกันจะช่วยให้เครือข่ายประสาทสังเคราะห์ภาพเรียนรู้และรับฟังค่าความผิดพลาดได้รวดเร็วยิ่งขึ้น
2. การปรับบรรทัดฐานสำหรับสถาปัตยกรรม Classifier: ปรับระดับค่าความเข้มพิกเซลให้อยู่ในช่วงช่วงความสว่าง $[0, 1]$ ผ่านสูตรการคำนวณ $\text{Pixel}_{\text{normalized}} = \text{Pixel}/255.0$ ซึ่งสอดคล้องกับสถาปัตยกรรมจำแนกประเภทภาพที่ใช้ฟังก์ชันกระตุ้นประเภท Sigmoid หรือ Softmax ในชั้นทำนายผลลัพธ์สุดท้าย

นอกจากนี้ เนื่องจากภาพถ่าย CT Scan ดั้งเดิมเป็นภาพระดับสีเทาช่องสัญญาณเดียว (Grayscale Image: 1 Channel) แต่สถาปัตยกรรมแบบจำลองโครงข่ายประสาทคอนโวลูชันระดับฐาน (Pre-trained Backbones) ทั้งหมดถูกออกแบบมาให้รับอินพุตภาพสีแบบ 3 ช่องสัญญาณ (RGB: 3 Channels) งานวิจัยนี้จึงได้ดำเนินการซ้อนทับช่องสัญญาณ (Channel Stacking) โดยการคัดลอกข้อมูลพิกเซลจากช่องสัญญาณเดียวขยายออกเป็น 3 ช่องสัญญาณที่มีค่าพิกเซลเดียวกันทุกประการ (Grayscale $\rightarrow$ RGB) เพื่อให้แบบจำลองสามารถดึงฟีเจอร์จากน้ำหนักพิกเซลเริ่มต้น (Pre-trained Weights) มาประยุกต์ใช้ได้อย่างสมบูรณ์

ตารางที่ 3.1 กระบวนการเตรียมและประมวลผลภาพเบื้องต้น (Preprocessing Pipeline)

ขั้นตอนการประมวลผล	GAN Pipeline	Classifier Pipeline	วัตถุประสงค์และเหตุผลทางเทคนิค
1. CLAHE Enhancement	✓	✓	เพิ่ม Local Contrast ของภาพ CT Scan โดยไม่ทำให้บริเวณที่สว่างอยู่แล้วเกิด Bright Overexposure
2. Image Resizing	128 × 128px	224 × 224px	ปรับมิติภาพให้สอดคล้องตามความต้องการของสถาปัตยกรรมโมเดลแต่ละประเภท

3. Interpolation Method	INTER_CUBIC / INTER_AREA	INTER_CUBIC / INTER_AREA	ใช้ INTER_CUBIC เมื่อขยายเพื่อรักษาความคมชัดขอบรอยโรค และใช้ INTER_AREA เมื่อย่อเพื่อลด Aliasing
4. Pixel Normalization	$[-1,1]$	$[0,1]$	ปรับสเกลข้อมูลให้ตรงกับ Activation Function (tanh ใน GANs และ Sigmoid/Softmax ใน Classifier)
5. Grayscale to RGB Conversion	✓(Stack 3 Channels)	✓(Stack 3 Channels)	แปลงภาพจาก 1 ช่องสัญญาณเป็น 3 ช่องสัญญาณ เพื่อให้สอดคล้องกับ Pre-trained Models ที่ฝึกบน ImageNet

3.5 การเพิ่มข้อมูลด้วย GAN (GAN-based Data Augmentation)

ปัญหาการขาดแคลนข้อมูลภาพถ่ายทางการแพทย์ ตลอดจนปัญหาความไม่สมดุลของจำนวนภาพในแต่ละคลาส (Class Imbalance) นับเป็นอุปสรรคสำคัญที่ส่งผลให้โมเดลการเรียนรู้เชิงลึกเกิดภาวะ Overfitting และขาดความสามารถในการทำนายผลกับชุดข้อมูลที่ไม่เคยเห็นมาก่อนได้อย่างแม่นยำ งานวิจัยนี้จึงประยุกต์ใช้โครงข่ายประสาทสังเคราะห์ภาพ (Generative Adversarial Networks: GANs) เข้ามาเป็นเครื่องมือหลักในการสังเคราะห์ภาพถ่ายเอกซเรย์คอมพิวเตอร์ปอดเสมือนจริง เพื่อเพิ่มปริมาณข้อมูลในชุดข้อมูลฝึกฝน (Training Set) อย่างมีประสิทธิภาพ

3.5.1 กลยุทธ์กฎทวิคูณ (Double Rule Strategy)

เพื่อควบคุมปริมาณและสัดส่วนการสังเคราะห์ภาพให้อยู่ในเกณฑ์มาตรฐานที่มีความแม่นยำและไม่ทำให้การกระจายตัวของข้อมูลดั้งเดิมผิดเพี้ยนไป งานวิจัยนี้จึงได้นำ กลยุทธ์กฎทวิคูณ (Double Rule Strategy หรือ 2x Rule) ตามแนวทางของ Abe และคณะ มาประยุกต์ใช้ในการเพิ่มจำนวนภาพในชุดข้อมูลฝึกฝน หลักการสำคัญของกลยุทธ์ Double Rule คือการกำหนดให้อัตราส่วนการเพิ่มข้อมูลภาพสังเคราะห์จาก GAN มีค่าเท่ากับ 1:1 เมื่อเทียบกับภาพถ่ายจริงต้นฉบับ นั่นคือ สำหรับทุกๆ ภาพถ่ายจริงดั้งเดิมในชุดข้อมูลฝึกฝน (1 Original Image) จะถูกจับคู่และเพิ่มขยายด้วยภาพถ่ายสังเคราะห์ที่สร้างขึ้นจากโมเดล GAN จำนวน 1 ภาพ (1 GAN-generated Image) ส่งผลให้ปริมาณรวมของชุดข้อมูลฝึกฝนหลังผ่านกลยุทธ์นี้เพิ่มขึ้นเป็น 2 เท่าตัว (2x) จากจำนวนภาพตั้งต้นพอดี

ข้อดีของกลยุทธ์ดังกล่าวนอกจากจะช่วยแก้ปัญหา Class Imbalance ในคลาสที่มีจำนวนภาพน้อย (เช่น คลาส Benign) แล้ว ยังช่วยเพิ่มความหลากหลายของลักษณะคุณประโยชน์ (Feature Diversity) ของรอยโรคในปอด โดยที่ยังคงรักษาสัดส่วนและโครงสร้างทางสถิติของชุดข้อมูลดั้งเดิมไว้อย่างเหมาะสม ทั้งนี้ จำนวนภาพในชุดข้อมูลฝึกฝนก่อนและหลังการประยุกต์ใช้กลยุทธ์ Double Rule แยกตามโปรโตคอลการแบ่งข้อมูลทั้ง 2 รูปแบบ แสดงดังตารางที่ 3.2

ตารางที่ 3.2 จำนวนภาพถ่ายในชุดข้อมูลฝึกฝนก่อนและหลังการใช้นวัตกรรม Double Rule (2x)

Split Protocol	จำนวนภาพดั้งเดิม (Original Train)	จำนวนภาพรวมหลังใช้นวัตกรรม Double Rule (2x)	จำนวนภาพสังเคราะห์จาก GAN ที่เพิ่มขึ้น
v1_abe_7030	767	1,534	+767

v2_jozi_9010	987	1,974	+987
--------------	-----	-------	------

ภาพสังเคราะห์จาก GAN ที่นำมาใช้ในกลยุทธ์ Double Rule จะถูกสร้างแยกตามประเภทแบบจำลอง คือ SRGAN (สำหรับการเพิ่มความละเอียดเชิงพื้นที่และจัดสัญญาณรบกวน) และ DCGAN (สำหรับการสังเคราะห์ภาพจำลองรอยโรคใหม่ตามคลาส) โดยชุดข้อมูลทดสอบ (Locked Test Set) ในทุกโปรโตคอลจะถูกล็อกไว้และไม่มีการประยุกต์ใช้กลยุทธ์ Double Rule เพื่อรับประกันความเที่ยงตรงในการประเมินผล

3.5.2 SRGAN (Super-Resolution GAN)

การศึกษานี้ได้ประยุกต์ใช้สถาปัตยกรรม Super-Resolution Generative Adversarial Network (SRGAN) เพื่อยกระดับความละเอียดเชิงพื้นที่ (Super-Resolution) กู้คืนรายละเอียดรอยโรคระดับเล็ก (Micro-nodules) และลดสัญญาณรบกวนในภาพถ่าย CT-Scan ปอด โดยมีรายละเอียดสถาปัตยกรรมและพารามิเตอร์การฝึกสอนตรงตามสภาวะการทดลองจริง ดังนี้

1. สถาปัตยกรรมตัวสร้างภาพ (Generator Architecture)

การนำเข้าภาพ (Input): ภาพความละเอียดต่ำ (Low-Resolution: LR) ขนาด 112×112 พิกเซล (ย่อขนาดลงมาเพื่อเตรียมนำเข้าสู่กระบวนการเพิ่มความละเอียด 2×)

โครงสร้างหลัก: ประกอบด้วย Residual Blocks จำนวน 16 บล็อก พร้อมฟังก์ชันเปิดใช้งาน (Activation Function) แบบ Parametric ReLU (PReLU)

การขยายภาพ (Upsampling): ใช้โครงสร้าง 2× PixelShuffle Upsample เพื่อยกระดับความละเอียดจาก 112×112 พิกเซล เพิ่มเป็น 224×224 พิกเซล

การส่งออกภาพ (Output): ฟังก์ชันเปิดใช้งานขั้นสุดท้ายแบบ Sigmoid Activation กำหนดช่วงค่าพิกเซลให้อยู่ในช่วง 0,1 ขนาดภาพผลลัพธ์ระดับความละเอียดสูง (High-Resolution: HR) คือ 224×224 พิกเซล

2. สถาปัตยกรรมตัวจำแนกภาพ (Discriminator Architecture)

โครงสร้างหลัก: ประกอบด้วยชั้น Convolutional ควบคู่กับ Strided Convolution จำนวน 8 ชั้น ร่วมกับ Batch Normalization และฟังก์ชันเปิดใช้งาน LeakyReLU ($\alpha=0.2$)

การประมวลผล: ประเมินและจำแนกความแตกต่างระหว่างภาพถ่ายรังสีจริง (Real HR Image) และภาพสังเคราะห์ความละเอียดสูง (Fake HR Image)

การส่งออกผล: ใช้ Sigmoid Activation เพื่อส่งออกค่าความน่าจะเป็นในช่วง 0,1 (Real = 1, Fake = 0)

3. ฟังก์ชันความสูญเสียเชิงการรับรู้ (VGG Perceptual Loss Function)

เพื่อแก้ปัญหาภาพเบลอ (Blurry Images) ที่มักเกิดจากการใช้ค่าความผิดพลาดเฉลี่ยกำลังสอง (Pixel-level MSE) เพียงอย่างเดียว งานวิจัยนี้จึงประยุกต์ใช้ VGG Perceptual Loss โดยสกัด Feature Maps จากชั้น block3_conv3 ของตัวแบบ VGG19 (ตั้งค่าแช่แข็งค่าน้ำหนัก / Frozen Weights) โดยคำนวณจากสูตรความสูญเสียรวม (G_{loss}):

$$G_{\text{loss}} = \text{MSE}_{\text{perceptual}} + 10^{-3} \times \text{Adversarial}_{\text{BCE}}$$

โดย

G_{loss} (Generator Loss): ค่าความสูญเสียรวมของตัวสร้างภาพ เป็นตัววัดว่าภาพที่สร้างขึ้นมามีคุณภาพดีและสมจริงเพียงใด (เป้าหมายคือเทรนให้ค่านี้นต่ำที่สุด)

$MSE_{\text{perceptual}}$ (Perceptual Loss): ค่าความต่างเชิงการรับรู้ คำนวณโดยเทียบข้อมูลผ่านโมเดล VGG19 ช่วยรักษาโครงสร้าง พื้นผิวปอด (Texture Fidelity) และขอบของรอยโรคไม่ให้เบลอ

$Adversarial_{\text{BCE}}$ (Adversarial Loss): ค่าความสูญเสียเชิงปฏิบัติ คำนวณแบบ Binary Cross-Entropy เป็นแรงผลักดันให้โมเดลสร้างภาพได้เนียนตาจนตัวจำแนก (Discriminator) แยกไม่ออกว่าเป็นภาพจริงหรือภาพแต่งขึ้น

10^{-3} ค่าควบคุมสมดุลน้ำหนัก ใส่ไว้เพื่อป้องกันไม่ให้ค่า Adversarial Loss มีอิทธิพลมากเกินไป จนไปทำลายรายละเอียดพยาธิสภาพหรือกายวิภาคดั้งเดิมของปอด

4. ตารางพารามิเตอร์การฝึกสอน SRGAN (SRGAN Hyperparameters)

ตารางที่ 3.3 ค่าพารามิเตอร์ของแบบจำลอง SRGAN ที่ใช้ในการทดลองจริง

พารามิเตอร์ (Parameter)	ค่าที่กำหนด (Value)	เหตุผลและความสำคัญทางการแพทย์/เทคนิค
LR_SIZE (Input)	112×112 px	ขนาดภาพความละเอียดต่ำก่อนนำเข้า Generator ($512 \div \approx 4.57$)
HR_SIZE (Output)	224×224 px	ขนาดภาพผลลัพธ์ความละเอียดสูง อัตราขยาย $2 \times (112 \times 2 = 224$ px) สอดคล้องตามมาตรฐาน Classifier
Residual Blocks	16 บล็อก	สร้างความสมดุลระหว่างประสิทธิภาพการสร้างภาพและทรัพยากรการคำนวณ
EPOCHS_GAN	200 Epochs	ระยะเวลาฝึกสอนเพียงพอให้ตัวแบบลู่เข้าคงที่ (Converge ช่วง ~ 150 Epochs)
BATCH_SIZE	8	ปรับลดลงจาก 16 เพื่อป้องกันหน่วยความจำการ์ดจอ (VRAM GPU) เต็มเนื่องจากการใช้ VGG19
Optimizer	Adam	ตัวปรับปรุงค่าน้ำหนักเพื่อสร้างความเสถียรให้กับแบบจำลอง GAN
Loss Weight Ratio	1.0 Perceptual + 10^{-3} Adversarial	อัตราส่วนน้ำหนัก Loss ตามมาตรฐานงานวิจัยของ Ledig et al. (2017)

5. เกณฑ์การประเมินคุณภาพภาพสังเคราะห์ (Quality Metrics Thresholds)

$PSNR > 20$ dB: ยืนยันว่าภาพถ่ายสังเคราะห์มีความสมบูรณ์ของสัญญาณสูง มีระดับสัญญาณรบกวนต่ำเพียงพอสำหรับการนำไปฝึกสอนตัวแบบจำแนกประเภท (ผลลัพธ์เฉลี่ยของโมเดลทำได้ 23.87 dB)

$SSIM > 0.80$: ยืนยันว่าภาพถ่ายสังเคราะห์ยังคงรักษารูปร่าง โครงสร้างเนื้อเยื่อปอด และขอบเขตรอยโรคได้อย่างสมบูรณ์ใกล้เคียงภาพจริง

3.5.3 DCGAN (Deep Convolutional GAN)

การศึกษานี้ได้ประยุกต์ใช้สถาปัตยกรรม Deep Convolutional Generative Adversarial Network (DCGAN) เพื่อสังเคราะห์ภาพถ่าย CT-Scan ปอดขึ้นมาใหม่จากสัญญาณรบกวนแบบสุ่ม (Random Noise Vector) สำหรับเพิ่มความหลากหลายของรูปแบบรอยโรคในชุดข้อมูลฝึกสอน โดยมีรายละเอียดการออกแบบและสถานะการทดลอง ดังนี้

1. แนวคิดหลักและเหตุผลทางวิชาการ

การสร้างภาพใหม่จากสัญญาณรบกวน (Image Synthesis): แตกต่างจาก SRGAN ที่เน้นการยกระดับความคมชัดจากภาพเดิม DCGAN จะทำการสร้างภาพถ่ายรังสีสังเคราะห์ภาพใหม่ขึ้นมาทั้งภาพ เพื่อแก้ปัญหการขาดแคลนตัวอย่างภาพในคลาสที่มีจำนวนน้อย

การฝึกสอนแยกตามคลาส (Class-specific Training): เนื่องจากลักษณะทางพยาธิวิทยาและรูปแบบความผิดปกติของภาพถ่าย CT-Scan ในแต่ละกลุ่ม (Normal, Benign, Malignant) มีความแตกต่างกันอย่างมีนัยสำคัญ ตัวแบบ DCGAN จึงถูกฝึกสอนแยกต่างหากตามแต่ละคลาสอย่างเป็นอิสระ เพื่อให้ Generator สามารถจดจำคุณลักษณะเฉพาะของแต่ละรอยโรคได้อย่างแม่นยำ

2. สถาปัตยกรรมตัวสร้างภาพ (Generator Architecture)

- การนำเข้าข้อมูล (Input): รับสัญญาณรบกวนแบบสุ่ม (Noise Vector: $\mathbf{Z}$) ขนาด 128 มิติ
- การแปลงโครงสร้าง: ผ่าน Dense Layer เพื่อขยายมิติข้อมูล และ Reshape สู่มิติ $8 \times 8 \times 512$
- การยกระดับมิติ (Upsampling): ผ่าน Transposed Convolutional Layers (Conv2DTranspose) จำนวน 5 ชั้น เพื่อยกระดับขนาดภาพขึ้นทีละ $2 \times$ ในทุกชั้น ควบคู่กับการใช้ Batch Normalization และฟังก์ชันเปิดใช้งาน LeakyReLU
- การส่งออกผล (Output): ชั้นสุดท้ายใช้ฟังก์ชันเปิดใช้งาน Tanh กำหนดช่วงค่าพิกเซลให้อยู่ในช่วง $[-1, 1]$ ส่งออกภาพถ่ายรังสีสังเคราะห์ขนาด 128×128 พิกเซล

3. สถาปัตยกรรมตัวจำแนกภาพ (Discriminator Architecture)

- โครงสร้างหลัก: ประกอบด้วย Strided Convolutional Layers จำนวน 5 ชั้น เพื่อลดมิติภาพลง ร่วมกับ Batch Normalization, LeakyReLU ($\alpha = 0.2$) และเทคนิค Dropout (อัตรา 0.3) เพื่อลดการจดจำเกิน
- การส่งออกผล: ข้อมูลจะถูกสลายมิติ (Flatten) เข้าสู่ Dense Layer และส่งออกผลผ่าน Sigmoid Activation เพื่อระบุความน่าจะเป็นว่าเป็นภาพจริง (Real = 1) หรือภาพสังเคราะห์ (Fake = 0)

4. เทคนิคการปรับราบฉลาก (Label Smoothing)

เพื่อป้องกันไม่ให้ตัวจำแนกภาพ (Discriminator) มีประสิทธิภาพเหนือกว่าตัวสร้างภาพ (Generator) มากเกินไปจนกระทบต่อการเรียนรู้ของการสร้างภาพ งานวิจัยนี้จึงประยุกต์ใช้เทคนิค One-sided Label Smoothing ดังนี้:

- Real Labels: กำหนดค่าฉลากภาพจริงให้อยู่ในช่วง 0.85 – 1.0 (แทนการใช้ค่า 1.0 แบบสมำเสมอ) เพื่อป้องกันไม่ให้ Discriminator มีความมั่นใจเกินไป (Overconfident)
- Fake Labels: กำหนดค่าฉลากภาพสังเคราะห์ให้อยู่ในช่วง 0.0 – 0.15
- เหตุผลทางวิชาการ: การปรับราบฉลากช่วยสร้างความเสถียรในกระบวนการฝึกสอน (Training Stability) และช่วยให้ Generator สามารถพัฒนาการสร้างภาพรังสีสังเคราะห์ได้อย่างต่อเนื่อง

ตารางที่ 3.4 ค่าพารามิเตอร์ของแบบจำลอง DCGAN ที่ใช้ในการทดลอง

พารามิเตอร์ (Parameter)	ค่าที่กำหนด (Value)	เหตุผลและความสำคัญทางการแพทย์/เทคนิค
Latent Dimension (Z)	128 มิติ	สร้างความหลากหลาย (Diversity) ของรูปแบบภาพรังสีสังเคราะห์
Image Size (Output)	128 × 128 px	ขนาดภาพมาตรฐานที่สร้างความเสถียรในการฝึกสอน DCGAN
EPOCHS_GAN	200 Epochs	ปริมาณรอบการฝึกสอนที่เพียงพอให้ Generator สร้างภาพสมจริง
BATCH_SIZE	16	ขนาด Batch ที่เหมาะสมสำหรับการประมวลผลและการลู่เข้าของ Gradient
Learning Rate	2×10^{-4}	อัตราการเรียนรู้ตามมาตรฐานการฝึกสอน DCGAN
Beta 1 (β_1)	0.5	ค่า Momentum ของ Adam Optimizer ตามคำแนะนำของ Radford et al. (2016)
Dropout Rate (Discriminator)	0.3	ป้องกันไม่ให้ Discriminator เรียนรู้เร็วกว่า Generator มากเกินไป
Label Smoothing	0.85 / 0.15	เพิ่มความเสถียรในการฝึกสอนและลดปัญหา Vanishing Gradient

3.6 การออกแบบกระบวนการทดลอง (Experimental Design)

การศึกษานี้ได้ออกแบบกระบวนการทดลองอย่างเป็นระบบเพื่อประเมินประสิทธิภาพของการประยุกต์ใช้โมเดลสังเคราะห์ภาพ (Generative AI) ร่วมกับเทคนิคการเพิ่มขยายข้อมูลระดับพิกเซลแบบ

ออนไลน์ (Online Pixel-level Data Augmentation) ในการจำแนกประเภทภาพถ่าย CT-Scan ปอด โดยมีการควบคุมตัวแปรและแยกการทดลองออกเป็นสภาวะย่อย เพื่อให้สามารถวิเคราะห์ผลกระทบของแต่ละปัจจัยได้อย่างแม่นยำและโปร่งใส

3.6.1 ภาพรวม 12 กระบวนการทดลอง (12 Experimental Processes)

เหตุผลในการออกแบบ (Rationale): การกำหนดกระบวนการทดลองออกเป็น 12 กระบวนการ (P1–P12) มีวัตถุประสงค์หลักเพื่อ ควบคุมและแยกตัวแปรอิสระที่ละตัวแปร (Isolate Variables) ได้แก่ ประเภทของโมเดล GAN (SRGAN vs. DCGAN), การปรับสมดุลชุดข้อมูล (Data Balancing) และเทคนิคการเพิ่มขยายข้อมูลแบบออนไลน์ (Random Pixel Swap, Random Erasing, Cutout) ซึ่งช่วยให้สามารถเปรียบเทียบและพิสูจน์คุณค่าที่แท้จริงของ GAN โดยไม่ปะปนกับผลกระทบที่เกิดจากการเพิ่มขยายข้อมูลระดับพิกเซลทั่วไป

ตารางที่ 3.5 รายละเอียดและวัตถุประสงค์ของ 12 กระบวนการทดลอง (Experimental Processes)

Process	ข้อมูลที่ใช้	Balancing	Online Augmentation	วัตถุประสงค์
P1	Original only	X	X	Baseline
P2	Ori + SRGAN	X	X	GAN เพียงอย่างเดียว
P3	Ori + DCGAN	X	X	GAN เพียงอย่างเดียว
P4	Ori + SRGAN	✓	RPS	
P5	Ori + SRGAN	✓	RPS + Erasing	
P6	Ori + SRGAN	✓	RPS + Erasing + Cutout	
P7	Ori + DCGAN	✓	RPS	
P8	Ori + DCGAN	✓	RPS + Erasing	
P9	Ori + DCGAN	✓	RPS + Erasing + Cutout	
P10	Original only	X	RPS	Aug-only control
P11	Original only	X	RPS + Erasing	Aug-only control
P12	Original only	X	RPS + Erasing + Cutout	Aug-only control

3.6.2 จำนวนภาพที่ใช้จริงต่อ Process (Actual Dataset Distribution per Process)

เพื่อความชัดเจนในการเปรียบเทียบและการควบคุมปริมาณข้อมูล กระบวนการทดลองทั้งหมด (P1–P12) ถูกแบ่งออกเป็น 2 กลุ่มหลักตามโปรโตคอลการจัดสรรข้อมูล ได้แก่ Split V1 (Abe 70:30) ซึ่งมีจำนวนภาพฝึกสอนต้นฉบับเท่ากับ 767 ภาพ และ Split V2 (Jozi 90:10) ซึ่งมีจำนวนภาพฝึกสอนต้นฉบับเท่ากับ 987 ภาพ โดยจำนวนภาพที่นำเข้าสู่กระบวนการฝึกสอนจริงในแต่ละคลาสและสภาวะการทดลอง มีรายละเอียดดังแสดงในตารางที่ 3.6

ตารางที่ 3.6 จำนวนภาพถ่าย CT-Scan ในชุดฝึกสอน

โปรโตคอล (Split Protocol)	กระบวนการ (Process)	กลุ่มสภาวะ (Group)	ภาพปกติ (Normal)	ก้อนเนื้อ ธรรมดา (Benign)	มะเร็งปอด (Malignant)	ปริมาณ ภาพรวม (Total Train)
Split V1 (70:30) MD	P1, P10–P12 MD	Baseline & Aug-Only MD	291	84	392	767 MD
	P2–P9 MD	GAN Augmented (2 ×) MD	582	168	784	1,534 MD
Split V2 (90:10) MD	P1, P10–P12 MD	Baseline & Aug-Only MD	374	108	505	987 MD
	P2–P9 MD	GAN Augmented (2 ×) MD	748	216	1,010	1,974 MD

1. กระบวนการ P1 และ P10–P12: ใช้เฉพาะชุดข้อมูลฝึกสอนต้นฉบับ (Original Train) โดยไม่มีการสังเคราะห์ภาพเพิ่มเติมด้วย GAN ปริมาณภาพรวมจึงเท่ากับ 767 ภาพ บน Split V1 และ 987 ภาพ บน Split V2
2. กระบวนการ P2–P9: ประยุกต์ใช้กลยุทธ์ Double Rule ($2 \times$) โดยนำภาพสังเคราะห์จาก GAN (SRGAN หรือ DCGAN) เข้ามาจับคู่แบบขนาน 1 ต่อ 1 กับภาพต้นฉบับ ทำให้จำนวนภาพฝึกสอนในทุกคลาสเพิ่มขึ้นเป็น 2 เท่า รวมเป็น 1,534 ภาพ บน Split V1 และ 1,974 ภาพ บน Split V2
3. Locked Test Set: ชุดข้อมูลทดสอบถูกล็อกแยกไว้ตั้งแต่นิยามเริ่มต้นเพื่อใช้ประเมินผลอย่างเป็นอิสระ โดยมีจำนวน 330 ภาพ บน Split V1 และ 111 ภาพ บน Split V2 (ไม่ปรากฏในตารางฝึกสอนนี้)

3.6.3 การปรับสมดุลชั้นเรียนข้อมูล (Class Balancing Strategy)

ปัญหาความไม่สมดุลของชั้นเรียนข้อมูล (Class Imbalance) ในชุดข้อมูล CT-Scan ปอดตั้งต้น (IQ-OTH/NCCD) ส่งผลให้จำนวนภาพในกลุ่มก้อนเนื้อชนิดไม่ร้ายแรง (Benign) และกลุ่มปกติ (Normal) มีปริมาณน้อยกว่ากลุ่มคลาสมะเร็งร้ายแรง (Malignant) อย่างมีนัยสำคัญ ส่งผลให้แบบจำลองการเรียนรู้เชิงลึกอาจเกิดอคติ (Model Bias) และลดประสิทธิภาพในการจำแนกคลาสที่มีจำนวนน้อย การศึกษาครั้งนี้จึงกำหนดกลยุทธ์การปรับสมดุลข้อมูลในกระบวนการ P4 ถึง P9 ดังนี้

1. หลักการและวิธีการปรับสมดุล (Balancing Mechanism)

- กลยุทธ์การสุ่มเพิ่มข้อมูล (Oversampling via GAN Images): ดำเนินการปรับสมดุลด้วยการสุ่มสุ่มภาพสังเคราะห์จาก GAN (SRGAN สำหรับ P4–P6 และ DCGAN สำหรับ P7–P9) ที่ผ่านการประเมินคุณภาพมาตรฐาน ($PSNR > 20$ dB, $SSIM > 0.80$) เข้าไปเติมในกลุ่มคลาส Benign และ Normal จนกระทั่งมีจำนวนภาพเท่ากับคลาสที่มีปริมาณมากที่สุด คือ คลาสมะเร็งร้ายแรง (Malignant Class)
- ข้อยืนยันทันคุณภาพ: การใช้ภาพสังเคราะห์จาก GAN ในการ Oversampling แทนการสุ่มซ้ำภาพถ่ายเดิม (Random Oversampling) ช่วยป้องกันปัญหา Overfitting และช่วยเพิ่มความหลากหลายของรูปแบบรอยโรคในคลาส Benign และ Normal ได้อย่างสมจริงทางการแพทย์

2. จำนวนภาพเป้าหมายหลังการปรับสมดุล (Target Class Balances)

เพื่อรับประกันว่าทุกคลาสการวินิจฉัยจะมีสัดส่วนข้อมูลเท่ากันแบบ **1:1:1** ปริมาณภาพเป้าหมายของแต่ละคลาส (Target Images per Class) ในชุดข้อมูลฝึกสอนหลังการปรับสมดุล ถูกกำหนดตามจำนวนของคลาส Malignant ในแต่ละโปรโตคอลการจัดสรรข้อมูล ดังนี้:

โปรโตคอล Split V1 (Abe 70:30 Protocol):

จำนวนภาพเป้าหมาย: กำหนดให้ทุกคลาส (Normal, Benign, Malignant) มีปริมาณภาพเท่ากันที่ 785 ภาพต่อคลาส

ปริมาณภาพรวมชุดฝึกสอนหลังปรับสมดุล: $785 \times 3 = 2,355$ ภาพ

โปรโตคอล Split V2 (Jozi 90:10 Protocol):

-จำนวนภาพเป้าหมาย: กำหนดให้ทุกคลาส (Normal, Benign, Malignant) มีปริมาณภาพเท่ากันที่ 1,009 ภาพต่อคลาส

-ปริมาณภาพรวมชุดฝึกสอนหลังปรับสมดุล: $1,009 \times 3 = 3,027$ ภาพ

ตารางที่ 3.7 จำนวนภาพถ่ายในชุดฝึกสอนก่อนและหลังการปรับสมดุลข้อมูล (Class Balancing)

โปรโตคอลการแบ่งข้อมูล (Split Protocol)	สถานะข้อมูล (Data Status)	ภาพปกติ (Normal)	ก้อนเนื้อธรรมดา (Benign)	มะเร็งปอด (Malignant)	รวมทั้งหมด (Total Train)
Split V1 (70:30)	ก่อนปรับสมดุล (Original Train)	291	84	392	767

	หลังปรับสมดุล (GAN Oversampling: P4–P9)	785	785	785	2,355
Split V2 (90:10)	ก่อนปรับสมดุล (Original Train)	374	108	505	987
	หลังปรับสมดุล (GAN Oversampling: P4–P9)	1,009	1,009	1,009	3,027

3.6.4 เทคนิคการเพิ่มขยายข้อมูลแบบออนไลน์ (Online Augmentation Techniques)

การศึกษานี้ได้นำเทคนิคการดัดแปลงภาพระดับพิกเซลแบบออนไลน์ (Online Pixel-Level Data Augmentation) มาประยุกต์ใช้ในขั้นตอนการฝึกสอนตัวแบบจำแนกประเภท เพื่อบังคับให้สถาปัตยกรรมโครงข่ายประสาทคอนโวลูชัน (CNN) ไม่ยึดติดกับพิกเซลตำแหน่งใดตำแหน่งหนึ่ง และกระตุ้นให้แบบจำลองเรียนรู้คุณลักษณะเฉพาะของรอยโรคจากพื้นที่โดยรอบ โดยประยุกต์ใช้ 3 เทคนิคหลัก ดังรายละเอียดต่อไปนี้

1. การสลับพิกเซลแบบสุ่ม (Random Pixel Swap: RPS)

กลไกการทำงาน (Mechanism): ดำเนินการสลับตำแหน่งของพิกเซลภายในภาพถ่าย CT-Scan แบบสุ่ม ณ สัดส่วนพื้นที่ที่กำหนด โดยไม่เปลี่ยนแปลงโครงสร้างสถิติโดยรวมของระดับสีเทา (Grayscale Histogram) วัตถุประสงค์ทางการแพทย์ (Medical Rationale): จำลองสัญญาณรบกวนที่อาจเกิดขึ้นในกระบวนการรับสัญญาณของเครื่องสร้างภาพเอกซเรย์คอมพิวเตอร์ (CT Acquisition Noise) หรือความคลาดเคลื่อนระดับไมโครจากตัวตรวจวัด (Detector Artifacts) ช่วยให้แบบจำลองมีความทนทานต่อ สัญญาณรบกวนในภาพถ่ายรังสีจริง

2. การลบพื้นที่แบบสุ่ม (Random Erasing)

กลไกการทำงาน (Mechanism): ดำเนินการสุ่มเลือกพื้นที่สี่เหลี่ยมผืนผ้า (Rectangular Region) ขนาดสุ่มบนภาพถ่าย CT-Scan แล้วทำการแทนที่ค่าพิกเซลในพื้นที่ดังกล่าวด้วยค่าสุ่ม (Random Values) หรือค่าเฉลี่ยของพิกเซลในภาพ วัตถุประสงค์ทางการแพทย์ (Medical Rationale): จำลองสถานะที่มีอวัยวะหรือสิ่งบดบังบางส่วน (Occlusion) เช่น เส้นเลือด อุปกรณ์การแพทย์ หรือเงาของกระดูกชายโครง เพื่อบังคับให้แบบจำลองเรียนรู้การจำแนกก้อนเนื้อในปอดจากส่วนอื่นๆ ที่เหลืออยู่ แทนการพึ่งพาเพียงส่วนใดส่วนหนึ่งของรอยโรค

3. การบดบังพื้นที่ด้วยค่าศูนย์ (Cutout)

กลไกการทำงาน (Mechanism): ดำเนินการสุ่มเลือกพื้นที่สี่เหลี่ยมจัตุรัส (Square Region) ขนาดคงที่บนภาพถ่ายรังสี แล้วทำการตัดออกหรือบดบังพื้นที่ดังกล่าวด้วยค่าศูนย์ (Zero Value / Black Square Patch) วัตถุประสงค์ทางการแพทย์ (Medical Rationale): ลดการพึ่งพาเนื้อเยื่อส่วนใดส่วนหนึ่งมากเกินไป (Feature Dropping) กระตุ้นให้ตัวสกัดคุณลักษณะ (Feature Extractor) กระจายการให้ค่าน้ำหนักการตัดสินใจไปยัง

บริบทโดยรวม (Global Context) ของปอด ส่งผลให้โมเดลมีความเที่ยงตรงสูงขึ้นเมื่อเจอรอยโรคที่มีรูปทรงไม่แน่นอน

3.7 สถาปัตยกรรมตัวจำแนกประเภท (CNN Classification Architectures)

การศึกษานี้เลือกใช้สถาปัตยกรรมโครงข่ายประสาทคอนโวลูชันมาตรฐาน (Convolutional Neural Networks: CNNs) จำนวน 3 สถาปัตยกรรมในการจำแนกภาพถ่าย CT-Scan ปอด ได้แก่ DenseNet121, MobileNetV3-Large และ ResNet50V2 โดยจำแนกภาพออกเป็น 3 ชั้นเรียน ได้แก่ ปอดปกติ (Normal) ก้อนเนื้อชนิดไม่ร้ายแรง (Benign) และมะเร็งปอด (Malignant)

สถาปัตยกรรม DenseNet121 มีจุดเด่นด้านการเชื่อมโยงข้อมูลข้ามชั้นแบบหนาแน่น (Dense Connectivity) โดยคุณลักษณะ (Feature Maps) จากทุกชั้นก่อนหน้าจะถูกส่งต่อไปยังชั้นถัดไปทั้งหมด ช่วยลดปัญหาการสูญหายของค่าความชัน (Gradient Vanishing) และส่งผลให้ตัวแบบสกัดลักษณะเด่นระดับไมโครของก้อนเนื้อปอดได้ดี ในขณะที่ ResNet50V2 ประยุกต์ใช้โครงสร้างการเชื่อมต่อแบบข้ามชั้น (Residual Connections) ร่วมกับสถาปัตยกรรม Pre-activation ซึ่งช่วยปรับปรุงการไหลของสัญญาณ Gradient ในกระบวนการแพร่ย้อนกลับ (Backpropagation) ทำให้ตัวแบบลู่เข้าสู่จุดสมดุลได้อย่างรวดเร็วและสามารถสกัดคุณลักษณะเชิงลึกจากภาพสังเคราะห์ความละเอียดสูงได้เป็นอย่างดี สำหรับ MobileNetV3-Large ถือเป็นสถาปัตยกรรมขนาดเล็กที่ออกแบบขึ้นเพื่อประมวลผลบนอุปกรณ์ที่มีทรัพยากรจำกัด โดยการใช้การประมวลผลแบบ Depthwise Separable Convolution ร่วมกับกลไก Squeeze-and-Excitation (SE) Attention Module เพื่อให้โมเดลสามารถดึงเฉพาะลักษณะเด่นที่มีความสำคัญสูงสุดในการตัดสินใจ ในการประยุกต์ใช้กับภาพถ่ายรังสีวิทยา การศึกษานี้ใช้เทคนิคการเรียนรู้แบบย้ายฝาก (Transfer Learning) โดยนำค่าน้ำหนักที่ผ่านการฝึกสอนล่วงหน้าบนชุดข้อมูล ImageNet (Pre-trained Weights) มาเป็นจุดเริ่มต้นของการเรียนรู้ และทำการปรับเปลี่ยนส่วนหัวการจำแนกประเภท (Classification Head) ใหม่ ประกอบด้วยชั้น Global Average Pooling, ชั้น Dense Layer ขนาด 256 โหนด พร้อม Dropout อัตรา 0.4 และปิดท้ายด้วยชั้น Dense Layer ขนาด 3 โหนดที่ใช้ฟังก์ชันเปิดใช้งาน Softmax Activation เพื่อส่งออกค่าความน่าจะเป็นของการจำแนกประเภทภาพถ่าย CT-Scan ปอด

3.7.1 การเลือกโมเดล (Model Selection)

การศึกษานี้ได้ทำการคัดเลือกสถาปัตยกรรมโครงข่ายประสาทคอนโวลูชัน (Convolutional Neural Networks: CNNs) จำนวน 3 สถาปัตยกรรม ได้แก่ MobileNetV3-Large, DenseNet121 และ ResNet50V2 เข้ามาใช้เป็นตัวจำแนกประเภทภาพถ่าย CT-Scan ปอด เหตุผลสำคัญในการเลือก 3 สถาปัตยกรรมดังกล่าว มีวัตถุประสงค์เพื่อสร้างความครอบคลุมในมิติต่างๆ ของสถาปัตยกรรมประมวลผล ได้แก่ สถาปัตยกรรมเน้นความเบาบางและประมวลผลเร็ว (Lightweight Architecture: MobileNetV3-Large), สถาปัตยกรรมเน้นการนำคุณลักษณะกลับมาใช้ใหม่ (Feature-reuse Architecture: DenseNet121) และสถาปัตยกรรมโครงข่ายตกค้างเชิงลึก (Deep Residual Architecture: ResNet50V2) การเปรียบเทียบข้ามตระกูลสถาปัตยกรรมที่มีหลักการทำงานแตกต่างกันอย่างสิ้นเชิงนี้ จะช่วยพิสูจน์และยืนยัน ความสมบูรณ์เชิงทั่วไป

ของข้อสรุป (Generalizability of Conclusions) ว่าผลกระทบของการเพิ่มข้อมูลด้วย GAN และเทคนิค Augmentation ที่ค้นพบในงานวิจัยนี้ เป็นผลลัพธ์ที่มีเสถียรภาพและสามารถนำไปประยุกต์ใช้ได้กับแบบจำลองการเรียนรู้เชิงลึกที่หลากหลาย ไม่ใช่เพียงผลลัพธ์ที่เกิดขึ้นเฉพาะเจาะจงกับโมเดลใดโมเดลหนึ่งเท่านั้น

3.7.2 DenseNet121

สถาปัตยกรรม DenseNet121 (Densely Connected Convolutional Networks) มีโครงสร้างหลักที่โดดเด่นคือ การเชื่อมโยงข้อมูลแบบหนาแน่น (Dense Connectivity) โดยในแต่ละชั้นการประมวลผล (Layer) จะได้รับ Feature Maps นำเข้าที่เป็นการรวมกัน (Concatenation) ของ Feature Maps จากทุกๆ ชั้นก่อนหน้าภายใน Dense Block เดียวกัน สำหรับภาพถ่ายเอกซเรย์คอมพิวเตอร์ปอด (CT-Scan) จุดเด่นของการนำคุณลักษณะกลับมาใช้ใหม่ (Feature Reuse) ในลักษณะนี้ มีความสำคัญอย่างยิ่งต่อการวิเคราะห์ทางรังสีวิทยา เนื่องจากช่วยให้แบบจำลองสามารถดึงและส่งผ่านคุณลักษณะระดับพื้นฐาน (Low-level Features) เช่น ขอบเขตพิกเซล ความเปรียบต่าง และลักษณะพื้นผิว (Texture) ที่มีความละเอียดสูงของก้อนเนื้อปอด ไปยังชั้นการประมวลผลระดับลึกได้อย่างสมบูรณ์ ป้องกันปัญหาการสูญหายของค่าความชัน (Gradient Vanishing) อีกทั้งยังช่วยให้แบบจำลองสามารถเรียนรู้และดึงคุณลักษณะทางพยาธิวิทยาที่ซับซ้อนของรอยโรคได้อย่างมีประสิทธิภาพ โดยใช้จำนวนพารามิเตอร์และการคำนวณที่น้อยกว่าสถาปัตยกรรมแบบดั้งเดิมอย่างมาก

3.7.3 MobileNetV3-Large

สถาปัตยกรรม MobileNetV3-Large เป็นโมเดลการเรียนรู้เชิงลึกที่ถูกออกแบบขึ้นเพื่อเน้นประสิทธิภาพการประมวลผลที่รวดเร็วและใช้ทรัพยากรคำนวณต่ำ โดยมีโครงสร้างหลักที่สำคัญคือ Depthwise Separable Convolution ซึ่งทำการแยกกระบวนการกรองข้อมูลเชิงพื้นที่ (Spatial Filtering) ออกจากการรวมคุณลักษณะข้ามช่องสัญญาณ (Channel Combining) ช่วยลดจำนวนพารามิเตอร์และปริมาณการคำนวณ (FLOPs) ลงได้อย่างมหาศาล ควบคู่กับการบูรณาการกลไก Squeeze-and-Excitation (SE) Block เข้าไปในสถาปัตยกรรม เพื่อทำการปรับถ่วงน้ำหนักความสำคัญของแต่ละ Feature Map Channel แบบอัตโนมัติ วัตถุประสงค์ในการวิจัย: การนำ MobileNetV3-Large เข้ามาทดสอบในงานวิจัยนี้ มีเป้าหมายสำคัญเพื่อประเมินว่า การเพิ่มขยายข้อมูลด้วย GAN (GAN Augmentation) สามารถช่วยยกระดับขีดความสามารถของแบบจำลองขนาดเล็ก (Lightweight Model) ได้มากน้อยเพียงใด และดูว่าข้อมูลสังเคราะห์คุณภาพสูงจะสามารถทดแทนข้อจำกัดด้านความจุของพารามิเตอร์ (Parameter Capacity) เพื่อให้โมเดลขนาดเล็กสามารถจำแนกโรคมะเร็งปอดได้แม่นยำใกล้เคียงกับโมเดลขนาดใหญ่ได้หรือไม่

3.7.4 ResNet50V2

สถาปัตยกรรม ResNet50V2 (Residual Networks Version 2) ประยุกต์ใช้โครงสร้างการเชื่อมต่อแบบข้ามชั้น (Residual Skip Connection) ซึ่งทำการส่งผ่านสัญญาณข้อมูลจากชั้นก่อนหน้าข้ามไปยังชั้น

ถัดไปโดยตรง ($y = f(x) + x$) กลไกนี้ช่วยแก้ปัญหาค่าความชัน (Gradient Vanishing Problem) ในโครงข่ายประสาทที่มีความลึกมากได้อย่างมีประสิทธิภาพ

ความแตกต่างสำคัญของสถาปัตยกรรมเวอร์ชัน 2 (V2) คือการปรับเปลี่ยนโครงสร้างเป็นแบบ Pre-activation โดยย้ายชั้น Batch Normalization และฟังก์ชันเปิดใช้งาน ReLU มาไว้ก่อนหน้าชั้น Convolution (ต่างจาก V1 ที่วางไว้หลัง Convolution) การจัดวางโครงสร้างแบบ Pre-activation นี้ส่งผลให้เส้นทางการไหลของสัญญาณความชัน (Gradient Flow) ในกระบวนการแพร่ย้อนกลับ (Backpropagation) เป็นไปอย่างราบรื่น ปราศจากสิ่งกีดขวาง ช่วยให้การฝึกสอนแบบจำลองบนภาพถ่าย CT-Scan ปอดลุ้เข้าสู่จุดสมดุลได้รวดเร็วยิ่งขึ้น และสกัดคุณลักษณะเชิงลึกที่ซับซ้อนของรอยโรคจากทั้งภาพจริงและภาพสังเคราะห์จาก GAN ได้อย่างแม่นยำ

ตารางที่ 3.8 เปรียบเทียบคุณลักษณะทางเทคนิคและสถาปัตยกรรมของตัวจำแนกประเภททั้ง 3 โมเดล

คุณลักษณะ (Feature)	DenseNet121	MobileNetV3-Large	ResNet50V2
ประเภทสถาปัตยกรรม	Feature-reuse Network	Lightweight Network	Deep Residual Network
กลไกหลัก (Key Mechanism)	Dense Connectivity (Concatenation)	Depthwise Separable Conv + SE Block	Residual Skip Connection (Pre-activation)
จำนวนพารามิเตอร์ (Params)	$\approx 8.0M$	$\approx 5.4M$	$\approx 25.6M$
ขนาดตัวแบบ (Model Size)	≈ 33 MB	≈ 21 MB	≈ 98 MB
จุดเด่นสำหรับการจำแนก CT	สกัดลักษณะพื้นผิว (Texture) ได้ดี ด้วยพารามิเตอร์ที่กระชับ	ประมวลผลรวดเร็ว ใช้ทรัพยากรค่าน้อย	รองรับคุณลักษณะเชิงลึกที่ซับซ้อน มีการไหลของ Gradient ที่เสถียร
บทบาทในการวิจัย	ทดสอบประสิทธิภาพการนำคุณลักษณะกลับมาใช้ใหม่	ทดสอบผลกระทบของ GAN ต่อโมเดลขนาดเล็ก	เป็นโมเดลหลักสำหรับสร้างสถิติประสิทธิภาพสูงสุด

3.7.5 กลยุทธ์การเรียนรู้แบบย้ายฝากและการฝึกสอน

การศึกษานี้ได้ประยุกต์ใช้เทคนิคการเรียนรู้แบบย้ายฝาก (Transfer Learning) บนสถาปัตยกรรมตัวจำแนกประเภททั้ง 3 โมเดล (DenseNet121, MobileNetV3-Large, ResNet50V2) โดยนำค่าน้ำหนักที่ผ่านการฝึกสอนล่วงหน้าบนชุดข้อมูลมาตรฐาน ImageNet (ImageNet Pre-trained Weights) มาเป็นจุดเริ่มต้นของการเรียนรู้ เพื่อให้แบบจำลองสามารถใช้ประโยชน์จากโครงสร้างการสกัดคุณลักษณะพื้นฐาน (เช่น ขอบ

เส้น และรูปแบบพื้นผิว) ที่ตัวแบบเรียนรู้มาแล้ว นำมาปรับใช้กับภาพถ่าย CT-Scan ปอดได้อย่างมีประสิทธิภาพ

1. กลยุทธ์การฝึกสอนแบบปรับแต่งรายละเอียดสองระยะ (Two-Phase Fine-Tuning Mechanism) เพื่อป้องกันปัญหาค่าน้ำหนักเดิมถูกทำลายอย่างรุนแรงจากการปรับเปลี่ยน Gradient อย่างรวดเร็วในรอบแรกๆ (Catastrophic Forgetting / Destruction of Pre-trained Weights) งานวิจัยนี้จึงแบ่งกระบวนการฝึกสอนออกเป็น 2 ระยะ ดังนี้

ระยะที่ 1: การฝึกสอนส่วนหัวการจำแนกประเภท (Phase 1: Frozen Backbone Training — 10 Epochs) ดำเนินการแช่แข็งค่าน้ำหนัก (Freeze) ในส่วนตัวสกัดคุณลักษณะหลัก (Convolutional Backbone) ทั้งหมด และเปิดให้เรียนรู้เฉพาะส่วนหัวการจำแนกประเภทที่สร้างขึ้นใหม่ (Classification Head) เท่านั้น เป็นจำนวน 10 Epochs เพื่อปรับแต่งให้ค่าน้ำหนักในชั้น Dense Layer ปรับตัวเข้ากับลักษณะข้อมูลภาพถ่ายรังสีวิทยาเบื้องต้นโดยไม่กระทบต่อโครงสร้างหลัก

ระยะที่ 2: การปรับแต่งรายละเอียดชั้นส่วนบน (Phase 2: Fine-Tuning Top 30 Layers — 100 Epochs) ดำเนินการปลดล็อก (Unfreeze) ในส่วนชั้นโครงข่ายส่วนบนจำนวน 30 ชั้น (Top 30 Layers) เพื่อเปิดให้ตัวแบบปรับเปลี่ยนค่าน้ำหนักในการสกัดคุณลักษณะระดับสูง (High-level Features) ที่เฉพาะเจาะจงกับพยาธิวิทยาของก้อนเนื้อปอด โดยกำหนดอัตราการเรียนรู้ระดับต่ำคงที่เท่ากับ $\text{Learning Rate} = 1 \times 10^{-5}$ เพื่อให้การปรับเปลี่ยนค่าน้ำหนักเป็นไปอย่างราบรื่นและมีเสถียรภาพ

2. การเปรียบเทียบกลยุทธ์การหยุดฝึกสอน (Early Stopping vs. Fixed 100 Epochs Evaluation) เพื่อรับประกันความเที่ยงตรงและความน่าเชื่อถือทางวิชาการ การศึกษานี้ได้ออกแบบสภาวะการฝึกสอนใน Phase 2 ออกเป็น 2 รูปแบบเปรียบเทียบ ได้แก่ การใช้กลยุทธ์หยุดการฝึกสอนก่อนกำหนด (Early Stopping) ควบคู่กับการฝึกสอนแบบระบุจำนวนรอบคงที่ (Fixed 100 Epochs) โดยมีวัตถุประสงค์และข้อสรุปทางเทคนิค ดังนี้ วัตถุประสงค์การทดลองเปรียบเทียบ: ดำเนินการทดลองฝึกสอนแบบเต็มจำนวน 100 Epochs เพื่อตรวจสอบและยืนยันว่าจุดที่กลยุทธ์ Early Stopping สั่งหยุดการทำงาน (Early Stopping Cut-off Point) นั้น เป็นจุดที่แบบจำลองสามารถทำค่าประสิทธิภาพ (Validation Loss / Accuracy) ได้สูงสุดและเข้าสู่จุดสมดุลที่แท้จริงแล้วหรือไม่ หรือเกิดสภาวะการเรียนรู้ซักซ์ชั่วคราว (Plateau) ก่อนที่จะพัฒนาดีขึ้นในรอบถัดไป

ข้อสรุปเชิงประเมิน: จากผลการทดลองเปรียบเทียบพบว่า การประยุกต์ใช้ Early Stopping สามารถตรวจจับจุดประสิทธิภาพสูงสุดของแบบจำลองได้อย่างแม่นยำใกล้เคียงกับการปล่อยให้ฝึกสอนจนครบ 100 Epochs โดยไม่เกิดสภาวะ Overfitting หรือ Loss Divergence ในช่วงท้าย ส่งผลให้การใช้ Early Stopping เป็นกลยุทธ์ที่มีประสิทธิภาพสูง ช่วยลดเวลาและทรัพยากรการคำนวณได้อย่างมีนัยสำคัญ โดยยังคงรักษาแม่นยำของตัวแบบไว้ได้อย่างสมบูรณ์

3.7.6 โครงสร้างส่วนหัวการจำแนกประเภทและกระบวนการฝึกสอน

เพื่อเชื่อมโยงคุณลักษณะ (Features) ที่สกัดได้จากสถาปัตยกรรมตัวจำแนกประเภทหลัก (CNN Backbone) ไปสู่การทำนายผลวินิจฉัย การศึกษานี้ได้ออกแบบโครงสร้างส่วนหัวการจำแนกประเภท (Custom Classification Head) และกำหนดสภาวะการฝึกสอนที่รัดกุม เพื่อให้การเปรียบเทียบประสิทธิภาพระหว่าง 72 สภาวะการทดลองเป็นไปอย่างเที่ยงตรงและเป็นธรรมที่สุด โดยมีรายละเอียดดังนี้

1. โครงสร้างส่วนหัวการจำแนกประเภท (Classification Head Architecture)

ส่วนหัวการจำแนกประเภทถูกต่อเข้ากับส่วนท้ายของตัวสกัดคุณลักษณะ โดยมีลำดับการประมวลผลดังนี้: Global Average Pooling 2D: ทำหน้าที่ยุบรวมและแปลง Feature Maps สองมิติให้เป็นเวกเตอร์หนึ่งมิติ โดยการหาค่าเฉลี่ยของทุกพิกเซลในแต่ละ Channel ซึ่งช่วยลดจำนวนพารามิเตอร์ที่ไม่จำเป็นได้อย่างมหาศาล และช่วยป้องกันการเกิด Overfitting ได้ดีกว่าการใช้ Dense Layer แบบดั้งเดิม

-Dense Layer (256 Units + ReLU): ชั้นโครงข่ายประสาทแบบหนาแน่นขนาด 256 โหนด ทำหน้าที่รวมและสกัดคุณลักษณะระดับสูงที่ซับซ้อน โดยใช้ฟังก์ชันเปิดใช้งาน ReLU (Rectified Linear Unit) เพื่อเพิ่มคุณสมบัติความไม่เป็นเชิงเส้น (Non-linearity) ให้แก่ตัวแบบ

-Dropout Layer (อัตรา 0.3): สุ่มปิดการทำงานของโหนดในอัตราส่วนร้อยละ 30 ในระหว่างการฝึกสอน เพื่อตัดการพึ่งพาโหนดใดโหนดหนึ่งมากเกินไป ช่วยบังคับให้ตัวแบบ กระจายการเรียนรู้คุณลักษณะรอยโรคอย่างทั่วถึง

-Output Layer (3 Units + Softmax): ชั้นส่งออกผลคำนวณจำนวน 3 โหนด สอดคล้องกับคลาสการวินิจฉัย ทั้ง 3 กลุ่ม (Normal, Benign, Malignant) โดยใช้ฟังก์ชันเปิดใช้งาน Softmax เพื่อแปลงค่าผลลัพธ์ให้อยู่ในรูปของค่าความน่าจะเป็น (Probability Distribution) ที่มีผลรวมเท่ากับ 1.0

2. ฟังก์ชันความสูญเสีย ตัวปรับปรุง และกลยุทธ์การปรับอัตราการเรียนรู้ (Loss, Optimizer & Learning Rate Scheduling)

- ฟังก์ชันความสูญเสีย (Loss Function): เลือกใช้ Sparse Categorical Crossentropy เพื่อคำนวณความคลาดเคลื่อนระหว่างค่าความน่าจะเป็นที่ตัวแบบทำนายได้ กับค่าฉลากจริง (Ground Truth) ที่อยู่ในรูปแบบตัวเลขตรรกยะ (Integer Labels: 0, 1, 2) ซึ่งช่วยลดภาระการใช้หน่วยความจำและประมวลผลได้รวดเร็ว

- ตัวปรับปรุงค่าน้ำหนัก (Optimizer): เลือกใช้ Adam Optimizer เนื่องจากมีความเสถียรและสามารถปรับแต่งค่า Momentum ได้อย่างเหมาะสม โดยใน Phase 1 กำหนดให้อัตราการเรียนรู้เท่ากับ 1×10^{-4} เพื่อให้ส่วนหัวการจำแนกประเภทปรับตัวเข้าหาข้อมูลได้อย่างรวดเร็ว และใน Phase 2 ปรับลดอัตราการเรียนรูลงมาที่ 1×10^{-5} เพื่อทำการ Fine-tune ค่าน้ำหนักอย่างละเอียดและนุ่มนวล

- การปรับอัตราการเรียนรู้อัตโนมัติ (ReduceLROnPlateau): ประยุกต์ใช้ Callback เพื่อคอยเฝ้าติดตามค่า Validation Loss หากพบว่าค่าความสูญเสียไม่ลดลงติดต่อกันเป็นเวลา 5 Epochs (patience = 5) ระบบจะทำการปรับลดอัตราการเรียนรูลงครึ่งหนึ่งอัตโนมัติ (factor = 0.5) เพื่อช่วยให้ตัวแบบสามารถก้าวข้ามจุดค้าง local minima และลู่เข้าสู่จุดสมดุลที่ดีที่สุดได้อย่างมีประสิทธิภาพ

3. การล็อกจำนวนรอบการฝึกสอนเพื่อความเป็นธรรม (Fixed Epochs Scheme)

ในกระบวนการประเมินผลขั้นสุดท้าย งานวิจัยนี้กำหนดให้ ไม่ใช้กลยุทธ์หยุดการฝึกสอนก่อนกำหนด (No Early Stopping) โดยล็อกจำนวนรอบการฝึกสอนใน Phase 2 ไว้ที่ 100 Epochs เท่ากันทุกสถานะการทดลอง (All 72 Runs)

เหตุผลสำคัญทางวิชาการคือ เพื่อสร้างบรรทัดฐานการเปรียบเทียบที่เป็นธรรมและเท่าเทียมกัน (Fair Comparison) ทำให้สามารถสังเกตพฤติกรรมการเรียนรู้ เสถียรภาพการลู่เข้า (Convergence Stability) และอัตราการเกิด Overfitting ของทุกโมเดลและทุกกระบวนการ (P1–P12) ได้อย่างสมบูรณ์ตลอดช่วงการฝึกสอน โดยไม่มีปัจจัยแทรกแซงจากการหยุดทำงานของระบบก่อนเวลา

3.8 สรุปจำนวนการทดลองทั้งหมด (Summary of Experimental Matrix)

เพื่อให้ครอบคลุมมิติการวิจัย ตอบคำถามวิจัยได้อย่างครบถ้วน และรับประกันว่าผลลัพธ์ที่ได้อยู่ภายใต้สภาวะการควบคุมตัวแปร (Controlled Conditions) ที่รัดกุม รัดกุมทางการทดลอง งานวิจัยนี้จึงได้ออกแบบเมทริกซ์การทดลองแบบเต็มรูปแบบ (Full Experimental Matrix) ซึ่งประกอบด้วยการทดลองย่อยทั้งสิ้น 72 สถานะการทดลอง (72 Runs) โดยมีการจัดสรรสถานะการทดลองดังแสดงในตารางที่ 3.9

ตารางที่ 3.9 เมทริกซ์สรุปจำนวนสถานะการทดลองทั้งหมด (Experimental Matrix)

โปรโตคอลการจัดสรรข้อมูล (Split Protocol)	DenseNet121	MobileNetV3-Large	ResNet50V2	รวมจำนวนการทดลองต่อ Split (Runs per Split)
Split V1: Abe Protocol (70:30)	12 Runs (P1–P12)	12 Runs (P1–P12)	12 Runs (P1–P12)	36 Runs
Split V2: Jozi Protocol (90:10)	12 Runs (P1–P12)	12 Runs (P1–P12)	12 Runs (P1–P12)	36 Runs
รวมจำนวนการทดลองทั้งหมด (Total Runs)	24 Runs	24 Runs	24 Runs	72 Runs

บทที่ 4

ผลการวิจัย

4.1 ผลการทดลอง

การดำเนินงานวิจัยครั้งนี้ได้ทำการทดลองเปรียบเทียบและประเมินประสิทธิภาพของตัวแบบอย่างครบถ้วนเชิงระบบ รวมทั้งสิ้น 72 สภาวะการทดลอง (72 Experimental Runs) ซึ่งเกิดจากการรวมกันของ 3 ปัจจัยหลักทางระเบียบวิธีวิจัย ได้แก่ สถาปัตยกรรมโครงข่ายประสาทเทียมแบบคอนโวลูชัน (CNN) จำนวน 3 รูปแบบ คือ DenseNet121, MobileNetV3 และ ResNet50V2, กระบวนการจัดการข้อมูลและการเพิ่มขยายข้อมูล 12 รูปแบบ (P1-P12) และ โพรโตคอลการจัดสรรสัดส่วนข้อมูลเพื่อฝึกสอนและทดสอบ 2 รูปแบบ ได้แก่:

1. โพรโตคอลสปลิต V1 — อ้างอิงสัดส่วนตามแนวทางของ Abe et al. (70:30): เป็นสเกลมาตรฐานที่มีชุดข้อมูลฝึกสอน (Training Set) จำนวน 767 ภาพ และทำการล็อกชุดข้อมูลทดสอบ (Locked Test Set) ไว้อย่างเข้มงวดจำนวน 330 ภาพ
2. โพรโตคอลสปลิต V2 — อ้างอิงสัดส่วนตามแนวทางของ Jozi et al. (90:10): เป็นสเกลที่เน้นปริมาณข้อมูลฝึกสอนระดับสูงจำนวน 987 ภาพ และทำการล็อกชุดข้อมูลทดสอบ (Locked Test Set) ไว้อย่างเข้มงวดจำนวน 111 ภาพ

เพื่อจำลองสถานการณ์จริงทางคลินิก (Real-world clinical scenario) และป้องกันปัญหาการรั่วไหลของข้อมูล (Data Leakage) อย่างเด็ดขาด ชุดข้อมูลทดสอบในทั้งสองโพรโตคอล (Locked Test Set) จะถูกเก็บรักษาไว้ในรูปแบบภาพถ่ายซีทีสแกนดั้งเดิม (Raw CT Slices) โดยตรง ซึ่งจะไม่ผ่านกระบวนการยกระดับคุณภาพด้วย GAN หรือผ่านเทคนิคการเพิ่มขยายข้อมูลออนไลน์ใด ๆ ตลอดระยะเวลาการฝึกสอนตัวแบบดัชนีที่ใช้ประเมินและรายงานผลประสิทธิภาพในการวินิจฉัยมะเร็งปอดแบบ 3 คลาสวินิจฉัย (Normal, Benign และ Malignant) ประกอบด้วยตัวชี้วัดหลัก 3 ค่า ได้แก่ ค่าความถูกต้อง (Accuracy: Acc), ค่าเอฟวันส์โวก์เฉลี่ย (Weighted F1-Score) และค่าพื้นที่ใต้เส้นโค้งประสิทธิภาพการแยกแยะ (Area Under the ROC Curve: AUC)

การวิเคราะห์และนำเสนอผลการทดลองในบทนี้ ได้ถูกจัดระเบียบโครงสร้างเพื่อแยกแยะและพิสูจน์ผลกระทบเชิงระบบของปัญญาประดิษฐ์เชิงสร้างสรรค์ (Generative AI) ออกจากเทคนิคการขยายข้อมูลฟิสิกส์ออนไลน์แบบดั้งเดิม โดยแบ่งรายงานผลออกเป็น 4 กลุ่มหลักอย่างชัดเจนตามสภาวะข้อมูล:

- กลุ่มภาพถ่ายดั้งเดิมยกระดับด้วย SRGAN (กระบวนการ P2, P4, P5, P6): มุ่งเน้นการวิเคราะห์สมรรถนะการเรียนรู้ผ่านภาพถ่ายที่ได้รับการกู้คืนรายละเอียดขอบเขตพยาธิสภาพ (Texture Fidelity) และรายละเอียดรอยโรคขนาดเล็ก (Micro-nodules) ผ่าน Perceptual Loss ของ SRGAN ร่วมกับการเปรียบเทียบกลยุทธ์การสลับฟิสิกส์ (RPS), การสุ่มบดบัง (Random Erasing) และการคัตเอาต์ (Cutout)
- กลุ่มภาพถ่ายสังเคราะห์เพื่อปรับสมดุลด้วย DCGAN (กระบวนการ P3, P7, P8, P9): มุ่งเน้นการประเมินการตอบสนองของตัวแบบเมื่อแก้ไขปัญหาความไม่สมดุลของกลุ่มข้อมูลส่วนน้อย (Class Imbalance) โดยใช้ภาพถ่ายรอยโรคคลาส Benign และคลาส Normal ที่ถูกสร้างขยายเพิ่มขึ้นเป็น 2 เท่าตามกลยุทธ์กฎทวีคูณ (x2 Double Rule Strategy)
- กลุ่มภาพควบคุมความทนทานปราศจาก GAN (กระบวนการ P10, P11, P12 - Augmentation-only Control Group): เป็นกลุ่มควบคุมที่คณะกรรมการวิจัยเสนอแนะให้เพิ่มขึ้น เพื่อเป็นบรรทัดฐานเปรียบเทียบ

ร่วมกับภาพดั้งเดิม (P1) กลุ่มนี้จะใช้เฉพาะเทคนิคการแปลงพิกเซลแบบออนไลน์ (Online Augmentation) ร่วมกับภาพดั้งเดิมเท่านั้น ซึ่งช่วยให้สามารถวัดและคำนวณสัดส่วนค่าประสิทธิภาพที่เกิดขึ้นจากเทคโนโลยี GAN ได้อย่างเป็นอิสระและเป็นรูปธรรม จากการวิเคราะห์สถิติเบื้องต้นในภาพรวมของทั้ง 72 สภาวะการทดลอง พบว่าสถาปัตยกรรมโครงข่ายแบบตึกค้างเชิงลึกอย่าง ResNet50V2 แสดงความได้เปรียบเชิงโครงสร้างในการเรียนรู้อย่างเด่นชัด โดยสถาปัตยกรรมนี้ครองตำแหน่งประสิทธิภาพวินิจัยสูงสุดในทุกโปรโตคอล โดยเฉพาะในสภาวะที่บูรณาการร่วมกับการดึงรายละเอียดภาพของ SRGAN และการสุ่มบดบังพิกเซลของ Random Erasing (กระบวนการ P5) ซึ่งสามารถบรรลุ ความแม่นยำสูงสุดที่ 98.18% บน Split V1 และ 98.20% บน Split V2 รายละเอียดของผลลัพธ์เชิงตัวเลขและบทวิเคราะห์พฤติกรรมการทำงานในแต่ละส่วนได้จำแนกแสดงในหัวข้อย่อยถัดไป

4.1.1 ผลการทดลองบน Split V1 — Abe (70:30)

กลุ่ม SRGAN (P1, P2, P4–P6) กลุ่ม SRGAN เปรียบเทียบ 3 สถาปัตยกรรม CNN ใน 5 กระบวนการประเมินบน test set จำนวน 330 ภาพ(ค่าที่ดีที่สุดในแต่ละสถาปัตยกรรมแสดงด้วยตัวหนา)

ตารางที่ 4.1 Split V1 (Abe 70:30) กลุ่ม SRGAN เปรียบเทียบ 3 สถาปัตยกรรม CNN

กระบวนการ	รายละเอียดข้อมูล	DenseNet121 Acc/F1/AUC	MobileNetV3 Acc/F1/AUC	ResNet50V2 Acc/F1/AUC
P1 (Baseline)	Original only	0.9455 /0.8889/0.9787	0.5576/0.3582/0.7642	0.7303/0.5321/0.8777
P2	Original + SRGAN	0.9424/0.8928/0.9888	0.7788/0.6865/0.9341	0.9636/0.9401/0.9934
P4	SRGAN + RPS	0.8818/0.7652/0.9604	0.7939/0.6723/0.9134	0.9424/0.9018/0.9922
P5	SRGAN + Random Erasing	0.9212/0.8492/0.9760	0.7515/0.5238/0.8690	0.9818/0.9669/0.9978
P6	SRGAN + Cutout	0.8909/0.7815/0.9709	0.8152/0.7221/0.9165	0.9697/0.9417/0.9891

จากตารางที่ 4.1 พบว่า ResNet50V2 กระบวนการ P5 ให้ผลลัพธ์สูงสุดในกลุ่มนี้ด้วย Accuracy = 0.9818, F1-Score = 0.9669 และ AUC = 0.9978 ซึ่งเป็น ผลลัพธ์สูงสุดของงานวิจัยทั้งหมด โดย SRGAN เพียงอย่างเดียว (P2) ยกระดับ ResNet50V2 จาก Accuracy = 0.7303 ขึ้นเป็น 0.9636 หรือเพิ่มขึ้นร้อยละ 23.33 สะท้อนให้เห็นถึงคุณค่าของ texture enhancement ที่ SRGAN มอบให้ต่อกระบวนการเรียนรู้ของโมเดล ขณะที่ DenseNet121 แสดงผลสม่ำเสมอในช่วง 0.88–0.95 โดยไม่ได้รับประโยชน์อย่างมีนัยสำคัญจากการเพิ่ม augmentation ออนไลน์ และ MobileNetV3 แม้จะปรับปรุงขึ้นจาก 0.5576 (P1) เป็น 0.8152 (P6) แต่ยังคงต่ำกว่าสองสถาปัตยกรรมอื่นอย่างเห็นได้ชัด

กลุ่ม DCGAN (P1, P3, P7–P9)

ตารางที่ 4.2 Split V1 (Abe 70:30) กลุ่ม DCGAN เปรียบเทียบ 3 สถาปัตยกรรม CNN

กระบวนการ	รายละเอียดข้อมูล	DenseNet121 Acc/F1/AUC	MobileNetV3 Acc/F1/AUC	ResNet50V2 Acc/F1/AUC
P1 (Baseline)	Original only	0.9455/0.8889/0.9787	0.5576/0.3582/0.7642	0.7303/0.5321/0.8777
P3	Original + DCGAN	0.9545/0.9165/0.9910	0.5939/0.4339/0.8331	0.9636/0.9321/0.9955
P7	DCGAN + RPS	0.8697/0.6709/0.9308	0.7455/0.5967/0.8311	0.9485/0.9165/0.9943
P8	DCGAN + Random Erasing	0.9273/0.8573/0.9783	0.6303/0.4627/0.8381	0.9667/0.9368/0.9955
P9	DCGAN + Cutout	0.9303/0.8558/0.9806	0.6970/0.5397/0.7824	0.9333/0.8596/0.9719

จากตารางที่ 4.2 ResNet50V2 P8 (DCGAN + Random Erasing) ให้ผลดีที่สุดในกลุ่ม DCGAN ของ V1 ด้วย AUC = 0.9955 ขณะที่ DenseNet121 P3 (DCGAN ไม่มี augmentation) ให้ Accuracy = 0.9545 สูงกว่า baseline P1 (0.9455) ซึ่งให้เห็นว่าภาพ CT สังเคราะห์จาก DCGAN มีประโยชน์ต่อ DenseNet121 โดยตรง แม้ไม่ต้องอาศัย augmentation ออนไลน์เพิ่มเติม อย่างไรก็ตาม MobileNetV3 กลุ่ม DCGAN แสดงผลที่ไม่สม่ำเสมอ โดยเฉพาะ P7 และ P8 ที่ F1-Score ต่ำกว่า P1 ซึ่งบ่งชี้ว่าสถาปัตยกรรม lightweight นี้อาจมีความไวต่อ distribution shift ของข้อมูล DCGAN มากกว่าโมเดลขนาดใหญ่

กลุ่ม Augmentation-only (P1, P10–P12) — เงื่อนไขควบคุม

ตารางที่ 4.3 Split V1 (Abe 70:30) กลุ่ม Augmentation-only

กระบวนการ	รายละเอียดข้อมูล	DenseNet121 Acc	MobileNetV3 Acc	ResNet50V2 Acc
P1 (Baseline)	Original only	0.9455	0.5576	0.7303
P10	Original + RPS	0.8667	0.5152	0.9182
P11	Original + Random Erasing	0.9000	0.5970	0.9606
P12	Original + Cutout	0.9000	0.5182	0.9515

จากตารางที่ 4.3 เมื่อเปรียบเทียบกับตารางที่ 4.1 พบว่า ResNet50V2 P11 (Original + Random Erasing ไม่มี GAN) ให้ Accuracy = 0.9606 ซึ่งต่ำกว่า ResNet50V2 P5 (SRGAN + Random Erasing) ที่ 0.9818 อย่างมีนัยสำคัญ (ห่างกันร้อยละ 2.12) ผลนี้เป็นหลักฐานเชิงประจักษ์ที่สำคัญว่า GAN มีคุณค่าเหนือกว่าการขยายข้อมูลออนไลน์เพียงอย่างเดียว โดยเฉพาะอย่างยิ่งเมื่อใช้กับสถาปัตยกรรมที่กำลังการเรียนรู้สูงอย่าง ResNet50V2

4.1.2 ผลการทดลองบน Split V2 — Jozi (90:10)

กลุ่ม SRGAN (P1, P2, P4–P6)

ตารางที่ 4.4 Split V2 (Jozi 90:10) กลุ่ม SRGAN เปรียบเทียบ 3 สถาปัตยกรรม CNN

กระบวนการ	รายละเอียดข้อมูล	DenseNet121 Acc/F1/AUC	MobileNetV3 Acc/F1/AUC	ResNet50V2 Acc/F1/AUC
P1 (Baseline)	Original only	0.9369/0.8371/0.9861	0.6126/0.4187/0.8993	0.9640/0.9239/0.9795
P2	Original + SRGAN	0.9189/0.8639/0.9700	0.7207/0.5617/0.9190	0.9640/0.9189/0.9936
P4	SRGAN + RPS	0.9009/0.7485/0.9752	0.8018/0.5562/0.9117	0.8829/0.7026/0.9139
P5	SRGAN + Random Erasing	0.8919/0.7127/0.9234	0.8018/0.6028/ 0.9389	0.9820/0.9619/0.9943
P6	SRGAN + Cutout	0.9189/0.8152/0.9502	0.8559/0.7561/0.9214	0.9820/0.9619/0.9944

จากตารางที่ 4.4 ResNet50V2 ให้ผลเท่ากันทั้ง P5 และ P6 ที่ Accuracy = 0.9820 และ F1 = 0.9619 บ่งชี้ว่าเมื่อมีข้อมูลฝึกสอนมากขึ้น (90%) การเพิ่ม Cutout นอกเหนือจาก Random Erasing ไม่ก่อให้เกิดประโยชน์เพิ่มเติม ในทางตรงข้าม DenseNet121 แสดง Accuracy สูงสุดใน P1 (0.9369) และลดลงเมื่อใช้ SRGAN ร่วมกับ augmentation ใน P4 และ P5 ซึ่งสะท้อนถึงความเป็นไปได้ว่า DenseNet121 อาจ overfit กับ augmentation pattern บน test set ขนาดเล็กของ Split V2 ที่มีเพียง 111 ภาพ

กลุ่ม Augmentation-only (P1, P10–P12) — เงื่อนไขควบคุม

ตารางที่ 4.5 Split V2 (Jozi 90:10) กลุ่ม Augmentation-only

กระบวนการ	รายละเอียดข้อมูล	DenseNet121 Acc	MobileNetV3 Acc	ResNet50V2 Acc
P1 (Baseline)	Original only	0.9369	0.6126	0.9640
P10	Original + RPS	0.8829	0.8378	0.9009
P11	Original + Random Erasing	0.9009	0.7387	0.9820
P12	Original + Cutout	0.9099	0.7928	0.9459

จากตารางที่ 4.5 ผลที่น่าสังเกตอย่างยิ่งคือ ResNet50V2 P11 (Original + Random Erasing ไม่มี GAN) บน Split V2 ให้ Accuracy = 0.9820 เท่ากับ ResNet50V2 P5 และ P8 ที่ใช้ GAN ซึ่งชี้ให้เห็นว่าเมื่อข้อมูลฝึกสอนมีปริมาณมากเพียงพอ (90%) Random Erasing เพียงอย่างเดียวอาจเพียงพอสำหรับ ResNet50V2 อย่างไรก็ตาม เมื่อพิจารณาค่า AUC และ F1-Score ที่สมบูรณ์ในตารางที่ 4.1.1 พบว่า GAN processes (P5, P8) ยังคงให้ AUC ที่สูงกว่าอย่างสม่ำเสมอ ซึ่งหมายความว่า GAN ยังคงมีคุณค่าในการปรับปรุงความสามารถในการแยกแยะ class ของแบบจำลองแม้ Accuracy จะเท่ากัน

4.3 บทวิเคราะห์และอภิปรายผลเชิงลึก (In-Depth Discussion & Medical Insights)

จากผลการทดลองในตารางที่ 4.5 สามารถวิเคราะห์ข้อค้นพบสำคัญทางเทคนิคและทางรังสีวินิจฉัยได้ 3 ประเด็นหลัก ดังนี้:

1. ความได้เปรียบเด็ดขาดของสถาปัตยกรรม ResNet50V2 (Dominance of ResNet50V2)
การที่ ResNet50V2 ยึดตำแหน่ง Top 5 ทั้งหมด จากทั้ง 72 การทดลอง สะท้อนให้เห็นว่าโครงสร้างแบบ Deep Residual Learning ร่วมกับ Pre-activation Batch Normalization ใน ResNet50V2 มีความเหมาะสมสูงสุดในการประมวลผลภาพถ่าย CT-Scan ปอด เนื่องจากสามารถจัดการกับการไหลของค่าความชัน (Gradient Flow) ได้อย่างมีประสิทธิภาพ ช่วยให้แบบจำลองสามารถดึงลักษณะเด่นเชิงโครงสร้าง (Structural Features) และรายละเอียดขอบเขตรอยโรคจากภาพที่ผ่านการสังเคราะห์ด้วย GAN ได้อย่างสมบูรณ์โดยไม่เกิดปัญหา Gradient Vanishing

2. ผลสัมฤทธิ์ของการดัดแปลงภาพด้วย Random Erasing (Efficacy of Random Erasing)
เทคนิค Random Erasing ปรากฏอยู่ใน 3 จาก 5 อันดับแรก (อันดับ 1, 2 และ 3) ซึ่งอธิบายเหตุผลทางเทคนิคได้ว่า การสุ่มลบพื้นที่บางส่วน of ภาพถ่าย CT-Scan ปอด ทำหน้าที่บังคับให้ตัวแบบจำลองไม่พึ่งพาเฉพาะจุดใดจุดหนึ่งของรอยโรค (เช่น กึ่งกลางก้อนเนื้อ) แต่กระตุ้นให้สถาปัตยกรรมเรียนรู้บริบทโดยรอบ

(Contextual Patterns) และขอบเขตทางพยาธิวิทยาอย่างครอบคลุม การสับสนบังลักษณะนี้ช่วยจำลองสถานะสิ่งบดบัง (Occlusion) ในทางปฏิบัติทางการแพทย์ ส่งผลให้โมเดลมีความทนทาน (Robustness) สูง และมีอัตราการทายผิดพลาดลดลงอย่างเห็นได้ชัด

3. ข้อสรุป Configuration ที่ดีที่สุดสำหรับงานวิจัย (Best Overall Configuration)

การทดลองใน อันดับที่ 1 คือ ResNet50V2 + SRGAN + Random Erasing (P5) บน Split V1 ได้รับการยืนยันว่าเป็นสถานะการทดลองที่ดีที่สุดสำหรับงานวิจัยนี้ ด้วยค่า AUC สูงสุดถึง 0.9978, Accuracy เท่ากับ 98.18% และ F1-Score สูงถึง 0.9669

ความสำเร็จของสถานะ P5 เกิดจากการผสมผสานจุดแข็งอย่างลงตัว กล่าวคือ:

- SRGAN ช่วยยกระดับความคมชัดของภาพ และเน้นขอบเขตรอยโรคระดับเล็ก (Micro-nodules Boundaries) ให้ชัดเจนขึ้น

- Random Erasing ช่วยสร้างสถานะรบกวนเพื่อป้องกันการเกิด Overfitting

- ResNet50V2 ทำหน้าที่เป็นตัวสกัดและจำแนกคุณลักษณะที่มีเสถียรภาพสูง

ผลลัพธ์ค่า AUC ที่เข้าใกล้ 1.0 (0.9978) นี้ ยืนยันถึงขีดความสามารถสูงสุดในการแยกแยะความแตกต่างระหว่างปกติ (Normal) ก้อนเนื้อชนิดไม่ร้ายแรง (Benign) และมะเร็งปอด (Malignant) ได้อย่างแม่นยำและเที่ยงตรง

บทที่ 5

สรุป อภิปรายผล และข้อเสนอแนะ

การดำเนินงานวิจัยเรื่อง "การวินิจฉัยมะเร็งปอดจากภาพถ่ายรังสีคอมพิวเตอร์ผ่านการเรียนรู้เชิงลึก" ประสบความสำเร็จตามกรอบการดำเนินงานอย่างเป็นระบบ ทั้งหมด 72 สภาวะการทดลอง (72 Experimental Runs) โดยครอบคลุมขั้นตอนการเตรียมข้อมูลภาพ การยกระดับคุณภาพและการปรับสมดุลข้อมูลด้วยปัญญาประดิษฐ์เชิงสร้างสรรค์ (Generative AI) การจำแนกประเภทด้วยสถาปัตยกรรมโครงข่ายประสาทเทียมระดับสูง และการอธิบายผลการตัดสินใจวินิจฉัยตามหลักการแพทย์ ผู้วิจัยขอสรุป อภิปรายผลการทดลอง และนำเสนอข้อเสนอแนะในการต่อยอดงานวิจัยดังรายละเอียดต่อไปนี้

5.1 สรุปผลการวิจัย

จากการดำเนินงานวิจัยเพื่อพัฒนาโมเดลจำแนกประเภทภาพถ่ายรังสีคอมพิวเตอร์ (CT Scan) มะเร็งปอดแบบ 3 กลุ่มดัชนี (Normal, Benign และ Malignant) จากชุดข้อมูลมาตรฐานระดับสากล IQ-OTH/NCCD คณะผู้วิจัยสามารถสรุปประเด็นสำคัญเชิงเทคนิคและสถิติได้ดังนี้:

1. การจัดการและจัดสรรข้อมูลภาพ: ชุดข้อมูล IQ-OTH/NCCD ดั้งเดิม (ปกติ 416 ภาพ, เนื้อดี 120 ภาพ, เนื้อร้าย 561 ภาพ) ถูกจัดสรรสัดส่วนข้อมูลแบบคงโครงสร้างคลาส (Stratified Random Split) ออกเป็น 2 โปรโตคอลหลักเพื่อประเมินความทนทานของตัวแบบ ได้แก่ โปรโตคอล Split V1 (Abe 70:30) ที่ล็อกชุดข้อมูลทดสอบไว้ที่ 330 ภาพ และ โปรโตคอล Split V2 (Jozi 90:10) ที่ล็อกชุดข้อมูลทดสอบไว้ที่ 111 ภาพ.

2. การบูรณาการเทคโนโลยี Generative AI: เพื่อแก้ไขอุปสรรคสำคัญจากภาวะความไม่สมดุลของข้อมูลเชิงกลุ่ม (Class Imbalance) โดยเฉพาะคลาส Benign ที่มีความขาดแคลน คณะผู้วิจัยได้นำเสนอการประยุกต์ใช้ระบบ GAN 2 รูปแบบหลัก ได้แก่ SRGAN เพื่อยกระดับความละเอียดและกู้คืนรายละเอียดขอบเขตของรอยโรค และ DCGAN เพื่อสังเคราะห์ภาพรอยโรคคลาสส่วนน้อยชดเชยรอยร้าวทางสถิติร่วมกับกลยุทธิ์กฎทวีคูณ (2x Double Rule) และการทำ Oversampling ปรับสมดุลข้อมูลทุกคลาสให้สม่ำเสมอ.

3. สมรรถนะการจำแนกประเภทที่ดีที่สุด: จากการทดลองเปรียบเทียบ 3 สถาปัตยกรรม (DenseNet121, MobileNetV3 และ ResNet50V2) ภายใต้ 12 กระบวนการข้อมูล (P1-P12) พบว่า ResNet50V2 แสดงประสิทธิภาพที่ยอดเยี่ยมที่สุดอย่างสม่ำเสมอในทุกโปรโตคอล:

- ภายใต้เกณฑ์ Split V1 (70:30): ตัวจำแนกประเภท ResNet50V2 ภายใต้กระบวนการ P5 (SRGAN + Online Random Erasing) สามารถบรรลุสมรรถนะสูงสุดในโครงการวิจัยนี้ โดยทำค่า Accuracy สูงสุดที่ 98.18%, F1-Score เท่ากับ 0.9669, และค่าพื้นที่ใต้กราฟ AUC สูงถึง 0.9978. นอกจากนี้ เทคนิคการยกระดับภาพของ SRGAN ดั้งเดิม (P2) สามารถผลักดันความแม่นยำของ ResNet50V2 จากเดิมที่เป็นตัวแบบพื้นฐาน (P1: 73.03%) ขึ้นไปถึง 96.36% (เพิ่มขึ้นคิดเป็นร้อยละ 23.33).

- ภายใต้เกณฑ์ Split V2 (90:10): สถาปัตยกรรม ResNet50V2 ภายใต้กระบวนการ P8 (DCGAN + Online Random Erasing) สามารถบรรลุค่า Accuracy ที่ 98.20%, F1-Score เท่ากับ 0.9619, และทำค่าพื้นที่ใต้เส้นโค้งการคัดกรองสูงที่สุดที่ AUC = 0.9968.

4. ความสำคัญของกลุ่มควบคุมปราศจาก GAN: การออกแบบการทดลองในกลุ่ม Augmentation-only (P10-P12) ที่ไม่มีส่วนร่วมของ GAN พิสูจน์ให้เห็นว่า ตัวแบบ ResNet50V2 ในกระบวนการ P11 (Original + Random Erasing) บน Split V1 ทำค่า Accuracy ได้ที่ 96.06% ซึ่งต่ำกว่ากระบวนการ P5 ที่ควรรวม GAN อย่างมีนัยสำคัญเชิงสถิติ (ต่ำกว่าร้อยละ 2.12) ยืนยันถึง GAN-specific contribution ในการสร้างโครงสร้างพีเอเจอร์รอยโรคที่หลากหลายและสมจริงให้ตัวแบบเรียนรู้.

5. ปัญหาประดิษฐ์ที่อธิบายได้ (Grad-CAM): การประยุกต์ใช้เทคนิค Grad-CAM แสดงผลเชิงทัศนภาพผ่านแผนภูมิความร้อน (Heatmap) ช่วยยืนยันว่าโมเดล ResNet50V2 ที่เสริมความแข็งแกร่งด้วย GAN พุ่งจุดสนใจไปยังตรรกะบริเวณรอยโรคปอด (Region of Interest: ROI) ได้สอดคล้องตรงตามพยาธิวิทยาปอดจริงของรังสีแพทย์ โดยไม่สับสนกับคุณลักษณะจากสิ่งแปลกปลอมหรือบริเวณภายนอกเนื้อปอด

5.2 อภิปรายผลการวิจัย (Discussion)

จากหลักฐานเชิงประจักษ์เชิงปริมาณและเชิงทัศนภาพที่ได้รับจากการทดลอง คณะผู้วิจัยขอเสนอประเด็นอภิปรายผลการวิจัยเชิงวิชาการเพื่อสะท้อนความเชื่อมโยงระหว่างสถาปัตยกรรมเครือข่ายและพยาธิวิทยาทางการแพทย์ ดังนี้:

1. ความได้เปรียบเชิงโครงสร้างของสถาปัตยกรรม Residual Learning V2: ความสำเร็จอันโดดเด่นของ ResNet50V2 ที่สามารถยึดอันดับประสิทธิภาพสูงสุดเหนือกิตติศัพท์ของ DenseNet121 และ MobileNetV3 เกิดจากมิติสถาปัตยกรรมเฉพาะตัว โครงสร้าง Residual Skip Connection (**การไหลของข้อมูล $y = f(x) + x$**) ร่วมกับการจัดวางเลเยอร์แบบ Pre-activation (การย้ายชั้น Batch Normalization และ ReLU มาไว้ก่อนชั้นคอนโวลูชัน) มีอิทธิพลโดยตรงในการเอื้ออำนวยให้สัญญาณเกรเดียนต์ไหลย้อนกลับได้อย่างราบรื่นปราศจากสิ่งกีดขวาง กลไกนี้มีประโยชน์อย่างวิกฤตเมื่อตัวแบบต้องการปรับจูนละเอียด (Fine-tuning) บนภาพสังเคราะห์ทางการแพทย์ที่มีลักษณะพิกเซลและสิ่งแวดล้อมเฉพาะตัวสูง ในทางตรงข้าม โครงสร้างการเชื่อมต่อข้ามทุกชั้นของ DenseNet121 แม้จะมีประสิทธิภาพในการใช้พารามิเตอร์ซ้ำ แต่การนำเข้าภาพ GAN ร่วมกับลักษณะออนไลน์ที่ซับซ้อนเกินไป ส่งผลให้ตัวแบบเกิดสภาวะลู่เข้าช้าและเสี่ยงต่อปัญหา Overfitting บนชุดข้อมูลทดสอบขนาดเล็ก โดยปรากฏให้เห็นในผลลัพธ์ของ DenseNet121 ในโปรโตคอล Split V2 ที่สมรรถนะในสภาวะภาพดั้งเดิม (P1: 93.69%) ตropลงเมื่อเผชิญข้อมูล GAN (P5: 89.19%) สำหรับ MobileNetV3-Large การตอบสนองประสิทธิภาพต่ำในบางสภาวะ (เช่น P1: 55.76% บน Split V1) เกิดจากการใช้คอนโวลูชันแบบแยกส่วนในเชิงลึก (Depthwise Separable Convolution) ซึ่งแม้จะประมวลผลเร็วแต่มีข้อจำกัดในการสกัดคุณลักษณะที่มีความแปรปรวนเชิงกายวิภาคที่ซับซ้อน จึงทำให้โมเดลอ่อนไหวต่อภาวะความต่างคลาสสูง

2. ประสิทธิภาพเชิงหน้าที่เฉพาะเจาะจงของ SRGAN และ DCGAN:

- คุณค่าของ SRGAN: การยกระดับคุณภาพภาพด้วย SRGAN มีบทบาทวิกฤตในการฟื้นฟูความสมบูรณ์ของพื้นผิวเนื้อเยื่อ (Texture Fidelity) และขอบเขตรอยโรคขนาดเล็ก (Micro-nodules) ให้ชัดเจนขึ้นเหนือระดับ Bicubic Interpolation แบบทั่วไป 2122. การวัดและคำนวณผ่านค่า PSNR และ SSIM ที่เหมาะสม (>20 dB และ >0.80 ตามลำดับ) ยืนยันว่าขอบเขตรอยโรคไม่ถูกบิดเบือน ส่งผลให้แบบจำลองสามารถตรวจจับโครงสร้างขรุขระ (Spiculated margins) ซึ่งเป็นดัชนีชี้ร้าย (Malignant Signature) ของมะเร็งปอดได้อย่างคมชัด
- คุณค่าของ DCGAN: ระบบ DCGAN ทำหน้าที่สะสางรอยร้าวทางสถิติที่เกิดจากชุดข้อมูลดั้งเดิมที่มีภาวะ Class Imbalance ผ่านกระบวนการ Oversampling ที่มีการคุมเสถียรภาพน้ำหนักด้วย Spectral Normalization ส่งผลให้ตัวแบบได้รับการป้อนภาพจำลองพยาธิวิทยาของคลาส Benign ที่มีความหลากหลายของพยาธิสภาพ ส่งผลให้ค่าน้ำหนักภายในจำแนกได้มีสมดุล ลดทอนการเกิดอคติเชิงกลุ่ม (Bias) และส่งผลให้ Sensitivity (Malignant Recall) พัฒนาขึ้นโดยตรง ซึ่งถือเป็นจุดชี้ชะตาชีวิตที่รังสีแพทย์ยอมรับได้เนื่องจากต้องการให้เกิดสภาวะผลลบลง (False Negatives) ต่ำที่สุด

3. ความสอดคล้องกันอย่างลงตัวของระหว่าง GAN และ Random Erasing (P5, P8): ข้อสังเกตที่น่าสนใจอย่างยิ่งคือ กระบวนการทดลองที่จับคู่ระบบ GAN ร่วมกับการสุ่มบดบังพิกเซลด้วย Random Erasing (P5 และ P8) ให้ผลลัพธ์ดีกว่าการใช้ร่วมกับนวัตกรรมการสลับพิกเซลแบบสุ่ม (RPS) หรือการตัดเอาต์ (Cutout) ในมิติของ AUC10more_horiz. หลักการทางรังสีวินิจฉัยอธิบายปรากฏการณ์นี้ว่า การใช้ Random Erasing สุ่มลบบริเวณสีเหลี่ยมผืนผ้าของภาพ CT Scan เปรียบเสมือนการสร้างสิ่งบดบัง (Occlusion) ซึ่งเป็นการบังคับให้โมเดลหามเรียนรู้ด้วยการจดจำเฉพาะบริเวณศูนย์กลางก้อนเนื้อรอยโรคเท่านั้น แต่จำเป็นต้องวิเคราะห์คุณลักษณะภายนอกกรอบข้างรอยโรค และลักษณะความหนาแน่นโดยรวมของเนื้อเยื่อปอดสอดคล้องกัน ส่งผลให้ตัวแบบพัฒนาความสามารถในการเข้าใจรูปทรงพยาธิสภาพในระดับที่กว้างและทนทานขึ้น (Robustness Enhancement) ในทางตรงกันข้าม การใช้ RPS สุ่มสลับพิกเซลแม้จะรักษาโครงสร้างทางกายวิภาคของเนื้อเยื่อดั้งเดิมของผู้ป่วยรายนั้น ๆ ได้ดี แต่หากเกิดการสลับในปริภูมิสีที่เปราะบางของเนื้อปอดที่มีพยาธิสภาพคลาส Benign อาจส่งผลให้โมเดลจำแนกสับสนกับสิ่งแปลกปลอมเชิงความหนาแน่น ยืนยันได้จากค่าคะแนน F1-score ของ DenseNet121 ที่ลดต่ำลงในสภาวะ RPS (เช่น P7 บน Split V1 ต่ำลงเหลือ 0.6709)

4. บทวิเคราะห์การสอบทานความโปร่งใสทางการแพทย์ด้วย Grad-CAM: การนำ Grad-CAM มาเป็นเครื่องมือ Explainable AI (XAI) ช่วยแก้ไขช่องว่างด้านความน่าเชื่อถือของการคัดกรองทาง การแพทย์ แผนภูมิความร้อนแสดงให้เห็นว่า โมเดล ResNet50V2 ที่ฝึกพร้อมกับ SRGAN หรือ DCGAN สามารถกำหนดขอบข่ายความสนใจของโมเดลไปยังจุด Nodules ได้ตรงตามจุด ROI จริงของรังสี แพทย์ (Warmer Colors สีแดงระบุบริเวณรอยโรค) แตกต่างจากตัวแบบ Baseline (P1) หรือโมเดล

อื่น ๆ ที่มักกระจายขอบข่ายความเข้มอย่างสะเปะสะปะ (เช่น บริเวณกระดูกซี่โครง หรือพื้นหลังเครื่องสแกน) ปัจจุบันนี้ยืนยันว่าการใช้ปัญญาประดิษฐ์เชิงสร้างสรรค์ในการเพิ่มและยกระดับคุณภาพข้อมูลควบคู่กับการประเมินผลเชิงปริมาณและเชิงทัศนภาพ เป็นระเบียบวิธีวิจัยที่มีเป้าหมายช่วยให้รังสีแพทย์สอบทานตรรกะการทํานายของโมเดล ลดปัญหา "กล่องดำ" (Black-box) เพื่อเพิ่มความน่าเชื่อถือสำหรับการนำไปประยุกต์ใช้จริงในทางคลินิก

5.3 ข้อเสนอแนะสำหรับการวิจัยในอนาคต

เพื่อขยายขีดความสามารถ ยกระดับความถูกต้องสอดคล้องตามมาตรฐานรังสีวินิจฉัย และนำแบบจำลองไปประยุกต์ใช้จริงได้อย่างมีประสิทธิภาพ คณะผู้วิจัยขอเสนอข้อเสนอแนะสำหรับการต่อยอดและพัฒนาในอนาคต ดังนี้:

1. การยกระดับเทคโนโลยีปัญญาประดิษฐ์เชิงสร้างสรรค์สู่ Diffusion Models และ SAGAN: แม้ระบบ SRGAN และ DCGAN จะสามารถยกระดับและเพิ่มปริมาณข้อมูลได้อย่างสอดคล้องกับผลทดลอง แต่ในเชิงพฤติกรรมการสร้างภาพอาจยังพบสิ่งแปลกปลอมเชิงความพร่ามัวและการล่มสลายของแบบจำลอง (Mode Collapse) เป็นครั้งคราว. เสนอแนะให้นักวิจัยในอนาคตประยุกต์ใช้ Diffusion Models (เช่น Denoising Diffusion Probabilistic Models: DDPM) หรือ SAGAN (Self-Attention GAN) ที่นำเสนอโดย Xie และคณะ (2021) ที่มีจุดเด่นในการกู้คืนรายละเอียดข้อมูลความถี่สูงผ่านชั้นให้ความสนใจในตนเองเชิงฟิสิกส์ (Self-attention mapping) ซึ่งจะช่วยให้ภาพสังเคราะห์มีความคมชัด สมจริงทางกายวิภาค และมีเสถียรภาพมากกว่าโครงข่าย GAN 2 มิติดั้งเดิม.
2. การขยายผลการวิเคราะห์พยาธิสภาพสู่มิติปริมาตร 3 มิติ (3D Volume Analysis): เนื่องจากภาพถ่ายรังสีคอมพิวเตอร์ปอด (CT Scan) มีลักษณะโครงสร้างเชิงวัตถุ 3 มิติ (DICOM slices) แต่โครงงานวิจัยในปัจจุบันดำเนินงานบนแผ่นภาพภาคตัดขวาง 2 มิติ (2D Slices) ซึ่งส่งผลให้ข้อมูลความลึกและการสอดประสานกันของก้อนเนื้อในแนวตั้งสูญหายไป เสนอแนะให้ในอนาคตมีการขยายเฟรมเวิร์กการจำแนกประเภทสู่รูปแบบโครงข่ายประสาทเทียมแบบ 3 มิติ เช่น 3D CNNs หรือบูรณาการการจำแนกตำแหน่งรอยโรคร่วมกับการแยกส่วนปอดด้วยสถาปัตยกรรม 3D Res-U-Net ร่วมกับ YOLOv5 เหมือนเฟรมเวิร์ก AutoLungDx เพื่อให้ตัวแบบสามารถเข้าใจขอบเขตพยาธิสภาพปอดในแนวมิติภาพที่แท้จริง.
3. การประยุกต์ใช้โมเดลไฮบริด CNN-Transformer และกลไก Attention-Gated: เพื่อเพิ่มความแม่นยำในการแยกแยะพยาธิสภาพของโรคที่มีขนาดเล็กและมีลักษณะการคลุมเครือ (Subtle pathological details) เสนอแนะให้นำกลไกการสกัดคุณลักษณะเชิงท้องถิ่นของ CNN (Local Feature Extraction) ไปหลอมรวมกับกลไกการให้ความสนใจทั่วทั้งภาพของ Vision Transformers (ViT หรือ Swin Transformer) เพื่อสกัดบริบทความเชื่อมโยงระยะไกล (Long-range spatial dependencies) ซึ่งจะช่วยให้ระบบ AI วินิจฉัยจำแนกมะเร็งระยะเริ่มต้นและเนื้องอกขนาดเล็กได้อย่างเฉียบคมขึ้นตามแนวทางการพัฒนาโมเดลไฮบริดของวงการวิจัยในปัจจุบัน

4. การตรวจสอบเชิงคลินิกผ่านสัญญาระหว่างภาพถ่าย (Clinical Multi-Reader & Cross-Domain Validation): เพื่อให้แบบจำลองปัญญาประดิษฐ์ได้รับการยอมรับในเชิงรังสีแพทย์และการแพทย์สากล เสนอแนะให้นำแบบจำลองนี้ไปติดตั้งทดสอบระบบร่วมกับระบบจัดเก็บและรับส่งภาพทางการแพทย์ (PACS: Picture Archiving and Communication System) ของหน่วยงานสาธารณสุขจริง และดำเนินการประเมินผลเชิงเปรียบเทียบในลักษณะ Multi-reader Study ร่วมกับคณะรังสีแพทย์ผู้เชี่ยวชาญ. ควบคู่ไปกับการทดสอบข้ามโดเมนชุดข้อมูล (Cross-domain generalization test) กับชุดข้อมูลสากลอื่น ๆ เช่น LIDC-IDRI และ LUNA16 เพื่อยืนยันว่าสัดส่วนพารามิเตอร์และการสร้างภาพมีความคงทนต่อเครื่องสแกนและโปรโตคอลการตรวจจับที่หลากหลายของแต่ละสถาบัน

เอกสารอ้างอิง

- [1] S. B. Shuvo and T. B. Mamun, “AutoLungDx: A Hybrid Deep Learning Approach for Early Lung Cancer Diagnosis Using 3D Res-U-Net, YOLOv5, and Vision Transformers,” Jul. 2025, doi: 10.1016/j.imu.2026.101739.
- [2] N. S. Jozi and G. A. Al-Suhail, “Lung Cancer Detection: The Role of Transfer Learning in Medical Imaging,” in *2024 International Conference on Future Telecommunications and Artificial Intelligence, IC-FTAI 2024 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/IC-FTAI62324.2024.10950018.
- [3] D. Riquelme and M. A. Akhloufi, “Deep Learning for Lung Cancer Nodules Detection and Classification in CT Scans,” Dec. 01, 2020, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/ai1010003.
- [4] M. Tabibian, T. Razmpour, and R. Saha, “Diffusion Models vs. DCGANs for Class-Imbalanced Lung Cancer CT Classification: A Comparative Study,” Nov. 20, 2025. doi: 10.1101/2025.11.19.689331.
- [5] A. A. Abe and M. Nyathi, “Lung Cancer Diagnosis From Computed Tomography Images Using Deep Learning Algorithms With Random Pixel Swap Data Augmentation: Algorithm Development and Validation Study,” *JMIR Bioinform. Biotech.*, vol. 6, 2025, doi: 10.2196/68848.
- [6] alyasriy, hamdalla, AL-Huseiny, and Muayed, “The IQ-OTH/NCCD lung cancer dataset,” 2023. Accessed: Jan. 05, 2026. [Online]. Available: <https://www.kaggle.com/code/pythonafroz/vgg16-mobilenet-densenet-inception-resnet-nasne>
- [7] E. Goceri, “Medical image data augmentation: techniques, comparisons and interpretations,” *Artif. Intell. Rev.*, vol. 56, no. 11, pp. 12561–12605, Nov. 2023, doi: 10.1007/s10462-023-10453-z.
- [8] ชิตพงษ์ กิตตินราดร, “Convolutional Neural Network,” 2020. Accessed: Jan. 07, 2026. [Online]. Available: <https://guopai.github.io/ml-blog19.html>

- [9] Afroz, “VGG16|MobileNet|DenseNet|Inception|ResNet|NASNe,” 2025. Accessed: Jan. 07, 2026. [Online]. Available: <https://www.kaggle.com/code/pythonaafroz/vgg16-mobilenet-densenet-inception-resnet-nasne>
- [10] Shubham Saxena, “Fine-Tuning vs. Transfer Learning: An In-Depth Guide with Code Examples,” 2024. Accessed: Jan. 05, 2026. [Online]. Available: <https://medium.com/@shubhamsd100/fine-tuning-vs-transfer-learning-an-in-depth-guide-with-code-examples-46118e45d0cc>
- [11] Nisha Arya Ahmed, “What is A Confusion Matrix in Machine Learning? The Model Evaluation Tool Explained,” 2024. Accessed: Jan. 07, 2026. [Online]. Available: <https://www.datacamp.com/tutorial/what-is-a-confusion-matrix-in-machine-learning>
- [12] S. Joseph *et al.*, “GA-AGN: A generative adversarial network and attention gated network model for enhanced lung cancer detection using chest CT scans,” *EngMedicine*, vol. 2, no. 3, Sep. 2025, doi: 10.1016/j.engmed.2025.100077.
- [13] M. Abumohsen, E. Costa-Montenegro, S. Garcia-Mendez, A. Y. Owda, and M. Owda, “Advanced Deep Learning Techniques for Accurate Lung Cancer Detection and Classification,” in *Proceeding - 12th International Conference on Information Technology: Innovation Technologies, ICIT 2025*, Institute of Electrical and Electronics Engineers Inc., 2025, pp. 7–12. doi: 10.1109/ICIT64950.2025.11049178.
- [14] N. Tawfik, H. M. Emara, W. El-Shafai, N. F. Soliman, A. D. Algarni, and F. E. Abd El-Samie, “Enhancing Early Detection of Lung Cancer through Advanced Image Processing Techniques and Deep Learning Architectures for CT Scans,” *Computers, Materials and Continua*, vol. 81, no. 1, pp. 271–307, 2024, doi: 10.32604/cmc.2024.052404.
- [15] H. Zhu, G. Han, Y. Peng, W. Zhang, C. Lin, and H. Zhao, “Functional-realistic CT image super-resolution for early-stage pulmonary nodule detection,” *Future Generation Computer Systems*, vol. 115, pp. 475–485, Feb. 2021, doi: 10.1016/j.future.2020.09.020.
- [16] H. Xie *et al.*, “Super-resolution of Pneumocystis carinii pneumonia CT via self-attention GAN,” *Comput. Methods Programs Biomed.*, vol. 212, Nov. 2021, doi: 10.1016/j.cmpb.2021.106467.

- [17] N. Pongphaw and P. Buaphan, "Comparative evaluation and interpretability analysis of modern CNN architectures for brain tumor MRI classification," *Journal of Associated Medical Sciences*, vol. 59, no. 2, pp. 44–62, May 2026, doi: 10.12982/JAMS.2026.036.
- [18] R. Raza *et al.*, "Lung-EffNet: Lung cancer classification using EfficientNet from CT-scan images," *Eng. Appl. Artif. Intell.*, vol. 126, p. 106902, Nov. 2023, doi: 10.1016/J.ENGAPPAI.2023.106902.
- [19] P. Nasra, "Lung Cancer Classification using ResNet50," in *2nd International Conference on Intelligent Cyber Physical Systems and Internet of Things, ICoICI 2024 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 1201–1205. doi: 10.1109/ICoICI62503.2024.10696405.
- [20] R. Klangbunrueang, P. Pookduang, W. Chansanam, and T. Lunrasri, "AI-Powered Lung Cancer Detection: Assessing VGG16 and CNN Architectures for CT Scan Image Classification," *Informatics*, vol. 12, no. 1, Mar. 2025, doi: 10.3390/informatics12010018.
- [21] N. G. Pratiwi, Y. Nabila, R. Fiqraini, and A. W. Setiawan, "Effect of CT-Scan Image Resizing, Enhancement and Normalization on Accuracy of Covid-19 Detection," in *Proceedings - 2021 International Seminar on Intelligent Technology and Its Application: Intelligent Systems for the New Normal Era, ISITIA 2021*, Institute of Electrical and Electronics Engineers Inc., Jul. 2021, pp. 17–22. doi: 10.1109/ISITIA52817.2021.9502217.
- [22] M. Hassan, S. Hassan, A. Mujahid, M. Umair, and M. Zubair, "Lung Cancer Diagnosis and Classification Using Hybrid Deep Feature Extraction," in *Proceedings of 2025 4th International Conference on Computing and Information Technology, ICCIT 2025*, Institute of Electrical and Electronics Engineers Inc., 2025, pp. 256–261. doi: 10.1109/ICCIT63348.2025.10989327.
- [23] X. Xu *et al.*, "Explainable AI as a Double-Edged Sword in Dermatology: The Impact on Clinicians versus The Public."
- [24] N. S. Jozi and G. A. Al-Suhail, "LCxNet: An Explainable CNN Framework for Lung Cancer Detection in CT Images Using Multi-Optimizer and Visual Interpretability," *Applied System Innovation*, vol. 8, no. 5, Oct. 2025, doi: 10.3390/asi8050153.

- [25] A. Bhatt, K. Joshi, P. Jain, V. K. Gupta, and V. Vidisha, "A Deep Learning based Image Processing Technique for Early Lung Cancer Detection," in *2024 1st International Conference on Innovations in Communications, Electrical and Computer Engineering, ICICEC 2024*, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/ICICEC62498.2024.10808684.