Science and Engineering Connect, ISSN 3027-7914 (Online)
Vol. 49 No. 2 pp. 123-148 (2026)

บทความวิจัย

การปรับแต่งสถาปัตยกรรมชั้นเชื่อมโยงแบบสมบูรณ์บนโครงข่ายประสาทแบบคอนโวลูชัน เพื่อการจำแนกวัณโรคจากภาพรังสีทรวงอก

Fine-tuning Fully Connected Layer Architecture in Convolutional Neural Networks for Tuberculosis Classification of Chest Radiographs

พงษ์ศธร เชิดสม
Pongsathorn Chedsom
คณะสถาปัตยกรรมศาสตร์ มหาวิทยาลัยขอนแก่น ขอนแก่น ประเทศไทย
Faculty of Architecture, Khon Kaen University, Khon Kaen, Thailand
** Corresponding author E-mail: pongche@kku.ac.th*

Received 17 December 2025; Revised 11 May 2026; Accepted 29 May 2026

บทคัดย่อ

ความเป็นมาและวัตถุประสงค์ : วัณโรคเป็นสาเหตุการเสียชีวิตอันดับต้น ๆ ของประชากรโลก สูงถึง 1.5 ล้านรายต่อปี การวินิจฉัยเพื่อคัดกรองเบื้องต้นที่ทันสมัยเพื่อการจำแนกที่รวดเร็วและเชื่อถือได้มีความจำเป็นสำหรับรังสีแพทย์และผู้ป่วย แม้เทคโนโลยีโครงข่ายประสาทแบบคอนโวลูชัน (CNN) จะมีประสิทธิภาพสูง แต่แบบจำลองขนาดใหญ่จะมีปัญหาความซับซ้อนเชิงคำนวณที่ล้นเกิน ซึ่งอาจไม่เหมาะกับทรัพยากรทางเทคโนโลยีที่จำกัด งานวิจัยนี้จึงมุ่งเน้นการศึกษาเชิงทดลองเพื่อปรับแต่งชั้นเชื่อมโยงแบบสมบูรณ์ 15 รูปแบบเพื่อค้นหาโครงสร้างส่วนการจำแนกประเภทที่มีความสมดุลระหว่างความประหยัดทรัพยากร และประสิทธิภาพสูงสุด

วิธีดำเนินการวิจัย : พัฒนาแบบจำลอง 225 แบบจำลองจากการจับคู่สถาปัตยกรรมที่ได้รับความนิยมสำหรับการสกัดคุณลักษณะเด่นของข้อมูล 15 สถาปัตยกรรมร่วมกับชั้นการเชื่อมโยงแบบสมบูรณ์ สำหรับจำแนกข้อมูลที่ออกแบบใหม่จำนวน 15 รูปแบบ ใช้ชุดข้อมูล Tuberculosis Chest X-rays Images และ Tuberculosis (TB) Chest X-ray ประกอบไปด้วยภาพที่เป็นวัณโรคและภาพที่ไม่เป็นวัณโรคอย่างละ 3,194 ภาพ รวม 6,388 ภาพ แบ่งออกเป็น 70:20:10 สำหรับเรียนรู้ร้อยละ 70 จำนวน 4,471 ภาพ สำหรับการประเมินร้อยละ 20 จำนวน 1,278 ภาพ และสำหรับการทดสอบร้อยละ 10 จำนวน 639 ภาพ และประเมินประสิทธิภาพของแบบจำลองด้วยเทคนิคการทวนสอบแบบไขว้ 5 ชุด กำหนดไฮเปอร์พารามิเตอร์พื้นฐานในการทดลอง ได้แก่ Adam Optimizer (LR = 0.001), 15 Epochs และ Batch size 128

ผลการวิจัย : แบบจำลอง VGG16 ร่วมกับชั้นเชื่อมโยงแบบสมบูรณ์รูปแบบที่ 11 (FC Pattern 11) ให้ประสิทธิภาพสูงสุดในการจำแนกวัณโรค มีค่าความถูกต้องร้อยละ 98.58 และมีค่าเฉลี่ยความถูกต้องจากการทวนสอบแบบไขว้ร้อยละ 97.52 ส่วนเบี่ยงเบนมาตรฐาน 0.01 มีค่าความสูญเสียร้อยละ 0.15 มีค่าเฉลี่ย

ความถูกต้องจากการทวนสอบแบบไขว้ร้อยละ 0.16 ส่วนเบี่ยงเบนมาตรฐาน 0.02 และประเมินด้วยการวัดทางการแพทย์พบว่ามีค่าความไวร้อยละ 99.05 และความจำเพาะร้อยละ 98.11 มีโครงสร้างลดจำนวนโหนดตามลำดับขั้นได้แก่ (384 -> 192 -> 96) ร่วมกับ Batch Normalization ที่ช่วยให้โครงสร้างภายในเสถียรมากขึ้นและ L2 Regularization ควบคุมไม่ให้แบบจำลองซับซ้อนเกินไป และป้องกันปัญหา Overfitting แบบจำลองนี้มีโครงสร้างที่ไม่ซับซ้อน มีความลึก 20 ชั้น ช่วยลดภาระการคำนวณและป้องกันปัญหาความซับซ้อนที่ล้นเกินได้ดีกว่าแบบจำลองขนาดใหญ่ที่เน้นความประหยัดทรัพยากร

สรุป : การออกแบบชั้นเชื่อมโยงแบบสมบูรณ์สามารถเพิ่มประสิทธิภาพของสถาปัตยกรรมมาตรฐานให้สูงขึ้นในระดับที่แสดงให้เห็นถึงศักยภาพเชิงทฤษฎีในการพัฒนาต่อยอด เป็นระบบสนับสนุนการตัดสินใจทางคลินิก แบบจำลองที่พัฒนาขึ้นมีความเหมาะสมเชิงประสิทธิภาพทางเทคนิคในการวินิจฉัยคัดกรองเบื้องต้นโดยพิจารณาจากค่าความไวที่สูงถึงร้อยละ 99.05 ช่วยลดความเสี่ยงในการคัดกรองผู้ป่วยได้เป็นอย่างดี

การนำไปใช้ประโยชน์ในเชิงปฏิบัติ : แบบจำลองที่นำเสนอสามารถประยุกต์ใช้เป็นเครื่องมือคัดกรองวัณโรคเบื้องต้นเพื่อช่วยลดภาระงานของรังสีแพทย์ในการคัดกรองเคสปกติออกจากเคสที่สงสัยได้อย่างมีประสิทธิภาพ อีกทั้งแนวทางนี้จะช่วยเพิ่มความรวดเร็วในกระบวนการวินิจฉัย โดยเฉพาะในโรงพยาบาลที่มีทรัพยากรและผู้เชี่ยวชาญที่จำกัด เพื่อให้ผู้ป่วยได้รับการรักษาที่ทันท่วงทีอย่างมีประสิทธิภาพ

คำสำคัญ : การจำแนกวัณโรค, โครงข่ายประสาทเทียมแบบคอนโวลูชัน, ภาพเอกซเรย์ทรวงอก

Abstract

Background and Objectives: Tuberculosis remains a leading cause of worldwide mortality, accounting for approximately 1.5 million deaths annually. Rapid diagnosis is essential for increasing accessibility to timely treatment. Although Convolutional Neural Network (CNN) technologies exhibit high performance, large-scale models often encounter over-parameterization, which may be unsuitable in resource-limited settings. The present research presents an experimental study to fine-tune 15 variations of Fully Connected (FC) layer architectures to identify a classification head that achieves an optimal balance between resource efficiency (lightweight) and maximum performance (effectiveness).

Methodology: The models were developed by pairing 15 popular feature extraction architectures (backbones) with 15 newly designed FC layer patterns, resulting in a total of 225 model combinations. A balanced dataset comprising 6,388 chest X-ray images (3,194 Tuberculosis and 3,194 Normal) was utilized to mitigate model bias. The data was partitioned into a training set (70%, 4,471 images), a validation set (20%, 1,278 images), and a test set (10%, 639 images). Performance was evaluated using 5-fold cross-validation to assess

stability. Hyperparameters were set as follows: Adam optimizer (LR = 0.001), 15 epochs, and a batch size of 128.

Main Results: VGG16 architecture combined with "FC Pattern 11" achieved the highest performance, yielding an accuracy of 98.58%, a mean cross-validation accuracy of 97.52% (SD = 0.01), and a loss of 0.15. Clinical evaluation metrics demonstrated a sensitivity of 99.05% and a specificity of 98.11%. Pattern 11 employs a sequential node reduction strategy (384 -> 192 -> 96) integrated with Batch Normalization and L2 Regularization, which effectively improves internal stability and mitigates overfitting. Furthermore, this model maintains a lightweight structure with a depth of 16 layers, effectively reducing computational load and preventing over-parameterization more efficiently than deeper, more complex architectures.

Conclusions: Fine-tuning fully connected layers can significantly enhance standard architectures to a level suitable for Clinical Decision Support Systems (CDSS). The developed model is technically appropriate for preliminary screening, particularly due to its high sensitivity (99.05%), which minimizes the risk of missing pathological cases during initial triage.

Practical Application: The proposed model can be implemented as a preliminary screening tool to effectively reduce the workload of radiologists by filtering normal cases from suspected ones. The approach could help accelerate the diagnostic process, particularly in hospitals with limited resources and specialists, ensuring that patients receive timely and effective treatment.

Keywords: Tuberculosis Classification, Convolutional Neural Networks, Chest X-ray

Introduction

วัณโรค (Tuberculosis : TB) เป็นโรคติดเชื้อที่มีชีวิตผู้คนมากที่สุดในโลกและยังเป็นสาเหตุการเสียชีวิตอันดับต้น ๆ ซึ่งเกิดจากเชื้อแบคทีเรีย (Mycobacterium tuberculosis) และส่งผลกระทบต่อปอดมากที่สุด โดยการแพร่กระจายผ่านอากาศผ่านการ ไอ จาม หรือถ่มน้ำลาย จากข้อมูลรายงานขององค์การอนามัยโลกปี 2567 พบว่ามีผู้ป่วยเป็นวัณโรคถึง 10 ล้านรายและมีผู้เสียชีวิตจากวัณโรคถึง 1.5 ล้านรายต่อปี [1] ปัจจุบันประเทศไทยติด 1 ใน 14 ประเทศ ที่มีอัตราผู้ป่วยวัณโรคสูงที่สุดในโลก ในปี 2567 ประเทศไทยมีผู้ป่วยวัณโรคสูงกว่า 113,000 ราย และมีผู้เสียชีวิตกว่า 13,000 ราย และในปี 2568 จากข้อมูลสำนักงานป้องกันควบคุมโรคพบว่ามีผู้ป่วยวัณโรคจำนวน 11,604 คน อยู่ระหว่างการรักษา 9,954 ราย เสียชีวิต 687 ราย (ข้อมูลจากระบบ NTIP วันที่ 5 มีนาคม 2568) [2]

ภาพเอกซเรย์ทรวงอก (Chest X-ray) มีความสำคัญในวงการแพทย์ในการเป็นเครื่องมือตรวจวินิจฉัยเบื้องต้นและติดตามผลการรักษาอวัยวะภายในช่องอก เช่น ปอด หัวใจ กระดูกซี่โครง กระดูก สันหลัง และหลอดเลือด ช่วยให้แพทย์สามารถตรวจหาความผิดปกติ เช่น ปอดอักเสบ มะเร็งปอด หรือโรคเกี่ยวกับหัวใจ รวมถึงประเมินความรุนแรงของการบาดเจ็บและวางแผนการรักษาได้อย่างมีประสิทธิภาพ นอกจากจะมีความสะดวกและรวดเร็วในการตรวจ มีความปลอดภัยต่อผู้ป่วยเนื่องจากใช้ปริมาณรังสีน้อย เหมาะสำหรับการตรวจคัดกรองและส่งผลต่อความแม่นยำในการคัดกรอง แม้ภาพเอกซเรย์ทรวงอกจะเป็นเครื่องมือพื้นฐานที่สำคัญในการวินิจฉัยโรคแต่ยังมีอุปสรรคหลายอย่างที่ลดประสิทธิภาพในการรักษา เช่น การขาดแคลนผู้เชี่ยวชาญโดยเฉพาะในพื้นที่ชนบทที่มีอัตรากำลังไม่เพียงพอ ซึ่งจำนวนแพทย์ที่ไม่สอดคล้องกับปริมาณผู้ป่วยส่งผลให้กระบวนการวินิจฉัยล่าช้า ความซับซ้อนของการแปลผลที่ต้องอาศัยทักษะและประสบการณ์สูง ข้อจำกัดด้านทรัพยากรและเทคโนโลยีที่ไม่มีประสิทธิภาพในบางสถานพยาบาลเป็นอุปสรรคต่อความคมชัดของภาพและการเข้าถึงข้อมูลส่งผลให้การตัดสินใจทางคลินิกในภาวะฉุกเฉินทำได้ไม่เต็มประสิทธิภาพ

การนำเทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligence: AI) ประยุกต์ใช้โดยเฉพาะอย่างยิ่ง การเรียนรู้เชิงลึก (Deep Learning) เข้ามามีบทบาทสำคัญและได้รับการยอมรับอย่างกว้างขวางในวงการการแพทย์มากยิ่งขึ้นในปัจจุบัน โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Networks: CNN) ได้กลายเป็นสถาปัตยกรรมหลักที่มีประสิทธิภาพสูงในการวิเคราะห์ภาพถ่ายทางการแพทย์ เช่น ภาพถ่ายรังสีทรวงอกหรือภาพเอกซเรย์คอมพิวเตอร์ (CT Scan) ความสามารถของ CNN ที่มีประสิทธิภาพมากกว่าเทคนิคการเรียนรู้ของเครื่องแบบเดิม (Traditional Machine Learning) ที่มีความสามารถสกัดคุณลักษณะเด่นของข้อมูลจากภาพได้ (Feature Extraction) ตั้งแต่คุณลักษณะพื้นฐาน เช่น ขอบ เส้น และพื้นผิว ไปจนถึงคุณลักษณะที่มีความซับซ้อนมากยิ่งขึ้นเช่น รูปร่างของอวัยวะ หรือลักษณะเฉพาะของรอยโรคทำให้มีประสิทธิภาพในการคัดแยกโรค (Classification) การระบุตำแหน่งความผิดปกติ (Localization) และการจำแนกส่วนภาพ (Segmentation) ด้วยความสามารถในการประมวลผลข้อมูลปริมาณมหาศาลและการเรียนรู้รูปแบบ (Pattern) ที่ซับซ้อนได้ดีกว่าการสังเกตด้วยตาเปล่า จึงถูกนำมาประยุกต์ใช้ในทางการแพทย์มากยิ่งขึ้น

แม้เทคโนโลยีการเรียนรู้เชิงลึกจะมีความก้าวหน้าอย่างต่อเนื่อง ในปัจจุบันงานวิจัยมุ่งเน้นให้ความสำคัญกับการเพิ่มประสิทธิภาพของแบบจำลองโดยปรับแต่งในส่วนของโครงสร้างสถาปัตยกรรมหลัก (Backbone) เป็นหลัก ส่งผลให้มีความซับซ้อนเกินความจำเป็น ซึ่งการนำแบบจำลองที่มีสถาปัตยกรรมขนาดใหญ่มาใช้งานจะเจอกับปัญหาเรื่องความซับซ้อนเชิงคำนวณที่มากเกินไป (Over-parameterization) และอาจจะไม่สอดคล้องกับลักษณะเฉพาะของชุดข้อมูลที่มีความจำกัด เช่น ข้อมูลภาพรังสีทรวงอกในการวินิจฉัยวัณโรค งานวิจัยนี้จึงมุ่งเน้นไปที่การเลือกสถาปัตยกรรมชั้นเชื่อมโยงแบบสมบูรณ์ (Fully Connected Layer) ที่สามารถให้ความสำคัญระหว่างการประหยัดทรัพยากรในการคำนวณ (Lightweight) และประสิทธิภาพสูงสุด (Effectiveness) เพื่อแก้ปัญหาความซับซ้อนเชิงคำนวณที่มากเกินไปในการคัดกรองวัณโรคสำหรับพื้นที่ที่มีทรัพยากรจำกัด โดยการปรับแต่งสถาปัตยกรรมชั้นเชื่อมโยงแบบสมบูรณ์ 15 รูปแบบ เพื่อให้ได้โครงสร้างส่วนการจำแนกประเภท (Classification Head) ที่เหมาะสมที่สุดในข้อจำกัดของสถาปัตยกรรมมาตรฐานที่ได้

รับความนิยมในปัจจุบันที่มีความแตกต่างกัน รวมทั้งมีความซับซ้อนและการเลือกเทคนิคควบคุมพารามิเตอร์ เช่น การใช้ Dropout, Batch Normalization และ L2 Regularization เพื่อให้ได้โครงสร้างส่วนการจำแนกประเภทที่สามารถจำแนกภาพรังสีทรวงอกได้อย่างมีประสิทธิภาพสูงสุด และเพื่อเปรียบเทียบประสิทธิภาพการจำแนกวัณโรค โดยการจับคู่ระหว่างสถาปัตยกรรมโครงข่ายประสาทเทียมจำนวน 15 สถาปัตยกรรมกับรูปแบบชั้นเชื่อมโยงที่ออกแบบขึ้นรวมทั้งหมด 225 รูปแบบ/แบบจำลอง และใช้การประเมินประสิทธิภาพของแบบจำลองด้วยดัชนีชี้วัดที่มีความสำคัญกับการจำแนกภาพรังสีทรวงอก [4-5, 7-9] ที่บ่งชี้วัณโรค ได้แก่ ค่าความไว (Sensitivity) และค่าความจำเพาะ (Specificity) ที่เป็นดัชนีมาตรฐานที่ใช้ในการประเมินความสามารถของแบบจำลอง และมีการนำเทคนิคการทวนสอบแบบไขว้ 5 ชุด (5-Fold Cross Validation) อีกทั้งชุดข้อมูลที่มีความสมดุล (Balanced Dataset) จำนวน 6,388 ภาพ เพื่อให้สามารถนำไปประยุกต์ใช้ในการสนับสนุนการตัดสินใจของแพทย์ในสถานการณ์จริงได้อย่างมีประสิทธิภาพ

Literature Review

การคัดแยกวัณโรคด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชันจากภาพถ่ายรังสีทรวงอก (Chest X-ray) ของ Wajgi และคณะ [3] มีการประยุกต์ใช้เทคนิคการถ่ายโอนการเรียนรู้ (Transfer Learning) และการปรับแต่งไฮเปอร์พารามิเตอร์ (Hyperparameter Tuning) เพื่อเพิ่มประสิทธิภาพสูงสุดของแบบจำลอง จากชุดข้อมูล Tuberculosis (TB) Chest X-ray Database ประกอบด้วยภาพถ่ายรวม 4,200 ภาพ แบ่งเป็นภาพผู้ป่วยปกติ 3,500 ภาพ และผู้ป่วยที่ติดเชื้อวัณโรค 700 ภาพ ได้ใช้แบบจำลองจำลองของ VGG พบว่าการปรับแต่งที่มีค่าความถูกต้องมากที่สุดคือ Activation function Tanh, Optimizer SGD, Dropout ร้อยละ 30 จำนวน 50 รอบการเรียนรู้ ให้ค่าความถูกต้อง (accuracy) สูงถึงร้อยละ 98.11

การพัฒนาแบบจำลองการเรียนรู้เชิงลึกสำหรับการจำแนกวัณโรคของ Rim และคณะ [4] โดยใช้โครงสร้างสถาปัตยกรรมหลัก ได้แก่ EfficientNet และโครงสร้างส่วนการจำแนกประเภท คือ MLP-Mixer และใช้เทคนิคการถ่ายโอนการเรียนรู้ (Transfer Learning) ใช้ชุดข้อมูลเฉพาะ (SCH TB dataset) ของโรงพยาบาลชอนอันซุนชอนฮยาง (Cheonan Soonchunhyang Hospital) ประกอบด้วยภาพถ่ายรวม 3,828 ภาพ แบ่งเป็นภาพผู้ป่วยปกติ (Inactive TB) 2,115 ภาพ และผู้ป่วยที่ติดเชื้อวัณโรค (Active TB) 1,713 ภาพ และกำหนดพารามิเตอร์ได้แก่ Batch size 32 จำนวน 300 รอบการเรียนรู้ ให้ค่าความถูกต้อง (accuracy) สูงถึงร้อยละ 96.30

การพัฒนาแบบจำลองเพื่อตรวจหาวัณโรคของ Mirugwe และคณะ [5] โดยใช้แบบจำลอง 6 แบบจำลองได้แก่ VGG16, VGG19, ResNet50, ResNet101, ResNet152, และ Inception-ResNet-V2 ในการจำแนกภาพถ่ายรังสีทรวงอก (CXR) และมีพารามิเตอร์ Activation ReLU, Optimizer Adam, Batch size 16, จำนวน 100 รอบการเรียนรู้ ใช้ชุดข้อมูล Tuberculosis (TB) Chest X-ray Database ประกอบด้วยภาพถ่ายรวม 4,200 ภาพ แบ่งเป็นภาพผู้ป่วยปกติ 3,500 ภาพ และผู้ป่วยที่ติดเชื้อวัณโรค 700 ภาพ และได้เพิ่มจำนวนชุดข้อมูลด้วยเทคนิค Data Augmentation พบว่าแบบจำลองที่ให้ค่าความถูกต้องมากที่สุดคือ

แบบจำลอง VGG16 ร้อยละ 99.40 และแบบจำลองที่มีจำนวนพารามิเตอร์สูงและซับซ้อน เช่น ResNet152 และ Inception-ResNet-V2 ต้องใช้เวลาในการฝึกฝนที่ยาวนานกว่า โดยที่ไม่ได้ให้ประสิทธิภาพที่ดีขึ้นตามสัดส่วน นอกจากนี้ยังพบว่าการใช้เทคนิค Data Augmentation ไม่ได้ช่วยให้ประสิทธิภาพของแบบจำลองเพิ่มมากขึ้น การพัฒนาแบบจำลองในการจำแนกวัณโรคโดยใช้ภาพถ่ายรังสีทรวงอก (Chest X-ray: CXR) และใช้เทคนิคการเรียนรู้เชิงลึกจากโครงข่ายประสาทเทียมแบบคอนโวลูชัน (CNNs) ของ Haque และคณะ [6] ใช้ชุดข้อมูล Tuberculosis (TB) Chest X-ray Database จำนวน 4,200 ภาพ แบ่งเป็นภาพผู้ป่วยปกติ 3,500 ภาพ และผู้ป่วยที่ติดเชื้อวัณโรค 700 ภาพ ใช้เทคนิคการเรียนรู้ถ่ายโอน (Transfer Learning) ร่วมกับการปรับจูนพารามิเตอร์ (Hyperparameter Tuning) โดยใช้สถาปัตยกรรมที่ฝึกฝนล่วงหน้าคือ VGG19 จำนวน 50 รอบการเรียนรู้ ผลการทดลองพบว่าให้ค่าความถูกต้องร้อยละ 98.00

การพัฒนาแบบจำลองของ Fati และคณะ [7] ได้นำเสนอการวินิจฉัยวัณโรคในระยะเริ่มต้นจากภาพถ่ายรังสีทรวงอก (Chest X-ray) 2 แนวทางจากชุดข้อมูล 2 ชุดข้อมูล Shenzhen Dataset จำนวน 662 ภาพ และ Qatar University/Dhaka University Dataset จำนวน 4,200 ภาพ ภาพ ได้แก่ 1 การใช้แบบจำลอง CNN ร่วมกันได้แก่ ResNet-50 และ GoogLeNet ในการสกัดคุณลักษณะของข้อมูลและใช้ SVM ในการจำแนกประเภทเป็นวัณโรคหรือไม่เป็นวัณโรค และแนวทางที่ 2 ใช้โครงข่ายประสาทเทียม (ANN) ในการจำแนกประเภทเป็นวัณโรคหรือไม่เป็นวัณโรค จากการสกัดคุณลักษณะของข้อมูลจาก ResNet-50 และ GoogLeNet โดยใช้เทคนิคการสกัดคุณลักษณะของข้อมูลได้แก่ Gray Level Co-occurrence Matrix (GLCM), Discrete Wavelet Transform (DWT) และ Local Binary Pattern (LBP) พบว่าแนวทางที่ 2 ให้ค่าความถูกต้องร้อยละ 99.20 ในชุดข้อมูลที่ 1 และให้ค่าความถูกต้องร้อยละ 99.80 ในชุดข้อมูลที่ 2 (จากชุดข้อมูล Traning set)

การนำเสนอการพัฒนาแบบจำลองโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network: CNN) เพื่อจำแนกวัณโรคของ Genitha และคณะ [8] ในภาพถ่ายรังสีทรวงอก (Chest X-ray) จากชุดข้อมูล TB X-Ray แบ่งข้อมูลออกเป็นร้อยละ 80 สำหรับการเรียนรู้ของแบบจำลองและร้อยละ 20 สำหรับการตรวจสอบและทดสอบแบบจำลอง โดยมีเป้าหมายเพื่อเพิ่มความแม่นยำและความรวดเร็วในการวินิจฉัยแทนที่วิธีการแบบเดิมที่ใช้เวลานานและมีค่าใช้จ่ายสูง จากสถาปัตยกรรม ResNet ในการสกัดคุณลักษณะข้อมูลและจำแนกวัณโรคแบ่งออกเป็น TB-positive เป็นวัณโรคและ TB-negative ไม่เป็นวัณโรคพบว่าการใช้งาน CNN ให้ค่าความถูกต้องร้อยละ 92.00 และเปรียบเทียบกับสถาปัตยกรรม VGG พบว่า ให้ค่าความถูกต้องร้อยละ 90.85

การพัฒนาแบบจำลองโครงข่ายประสาทเทียมแบบไฮบริดเพื่อตรวจหาวัณโรคของ Hossain และคณะ [9] ได้นำเสนอแนวทางการรวมจุดแข็งของสถาปัตยกรรม CNN 3 รูปแบบ ได้แก่ VGG16, ResNet50 และ DenseNet121 เข้าด้วยกันผ่านเทคนิคการรวมคุณลักษณะ (Feature-level fusion) เพื่อเพิ่มความน่าเชื่อถือในการวินิจฉัยจากภาพถ่ายรังสีทรวงอก (CXR) โดยทดสอบกับชุดข้อมูลรวมจาก Montgomery County และโรงพยาบาล Shenzhen และชุดข้อมูลเพิ่มเติมรวมกว่า 4,300 ภาพ ผลการทดลองพบว่าแบบจำลอง

ไฮบริดให้ประสิทธิภาพสูงกว่าแบบจำลองเดี่ยวอย่างมีนัยสำคัญ โดยมีค่าความถูกต้อง (Accuracy) ร้อยละ 97.40 และมีค่าความแม่นยำ (Precision), ความครบถ้วน (Recall) และ F1-score อยู่ที่ร้อยละ 96 เท่ากันทั้งสามค่า ซึ่งเหมาะสำหรับการนำไปใช้เป็นระบบสนับสนุนการตัดสินใจทางคลินิกที่มีทรัพยากรทางเทคโนโลยีและงบประมาณที่จำกัด (Limited Technological Resources) และบุคลากรผู้เชี่ยวชาญที่มีความจำกัด (Shortage of Specialists)

จากการวิเคราะห์เปรียบเทียบระเบียบวิธีวิจัยใน Table 1 พบว่างานวิจัยที่เกี่ยวข้องส่วนใหญ่มุ่งเน้นการใช้สถาปัตยกรรมมาตรฐาน เช่น VGG16, ResNet50 และ DenseNet121 โดยเน้นการถ่ายโอนการเรียนรู้ (Transfer Learning) ในส่วนของสถาปัตยกรรมหลักเป็นหลักเพียงอย่างเดียว ซึ่งการใช้พารามิเตอร์พื้นฐานในสถาปัตยกรรมขนาดใหญ่เกิดความซับซ้อนเชิงคำนวณที่มากเกินไป งานวิจัยนี้คือการมุ่งเน้นออกแบบสถาปัตยกรรมชั้นเชื่อมโยงแบบสมบูรณ์ 15 รูปแบบ เพื่อหาโครงสร้างส่วนการจำแนกประเภทที่มีความสมดุลระหว่างความประหยัดทรัพยากร (Lightweight) และประสิทธิภาพสูงสุด ในการทดลองได้นำชุดข้อมูลภาพรังสีทรวงอกรวม 6,388 ภาพจาก Kaggle [11] และ Mendeley [10] มารวมเป็นชุดข้อมูลที่มีความสมดุล (Balanced Dataset) กลุ่มละ 3,194 ภาพ เพื่อลดอคติของแบบจำลองซึ่งเป็นปัจจัยสำคัญที่ทำให้ผลการทดลองมีความน่าเชื่อถือสูงกว่าการใช้ชุดข้อมูลที่ไม่สมดุล (Imbalanced Dataset) และมีการประเมินประสิทธิภาพได้แก่ Accuracy, Precision, Recall, Specificity, F1-score ให้เหมาะสมกับแบบจำลองในการจำแนกภาพรังสีทรวงอกและสอดคล้องกับ [4-5, 7-9] รวมทั้งการใช้เทคนิค 5-Fold Cross Validation เพื่อประเมินความถูกต้องและความน่าเชื่อถือของแบบจำลอง

Table 1 Comparison of Experimental Methodologies and Performance Results

Ref	Dataset	Backbones	Technique	Matrix	Accuracy
[3]	Tuberculosis (TB) Chest X - ray Database 4,200 images (Imbalanced)	VGG 19	Hyperparameter Tuning (Tanh, SGD)	Accuracy, ROC-AUC	98.11
[4]	SCH TB dataset 3,828 images (Specific set)	EfficientNet + MLP-Mixer	MLP-Mixer Classification Head	Accuracy, Sensitivity, Specificity	96.30
[5]	Tuberculosis (TB) Chest X - ray Database 4,200 images	6 Models (VGG, ResNet, Inception)	Data Augmentation	Precision, Recall, F1-score , AUC	99.40
[6]	Tuberculosis (TB) Chest X - ray Database 4,200 images	VGG19	Transfer Learning, Layer Freezing	Accuracy, ROC-AUC	98.00

Table 1 Comparison of Experimental Methodologies and Performance Results (continued)

Ref	Dataset	Backbones	Technique	Matrix	Accuracy
[7]	Shenzhen dataset 662 images + TB Chest Radiography Database 4,200 images	ResNet-50, GoogLeNet	Feature Fusion (GLCM, DWT, LBP)	Accuracy, Sensitivity, Specificity, AUC	99.80
[8]	Kaggle TB X-Ray dataset (Tuberculosis (TB) Chest X-ray Database) 4,200 images	ResNet, VGG	Deep Learning Algorithm / Preprocessing (ROI extraction)	Accuracy, Precision, Sensitivity, Specificity, F1-score	92.00
[9]	Montgomery County Chest X-ray dataset 4,300 images	Hybrid (VGG16, ResNet50, DenseNet121)	Feature-level Fusion	Accuracy, Precision, Recall, F1-score	97.40
Proposed Method	Tuberculosis Chest X-rays Images [10]+ Tuberculosis (TB) Chest X-ray Database [11] (6,388 images Balanced)	15 Architecture - EfficientNet B0-B7 - VGG16-19 - ResNet50 - ResNet101 - ResNet152 - Inception-ResNet-V2 - GoogLeNet	Fine-tune Fully Connected Layer (FC) (15 Architecture X FC 15 Patterns = 225 Experimental Models, 5-Fold Cross Validation	Accuracy, Loss, Precision, Recall (Sensitivity), Specificity, F1	98.58

Materials and Methods

เครื่องมือ

ในการศึกษานี้แบบจำลองในการจำแนกวัคโรคถูกพัฒนาขึ้นบนภาษา Python โดยใช้ TensorFlow และ Keras เป็นเฟรมเวิร์คหลักในการออกแบบและฝึกสอนโครงข่ายประสาทเทียม ฮาร์ดแวร์ประกอบด้วยหน่วยประมวลผลกลาง Intel Core i9-13900KS หน่วยความจำหลัก (RAM) ขนาด 32 GB หน่วยประมวลผลกราฟิก Nvidia RTX 4070 SUPER (7,168 CUDA cores, Ram 12 GB) อัตราการส่งข้อมูล (Memory bandwidth) อยู่ที่ 504 GB/s และใช้ระบบปฏิบัติการ Ubuntu 22.04 LTS

ชุดข้อมูลและการเตรียมข้อมูล

ชุดข้อมูลที่ใช้ในการทดลองคือ Dataset of Tuberculosis Chest X-rays Images [10] เป็นชุดข้อมูลภาพ chest X-ray ประกอบด้วยข้อมูลจำนวน 3,008 ภาพ แบ่งออกเป็นวัณโรค (Tuberculosis) 2,494 ภาพ และไม่เป็นวัณโรค (Normal) 514 ภาพ ร่วมกับชุดข้อมูล Tuberculosis (TB) Chest X-ray Database [11] เป็นชุดข้อมูลภาพ chest X-ray ประกอบด้วยข้อมูลจำนวน 4,200 ภาพ แบ่งออกเป็นวัณโรค 700 ภาพ และไม่เป็นวัณโรค 3,500 ภาพ นำชุดข้อมูลทั้ง 2 ชุดข้อมูลมากันรวม ได้ภาพที่เป็นวัณโรค 3,149 ภาพ และไม่เป็นวัณโรค 4,014 ภาพ ดังแสดงภาพตัวอย่างของชุดข้อมูลใน Figure 1 และเพื่อให้ทั้งกลุ่มข้อมูล (Tuberculosis/Normal) มีปริมาณที่เท่ากัน (Balanced dataset) จึงเลือกชุดข้อมูลละ 3,149 ภาพ มาใช้ในการทดลองนี้ โดยใช้กลุ่มข้อมูลที่น้อยที่สุดคือ 3,149 ภาพ ตามเกณฑ์กลุ่มข้อมูลที่น้อยที่สุด และใช้วิธีการสุ่มเลือกภาพ (Random) รวมทั้งไม่ใช้เทคนิคการเพิ่มข้อมูล (Data Augmentation) เพื่อต้องการประเมินประสิทธิภาพของแบบจำลองจากชุดข้อมูลต้นฉบับโดยตรงและเพื่อให้ผลการทดลองสะท้อนถึงความสามารถในการเรียนรู้ภายใต้สภาพแวดล้อมที่กำหนดอย่างชัดเจน

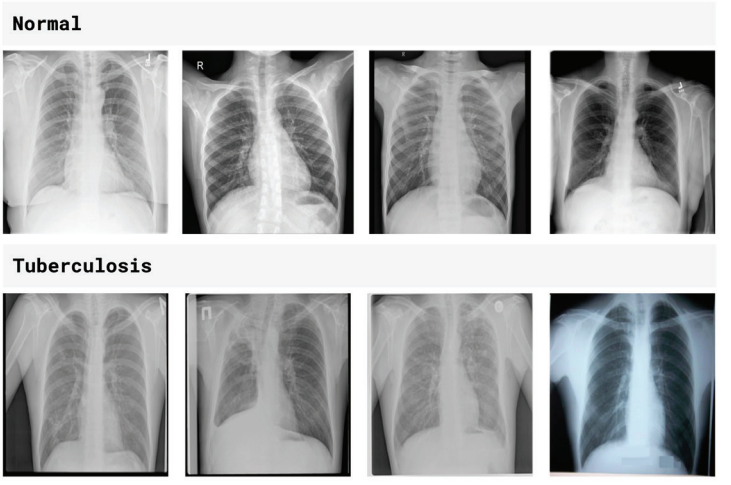

Figure 1 Example X-ray Images of Dataset

การเตรียมชุดข้อมูลในการทดลองได้แบ่งชุดข้อมูลออกเป็น 70:20:10 เป็นกลุ่มภาพวัณโรค (Tuberculosis) และกลุ่มภาพที่ไม่เป็นวัคโรค (Normal) กลุ่มละ 3,149 ภาพ รวมเป็น 6,388 ภาพ แบ่งข้อมูลสำหรับเรียนรู้ (Training set) ร้อยละ 70 คิดเป็น 4,470 ภาพ (กลุ่มละ 2,235 ภาพ) ชุดสำหรับการประเมิน (Validation set) ร้อยละ 20 คิดเป็น 1,276 ภาพ (กลุ่มละ 638 ภาพ) และชุดข้อมูลสำหรับการทดสอบ (Test set) ร้อยละ 10 คิดเป็น 640 ภาพ (กลุ่มละ 320 ภาพ) ดังแสดงรายละเอียดใน Figure 2 และในการทดลองได้นำเทคนิคการทวนสอบแบบไขว้ (5-Fold Cross-Validation) มาใช้เพื่อประเมินความเสถียร (Stability) และความทนทาน (Robustness) ของแบบจำลอง ซึ่งช่วยลดความลำเอียง (Bias) ที่อาจเกิดขึ้นจากการแบ่งชุดข้อมูลเพียงครั้งเดียว และเพื่อให้มั่นใจว่าข้อมูลในแต่ละส่วน (Fold) มีการกระจายตัวอย่างสม่ำเสมอ ทำให้การประเมินประสิทธิภาพของแบบจำลองที่พัฒนาขึ้นมีความแม่นยำและน่าเชื่อถือมากขึ้น

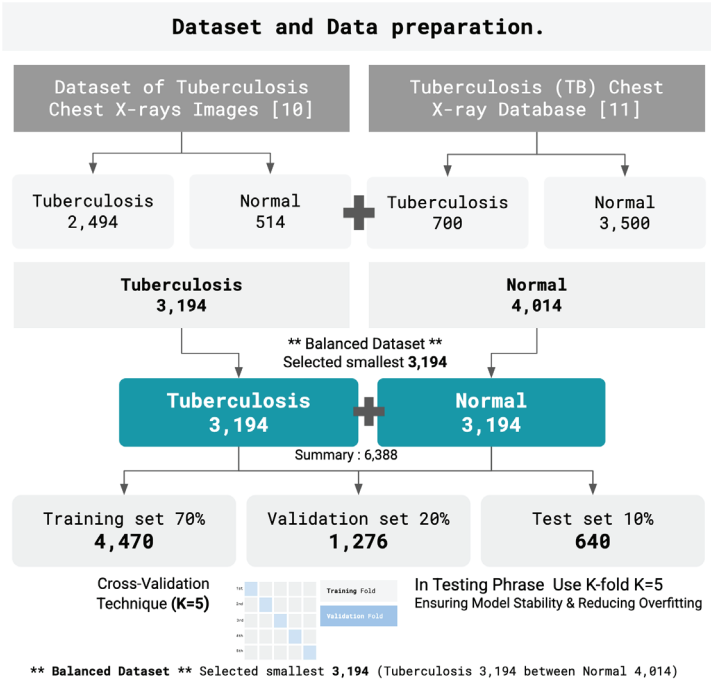

Figure 2 Splitting Dataset for Model Learning

การประเมินประสิทธิภาพแบบจำลอง

การวัดค่าความถูกต้อง (Accuracy)

เป็นการวัดความถูกต้องในการจำแนกข้อมูลของแบบจำลอง โดยเป็นสัดส่วนของผลลัพธ์ที่แบบจำลองทำนายได้ถูกต้องทั้งหมด (ทั้ง TP และ TN) หารด้วยจำนวนตัวอย่างทั้งหมดในชุดข้อมูลที่นำมาทดสอบ [9] โดยใช้การคำนวณดังนี้

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

โดยที่

True Positive (TP) จำนวนตัวอย่างที่เป็นจริง (Positive) และถูกแบบจำลองทำนายว่าเป็นจริง (Positive) ได้ถูกต้อง

True Negative (TN) จำนวนตัวอย่างที่เป็นไม่จริง (Negative) และถูกแบบจำลองทำนายว่าเป็นไม่จริง (Negative) ได้ถูกต้อง

False Positive (FP) จำนวนตัวอย่างที่เป็นไม่จริง (Negative) แต่ถูกแบบจำลองทำนายว่าเป็นจริง (Positive) ไม่ถูกต้อง

False Negative (FN) จำนวนตัวอย่างที่เป็นจริง (Positive) แต่ถูกแบบจำลองทำนายว่าเป็นไม่จริง (Negative) ไม่ถูกต้อง

การวัดค่าความแม่นยำ (Precision)

เป็นการวัดสัดส่วนของจำนวนตัวอย่างที่แบบจำลองทำนายว่าเป็นวัณโรคได้ถูกต้อง (TP) ต่อจำนวนตัวอย่างที่แบบจำลองทำนายว่าเป็นวัณโรคทั้งหมดทั้งที่ถูกและผิด (TP + FP) [9] โดยใช้การคำนวณดังนี้

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

การวัดค่าความครบถ้วน (Sensitivity หรือ Recall)

เป็นการวัดสัดส่วนของจำนวนตัวอย่างที่แบบจำลองทำนายว่าเป็นวัณโรคได้ถูกต้อง (TP) ต่อจำนวนตัวอย่างที่เป็นวัณโรคจริงทั้งหมดในชุดข้อมูลทั้งที่ทายถูกและทายผิดว่าปกติ (TP + FN) [9] โดยใช้การคำนวณดังนี้

$$\text{Sensitivity (Recall)} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

การวัดค่าความความจำเพาะ (Specificity)

เป็นการวัดสัดส่วนของจำนวนตัวอย่างที่แบบจำลองทำนายว่าปกติได้ถูกต้อง (TN) ต่อจำนวนตัวอย่างที่เป็นคนปกติจริงทั้งหมดทั้งที่ทายถูกและทายผิดว่าเป็นโรค (TN + FP) [4] โดยใช้การคำนวณดังนี้

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{4}$$

การวัดค่าเอฟวันสคอร์ (F1-Score)

เป็นตัววัดประสิทธิภาพที่ได้จากการคำนวณค่าเฉลี่ยแบบฮาร์มอนิก (Harmonic Mean) ระหว่างค่าความแม่นยำ (Precision) และค่าความครบถ้วน (Recall) เพื่อให้มีความสมดุลระหว่างการลดจำนวนผลบวกปลอม (FP) และผลลบปลอม (FN) [9] โดยใช้การคำนวณดังนี้

$$F1\text{-Score} = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{5}$$

การวัดความสูญเสีย (Loss Function)

เป็นฟังก์ชันนี้ที่ตัวชี้วัดประสิทธิภาพของแบบจำลองและเป็นกลไกหลักในกระบวนเรียนรู้ของแบบจำลอง (Training) โดยใช้ในการวัดความคลาดเคลื่อน (error) ของแบบจำลองการเรียนรู้ของเครื่อง (Machine Learning Model) เพื่อให้แบบจำลองสามารถปรับพารามิเตอร์ (Parameters) เพื่อลดค่าการสูญเสียลงให้ต่ำที่สุด โดยค่าการสูญเสียต่ำจะหมายความว่าแบบจำลองทำงานได้แม่นยำและถูกต้องมากขึ้น ในการจำแนกข้อมูลแบบ 2 กลุ่ม (Binary Classification) จะนิยมนำ Binary Cross-Entropy Loss [9] มาใช้โดยใช้การคำนวณดังนี้

$$\text{Binary Loss} = -\frac{1}{n}\sum_{i=1}^{n}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)] \tag{6}$$

โดยที่

n คือ จำนวนข้อมูลทั้งหมด (Sample Size)

i คือ ลำดับของข้อมูล (Index)

y_i คือ ค่าจริง (Actual Value)
$\hat{y}_i$ คือ ค่าที่ทำนายได้ (Predicted Value)
$\sum$ คือ ผลรวม

การออกแบบการทดลอง (Design of Experiment)

การทดลองได้ใช้ชุดข้อมูล Chest X-Ray Dataset [10] และ Tuberculosis (TB) Chest X-ray Database [11] เป็นชุดข้อมูลสาธารณะที่ได้รับการยอมรับและถูกนำไปใช้อย่างกว้างขวางในงานวิจัยระดับโลกที่สามารถนำเปรียบเทียบประสิทธิภาพของแบบจำลองกับงานวิจัยอื่นได้ เช่น [3, 5-8] (Benchmarking) เป็นชุดข้อมูลภาพ chest X-ray จำนวน 6,388 ภาพ แบ่งออกเป็น 2 กลุ่มภาพที่เป็นวัณโรค (Tuberculosis) จำนวน 3,194 ภาพ และภาพที่ไม่เป็นวัณโรค (Normal) จำนวน 3,194 ภาพ แบ่งชุดข้อมูลออกเป็น 70:20:10 เพื่อใช้สำหรับการเรียนรู้ของแบบจำลองและได้กำหนดไฮเปอร์พารามิเตอร์ (Hyper parameter) พื้นฐานในการทดลองทุกการทดลองในวิจัยนี้ ได้แก่ Optimizer = Adam, Learning rate = 0.001 (ค่าพื้นฐานของ Optimizer Adam) Epoch = 15 และ Batch size = 128 โดยใช้สถาปัตยกรรมจำนวน 15 สถาปัตยกรรม (Architecture) ในการสกัดคุณลักษณะของข้อมูล (Feature Extraction) ร่วมกับการออกแบบชั้นเชื่อมโยงแบบสมบูรณ์ที่แตกต่างกันจำนวน 15 รูปแบบ (Pattern) เพื่อใช้ในกระบวนการจำแนกข้อมูลแบบสองกลุ่ม (Binary Classification) ได้แก่กลุ่มภาพวัณโรคและกลุ่มภาพปกติ รวมทั้งสิ้น 225 แบบจำลอง ซึ่งได้แสดงเปรียบเทียบขั้นตอนกับโครงสร้างของโครงข่ายประสาทแบบคอนโวลูชันดังแสดงใน Figure 3 และได้แสดงถึงขั้นตอนภาพรวมในกระบวนการพัฒนาแบบจำลองดังแสดงใน Figure 4

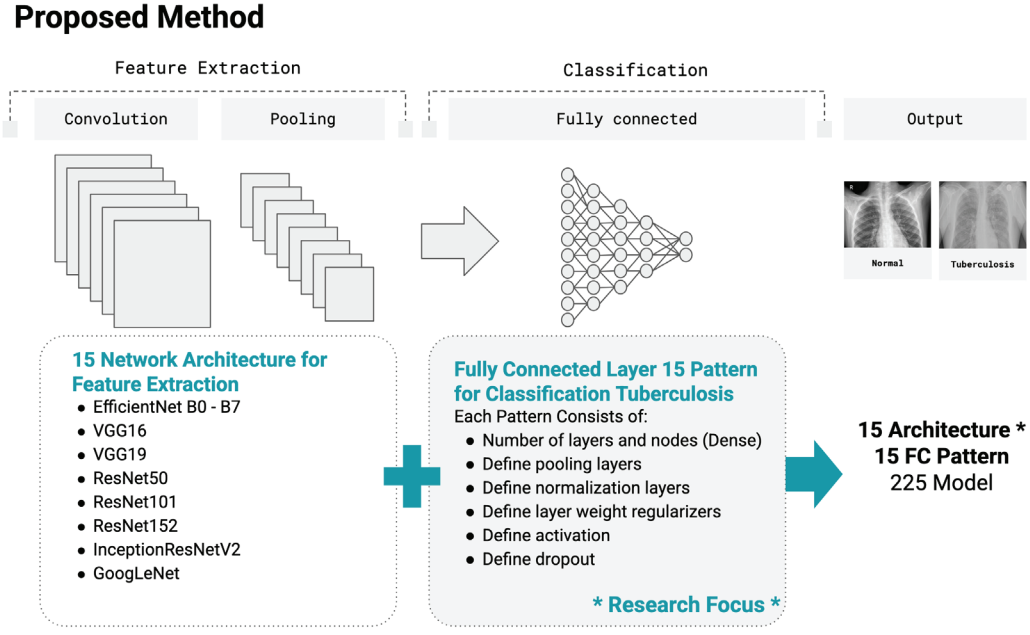

Figure 3 Proposed Methodologies

การทดลองได้นำสถาปัตยกรรมมาตรฐานที่ได้รับความนิยมจำนวน 15 สถาปัตยกรรม ที่มีความแตกต่างของโครงสร้าง เช่น จำนวนชั้น (Layer) จุดเด่น/ความสามารถ จำนวนพารามิเตอร์ (Parameter) ขนาด (Size) ของแต่ละของแต่ละสถาปัตยกรรม เพื่อให้มีเงื่อนไข ข้อจำกัดและมีความหลากหลายของสถาปัตยกรรมที่แตกต่างกัน รวมทั้งคุณลักษณะสำคัญของสถาปัตยกรรมหลัก ดังแสดงใน Table 2 มาใช้ในการสกัดคุณลักษณะเด่นของข้อมูลได้แก่

- EfficientNet B0-B7 [4]
- VGG16 [3] [5] [8]
- VGG19 [3] [6] [5] [8]
- ResNet50 [5] [7] [8]
- ResNet101 [5] [8]
- ResNet152 [5] [8]
- Inception-ResNet-V2 [5]
- GoogLeNet [7]

กลุ่มสถาปัตยกรรม EfficientNet (B0 - B7) เป็นแบบจำลองที่มีความยืดหยุ่นสูงเนื่องจากใช้ Compound Scaling ในการปรับสมดุลความกว้าง ความลึก และความละเอียดของภาพ ในรุ่น B0 - B3 มีจำนวนพารามิเตอร์ 5.3 - 12.3 ล้านพารามิเตอร์และขนาดไฟล์เล็ก (29 - 48 MB) ในรุ่น B4 - B7 ที่มีความลึกมากถึง 438 ชั้น และมีพารามิเตอร์ 66.7 ล้านพารามิเตอร์

กลุ่มสถาปัตยกรรม VGG (VGG16 และ VGG19) โครงสร้างไม่ซับซ้อนมีฟิลเตอร์ (Filter) ขนาด 3x3 ตลอดทั้งโครงสร้าง ง่ายต่อการปรับแต่งโครงสร้าง ในมิติของพารามิเตอร์มีจำนวนสูงถึง 138.4 - 143.7 ล้านพารามิเตอร์และมีขนาดไฟล์ใหญ่เกิน 500 MB

กลุ่มสถาปัตยกรรม ResNet (50, 101, 152) ออกแบบมาเพื่อแก้ปัญหาการหายไปของเกรเดียนต์ (Vanishing Gradient) โดยใช้เทคนิค Shortcut Connections ทำให้มีชั้นความลึกมากถึง 311 ชั้น ใน ResNet152 มีชั้นความลึก 311 ชั้น มีพารามิเตอร์สูงถึง 60.4 ล้านพารามิเตอร์ ซึ่งมีความซับซ้อนอาจก่อให้เกิดสภาวะความซับซ้อนเชิงคำนวณที่มากเกินไปได้

สถาปัตยกรรม Inception-ResNet-V2 เป็นแบบจำลองแบบไฮบริดที่รวมจุดแข็งของ Inception (การสกัดคุณลักษณะหลายระดับ) และ Residual connections เข้าด้วยกัน เพื่อให้สามารถเรียนรู้ลักษณะเด่นที่ซับซ้อนได้ดียิ่งขึ้นและมีโครงสร้างที่ซับซ้อนโดยมีความลึก 449 ชั้น มีพารามิเตอร์ 55.9 ล้านพารามิเตอร์ ส่งผลให้เป็นแบบจำลองมีความต้องการทรัพยากรระบบสูงมากและอาจเกิดปัญหาความซับซ้อนล้นเกินได้

สถาปัตยกรรม GoogLeNet มีจำนวนพารามิเตอร์ 6.8 ล้านพารามิเตอร์และขนาดไฟล์ที่ 96 MB เมื่อเทียบกับความสามารถในการสกัดคุณลักษณะโครงสร้างนี้จึงเหมาะสมกับสภาพแวดล้อมที่มีทรัพยากรจำกัด แต่มีความลึก 22 ชั้น ซึ่งอาจเป็นข้อจำกัดในการสกัดคุณลักษณะเชิงลึกที่ซับซ้อนได้ไม่ดี

Table 2 Comparison of Neural Network Architectures

Group	Architecture	Size (MB)	Depth	Parameter (M)
EfficientNet	EfficientNet B0	29	132	5.3
	EfficientNet B1	31	186	7.9
	EfficientNet B2	36	186	9.2
	EfficientNet B3	48	210	12.3
	EfficientNet B4	75	258	19.5
	EfficientNet B5	118	312	30.6
	EfficientNet B6	166	360	43.3
	EfficientNet B7	256	438	66.7
VGG	VGG16	528	16	138.4
	VGG19	549	19	143.7
ResNet	ResNet50	98	107	25.6
	ResNet101	171	209	44.7
	ResNet152	232	311	60.4
Inception	Inception-ResNet-V2	215	449	55.9
GoogLeNet	GoogLeNet	96	22	6.8

โดยการคัดเลือกสถาปัตยกรรมเน้นไปที่ความประหยัดทรัพยากร (Lightweight) และ ประสิทธิภาพสูงสุด (Effectiveness) จากโครงสร้างไม่ซับซ้อนและมีพารามิเตอร์ต่ำ เช่น VGG16 มีพารามิเตอร์ 14.7 ล้านพารามิเตอร์ สถาปัตยกรรมขนาดใหญ่ เช่น ResNet152 มีจำนวนพารามิเตอร์ 58.3 ล้านพารามิเตอร์ หรือ Inception-ResNet-V2 มีจำนวนพารามิเตอร์ 54.3 ล้านพารามิเตอร์ และสถาปัตยกรรมที่มีความลึกและซับซ้อนสูง (Deep and Complex Architectures) เพื่อทดสอบความสามารถในการสกัดคุณลักษณะเด่นที่ซับซ้อนและทดสอบปัญหาความซับซ้อนเชิงคำนวณที่เกินความจำเป็น สำหรับชุดข้อมูลที่มีจำกัดจำนวน 6,388 ภาพ อีกทั้งเป็นสถาปัตยกรรมมาตรฐานสากลที่ใช้ในการเปรียบเทียบประสิทธิภาพ (Benchmarking) ทำให้ผลการทดลองสามารถนำไปอ้างอิงและเทียบเคียงได้อย่าง และออกแบบชั้นเชื่อมโยงแบบสมบูรณ์ จำนวน 15 รูปแบบ ดังแสดงใน Figure 5 ซึ่งแต่ละรูปแบบอย่างน้อยประกอบไปด้วยจำนวนชั้น (Layer) จำนวนโหนด และอื่น ๆ ได้แก่กำหนด Pooling Layers, Normalization Layers, Weight Regularizers Layers, Activation Layers และกำหนด Dropout Layers ที่เป็นส่วนประกอบพื้นฐานของแต่ละชั้นของการเชื่อมโยงแบบสมบูรณ์ ได้ออกแบบโดยให้ความสำคัญกับความสัมพันธ์ระหว่างความลึกของโครงสร้างและจำนวนของพารามิเตอร์ที่เหมาะสมกับชุดข้อมูลภาพรังสีทรวงอก อีกทั้งให้มีความหลากหลายเชิงโครงสร้างสามารถเป็นตัวแทนของระดับความซับซ้อนที่แตกต่างกัน ตั้งแต่โครงสร้างแบบ Lightweight ที่มีจำนวนโหนดน้อยไปจนถึงที่มีความซับซ้อนสูง ซึ่งจำนวน 15 รูปแบบเพียงพอต่อการวิเคราะห์เชิงเปรียบเทียบร่วมกับสถาปัตยกรรมหลักทั้ง 15

สถาปัตยกรรม เพื่อให้ได้โครงสร้างส่วนการจำแนกประเภทที่ดีที่สุดสำหรับการใช้งานจริง

ในการปรับแต่งชั้นการเชื่อมโยงแบบสมบูรณ์ได้ออกแบบเพื่อทดสอบสมมติฐานทางเทคนิคในแต่ละรูปแบบ ได้แก่การควบคุมความซับซ้อนของข้อมูลมีการกำหนดจำนวนชั้น (Layer) และจำนวนโหนด (Dense units) ตั้งแต่ขนาด 16 ถึง 648 โหนดที่แตกต่างกัน และได้นำเทคนิค Dropout (0.25 - 0.6), L2 Regularization มาใช้งานเพื่อป้องกันการเรียนรู้ที่จดจำเฉพาะกับข้อมูลตัวอย่างมากเกินไป(Overfitting) ได้นำเทคนิค และ Batch Normalization เพื่อสร้างความเสถียรในการปรับค่าพารามิเตอร์ (Stability) นอกจากนี้ยังนำเทคนิคเฉพาะมาใช้ในการทดลอง ได้แก่ SpatialDropout1D, Orthogonal Regularizer, GaussianDropout เพื่อทดสอบความแม่นยำของผลลัพธ์และหาโครงสร้างที่เหมาะสมที่สุดกับชุดข้อมูลที่ใช้ในการทดลอง

ในการวัดประสิทธิภาพของแบบจำลองที่พัฒนาขึ้นได้ใช้เทคนิคต่าง ๆ ได้แก่ การวัดค่าความถูกต้อง (Accuracy) การวัดค่าความแม่นยำ (Precision) การวัดค่าความครบถ้วน (Recall หรือ Sensitivity) การวัดค่าความความจำเพาะ (Specificity) การวัดค่าเอฟวันสคอร์ (F1-Score) การวัดความสูญเสีย (Loss Function) รวมทั้งการนำเทคนิคการทดสอบแบบไขว้ (K-Fold Cross Validation) ในขั้นตอนการทดสอบแบบจำลอง (Testing model) ซึ่งจะนำมารายงานผลในรูปแบบค่าผลเฉลี่ย (Arithmetic Mean) ทั้ง 5-fold และค่าเบี่ยงเบนมาตรฐาน (Standard Deviation: SD)

เทคนิคการทดสอบแบบไขว้ (K-Fold Cross Validation) เป็นเทคนิคที่ในการประเมินความถูกต้องและความน่าเชื่อถือของแบบจำลอง โดยมีจุดประสงค์หลักคือการลดอคติ (Bias) และการเกิด Overfitting ต่อชุดข้อมูลใดชุดหนึ่ง โดยจะทำการแบ่งข้อมูลทั้งหมดออกเป็นชุดย่อย (Folds) ที่มีขนาดเท่า ๆ กัน [8] โดยทั่วไปมักใช้ K=5 หรือจำนวนคี่ในแต่ละรอบของการประเมิน กระบวนการนี้จะทำซ้ำจนกว่าทุกชุดข้อมูลจะถูกใช้เป็นชุดประเมินผลครบทุกรอบซึ่งการกระจายตัวทำให้ข้อมูลทุกชุดมีโอกาสเป็นทั้งชุดการเรียนรู้ของแบบจำลองและชุดประเมินผลทำให้ค่าความถูกต้องของแบบจำลองมีความเสถียรและน่าเชื่อถือ

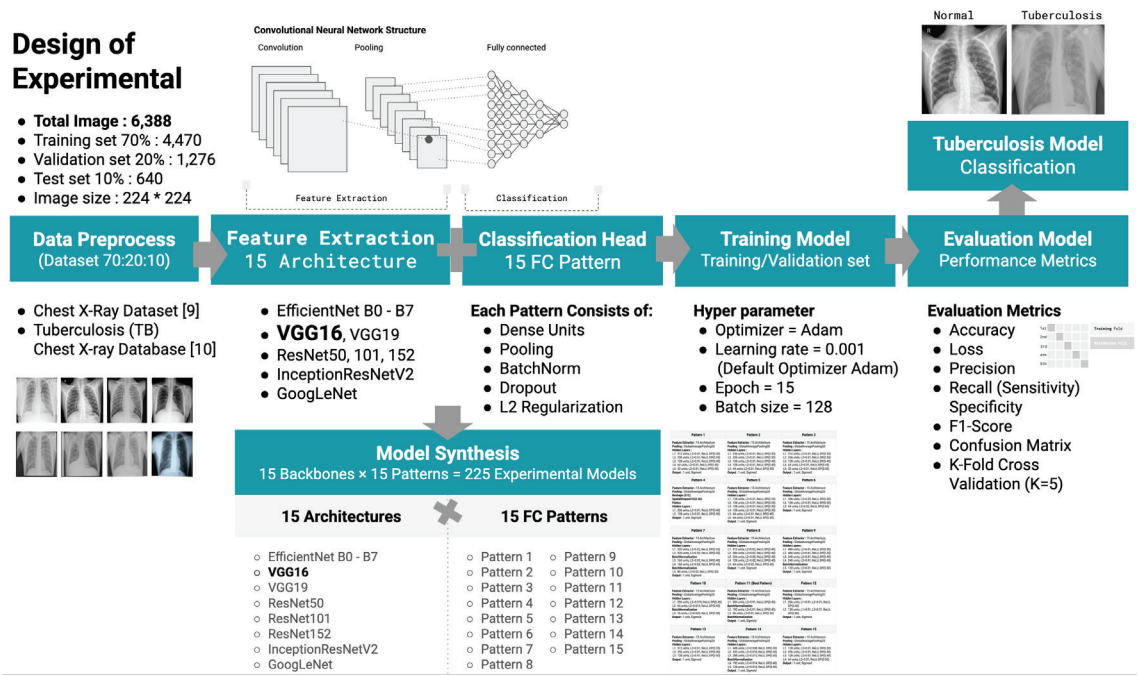

Figure 4 Model Development Steps

กระบวนการก่อนการเรียนรู้ของแบบจำลองในการจำแนกวัณโรค โดยใช้โครงข่ายประสาทแบบคอนโวลูชันจากภาพเอกซเรย์ทรวงอก มีขั้นตอนดังนี้

1. การรวบรวมและคัดกรองข้อมูล (Data Collection & Cleaning)

• รวบข้อมูลข้อมูลจากชุดข้อมูลต้นฉบับ (Tuberculosis Chest X-rays Images และ Tuberculosis (TB) Chest X-ray)

• จัดเตรียมข้อมูลให้อยู่ในรูปแบบการทำงานของฟังก์ชั่น ImageDataGenerator และ flow_from_directory ของ Keras โดยแบ่งเป็นโฟลเดอร์ (Main-Folder) Tuberculosis และ Normal ในแต่ละโฟลเดอร์ประกอบไปด้วยโฟลเดอร์ (Sub-Folder) Training_Set, Validation_Set, Test_Set

2. การปรับขนาดและรูปแบบภาพ (Resizing) เป็นการย่อขนาดภาพให้เท่ากันขนาด 224 × 224 เพื่อให้สอดคล้องกับ Input layer ของแบบจำลองที่นำมาใช้ในการทดลอง

3. การปรับค่ามาตรฐาน (Normalization) ให้เป็นแบบ One-hot Encoding ให้พิกเซล (Pixel) ของภาพให้มีค่าอยู่ระหว่าง 0 และ 1 ซึ่งเดิมภาพจะมีค่าระหว่าง 0 และ 255 ในแต่ละพิกเซล

4. การแบ่งชุดข้อมูล (Data Splitting)

• Training Set (70%) สำหรับการเรียนรู้ของแบบจำลอง

• Validation Set (20%) สำหรับการตรวจสอบและปรับจูนค่า parameters ในระหว่างกระบวนการเรียนรู้ของแบบจำลอง

• Test Set (10%) สำหรับทดสอบประสิทธิภาพของแบบจำลอง

• แบ่งชุดข้อมูลสำหรับการทดสอบประสิทธิภาพของแบบจำลองจากเทคนิค K-Fold Cross Validation จำนวน 5 ชุดข้อมูล (K=5)

5. การเรียนรู้ของแบบจำลองโดยใช้พารามิเตอร์ที่กำหนดและในทุกการทดลองกำหนดรอบการเรียนรู้ 15 รอบ (Epoch)

Pattern 1	Pattern 2	Pattern 3
Feature Extractor : 15 Architecture **Pooling :** GlobalAveragePooling2D **Hidden Layers :** L1. 512 units, L2=0.01, ReLU, DP(0.30) L2. 256 units, L2=0.01, ReLU, DP(0.35) L3. 128 units, L2=0.01, ReLU, DP(0.40) L4. 64 units, L2=0.01, ReLU, DP(0.45) L5. 32 units, L2=0.01, ReLU, DP(0.50) **Output :** 1 unit, Sigmoid	**Feature Extractor :** 15 Architecture **Pooling :** GlobalAveragePooling2D **Hidden Layers :** L1. 256 units, L2=0.01, ReLU, DP(0.30) L2. 256 units, L2=0.01, ReLU, DP(0.30) L3. 128 units, L2=0.01, ReLU, DP(0.40) L4. 128 units, L2=0.01, ReLU, DP(0.40) L5. 64 units, L2=0.01, ReLU, DP(0.50) **Output :** 1 unit, Sigmoid	**Feature Extractor :** 15 Architecture **Pooling :** GlobalAveragePooling2D **Hidden Layers :** L1. 512 units, L2=0.01, ReLU, DP(0.30) L2. 256 units, L2=0.01, ReLU, DP(0.35) L3. 128 units, L2=0.01, ReLU, DP(0.40) L4. 64 units, L2=0.01, ReLU, DP(0.45) L5. 32 units, L2=0.01, ReLU, DP(0.50) **Output :** 1 unit, Sigmoid
Pattern 4	**Pattern 5**	**Pattern 6**
Feature Extractor : 15 Architecture **Pooling :** GlobalAveragePooling2D **Reshape (512)** **SpatialDropout1D(0.30)** **Flatten** **Hidden Layers :** L1. 256 units, L2=0.01, ReLU, DP(0.40) L2. 128 units, L2=0.01, ReLU, DP(0.50) **Output :** 1 unit, Sigmoid	**Feature Extractor :** 15 Architecture **Pooling :** GlobalAveragePooling2D **Hidden Layers :** L1. 128 units, L2=0.01, ReLU, DP(0.25) L2. 128 units, L2=0.01, ReLU, DP(0.25) L3. 128 units, L2=0.01, ReLU, DP(0.30) L4. 128 units, L2=0.01, ReLU, DP(0.30) L5. 64 units, L2=0.01, ReLU, DP(0.40) L6. 64 units, L2=0.01, ReLU, DP(0.50) **Output :** 1 unit, Sigmoid	**Feature Extractor :** 15 Architecture **Pooling :** GlobalAveragePooling2D **Hidden Layers :** L1. 256 units, L2=0.02, ReLU, DP(0.50) L2. 128 units, L2=0.01, ReLU, DP(0.50) L3. 64 units, L2=0.02, ReLU, DP(0.60) **Output :** 1 unit, Sigmoid
Pattern 7	**Pattern 8**	**Pattern 9**
Feature Extractor : 15 Architecture **Pooling :** GlobalAveragePooling2D **Hidden Layers :** L1. 320 units, L2=0.02, ReLU, DP(0.35) L2. 320 units, L2=0.02, ReLU, DP(0.35) **BatchNormalization** L3. 160 units, L2=0.02, ReLU, DP(0.40) L4. 160 units, L2=0.02, ReLU, DP(0.45) **BatchNormalization** L5. 80 units, L2=0.02, ReLU, DP(0.50) **Output :** 1 unit, Sigmoid	**Feature Extractor :** 15 Architecture **Pooling :** GlobalAveragePooling2D **Hidden Layers :** L1. 512 units, L2=0.02, ReLU, DP(0.40) L2. 384 units, L2=0.02, ReLU, DP(0.40) L3. 256 units, L2=0.02, ReLU, DP(0.40) L4. 128 units, L2=0.02, ReLU, DP(0.40) L5. 64 units, L2=0.02, ReLU, DP(0.40) **Output :** 1 unit, Sigmoid	**Feature Extractor :** 15 Architecture **Pooling :** GlobalAveragePooling2D **Hidden Layers :** L1. 480 units, L2=0.01, ReLU, DP(0.30) L2. 480 units, L2=0.01, ReLU, DP(0.35) L3. 240 units, L2=0.01, ReLU, DP(0.40) L4. 240 units, L2=0.01, ReLU, DP(0.45) **BatchNormalization** L5. 120 units, L2=0.01, ReLU, DP(0.50) **Output :** 1 unit, Sigmoid
Pattern 10	**Pattern 11 (Best Pattern)**	**Pattern 12**
Feature Extractor : 15 Architecture **Pooling :** GlobalAveragePooling2D **Hidden Layers :** L1. 256 units, L2=0.010, ReLU, DP(0.40) L2. 64 units, L2=0.015, ReLU, DP(0.50) **BatchNormalization** L3. 16 units, L2=0.020, ReLU, DP(0.50) **Output :** 1 unit, Sigmoid	**Feature Extractor :** 15 Architecture **Pooling :** GlobalAveragePooling2D **Hidden Layers :** L1. 384 units, L2=0.01, ReLU, DP(0.40) **BatchNormalization** L2. 192 units, L2=0.01, ReLU, DP(0.40) L3. 96 units, L2=0.01, ReLU, DP(0.50) **BatchNormalization** **Output :** 1 unit, Sigmoid	**Feature Extractor :** 15 Architecture **Pooling :** GlobalAveragePooling2D **Hidden Layers :** L1. 256 units, L1=0.01, L2=0.01, ReLU, DP(0.40) L2. 128 units, L1=0.01, L2=0.01, ReLU, DP(0.40) **Output :** 1 unit, Sigmoid
Pattern 13	**Pattern 14**	**Pattern 15**
Feature Extractor : 15 Architecture **Pooling :** GlobalAveragePooling2D **Hidden Layers :** L1. 512 units, L2=0.01, ReLU, DP(0.35) L2. 256 units, L2=0.01, ReLU, DP(0.40) L3. 128 units, L2=0.01, ReLU, DP(0.50) **Output :** 1 unit, Sigmoid	**Feature Extractor :** 15 Architecture **Pooling :** GlobalAveragePooling2D **Hidden Layers :** L1. 648 units, L2=0.008, ReLU, DP(0.30) L2. 432 units, L2=0.010, ReLU, DP(0.30) L3. 288 units, L2=0.012, ReLU, DP(0.40) **BatchNormalization** L4. 192 units, L2=0.014, ReLU, DP(0.40) L5. 128 units, L2=0.016, ReLU, DP(0.50) **Output :** 1 unit, Sigmoid	**Feature Extractor :** 15 Architecture **Pooling :** GlobalAveragePooling2D **Hidden Layers :** L1. 128 units, L2=0.01, ReLU, DP(0.30) L2. 256 units, L2=0.01, ReLU, DP(0.40) L3. 128 units, L2=0.01, ReLU, DP(0.40) L4. 64 units, L2=0.01, ReLU, DP(0.50) **Output :** 1 unit, Sigmoid

Figure 5 Design of Fully Connected Layers

Results and Discussion

จากการพัฒนาและการเรียนรู้ของแบบจำลองพบว่าสถาปัตยกรรม VGG16 ร่วมกับรูปแบบที่ 11 ให้ประสิทธิภาพดีที่สุด มีค่าความถูกต้องร้อยละ 98.58 เมื่อทดสอบด้วยเทคนิค 5-Fold Cross Validation พบว่ามีค่าเฉลี่ยความถูกต้อง (K=5) ร้อยละ 97.52 ส่วนเบี่ยงเบนมาตรฐาน 0.01 และมีค่าความสูญเสีย (Loss)

ร้อยละ 0.15 มีค่าเฉลี่ยความสูญเสีย (K=5) 0.16 ส่วนเบี่ยงเบนมาตรฐาน 0.02 แสดงโครงสร้างใน Figure 5 ส่วนของ Pattern 11 และ Figure 6 ซึ่งสะท้อนถึงประสิทธิภาพและความเสถียรของแบบจำลองที่มีค่าความถูกต้องสูงกว่าร้อยละ 95 ในการทดสอบจำแนกวัณโรคในตางราง Confusion Matrix ประเมินประสิทธิภาพของแบบจำลองในการจำแนกประเภทแบบสองกลุ่ม (Binary Classification) โดยแสดงผลเปรียบเทียบระหว่างค่าจริง (Actual) ในแนวแถวและค่าพยากรณ์ (Predicted) ในแนวคอลัมน์ เพื่อระบุความถูกต้องและความผิดพลาดของแบบจำลองในรูปแบบตารางสองมิติ ได้ถูกต้องจำนวน 316 จำแนกผิดจำนวน 3 ภาพจากทั้งหมด 319 ภาพ ส่วนจำแนกที่ไม่ใช่วัณโรค (Normal) ได้ถูกต้องจำนวน 313 จำแนกผิดจำนวน 6 ภาพจากทั้งหมด 319 ภาพ ในการวัดค่าความแม่นยำ (Precision) มีค่าร้อยละ 98.13 แสดงว่าแบบจำลองมีความผิดพลาดแบบผลบวกปลอม (False Positive) หรือ การทำนายเคสปกติว่าเป็นโรคเพียง 6 เคสเท่านั้นซึ่งสามารถช่วยลดความกังวลและภาระในการตรวจวินิจฉัยซ้ำที่เกินความจำเป็นสำหรับเคสปกติได้มากขึ้น การวัดค่าความความจำเพาะ (Specificity) มีค่าร้อยละ 98.11 แสดงให้เห็นว่าแบบจำลองจะไม่ทำนายผิดโดยนำเคสที่ไม่เป็นโรคมาทำนายว่าเป็นโรค ซึ่งช่วยสร้างความมั่นใจให้กับแพทย์ในการตรวจมากยิ่งขึ้น นอกจากนั้นยังช่วยลดผลกระทบทางจิตใจแก่ผู้ป่วยที่มารับการตรวจอีกด้วย และในการวัดค่าความครบถ้วน (Recall) หรือในทางการแพทย์เรียกว่าค่า Sensitivity มีค่าร้อยละ 99.05 แสดงให้เห็นว่าแบบจำลองมีความแม่นยำที่จะไม่ทำนายผิดพลาดจากเคสที่เป็นโรคแต่ถูกทำนายว่าไม่เป็นโรค ซึ่งมีความสำคัญเป็นอย่างมากในการป้องกันไม่ให้ผู้ที่เป็นโรคหลุดจากการคัดกรองโดยสามารถยืนยันว่าแบบจำลองมีประสิทธิภาพในการเป็นเครื่องมือคัดกรองวัณโรคเบื้องต้น ซึ่งแบบจำลองนี้จะช่วยลดภาระงานของรังสีแพทย์ในการคัดกรองเคสปกติออกและเพิ่มความรวดเร็วในการวินิจฉัยคัดกรองเบื้องต้นผู้ป่วยวัณโรคในโรงพยาบาลที่มีทรัพยากรจำกัด ดังแสดงConfusion Matrix ใน Figure 7 และแสดงกราฟแสดงค่าความถูกต้องและค่าความสูญเสียในระหว่างการเรียนรู้ของแบบจำลองใน Figure 8

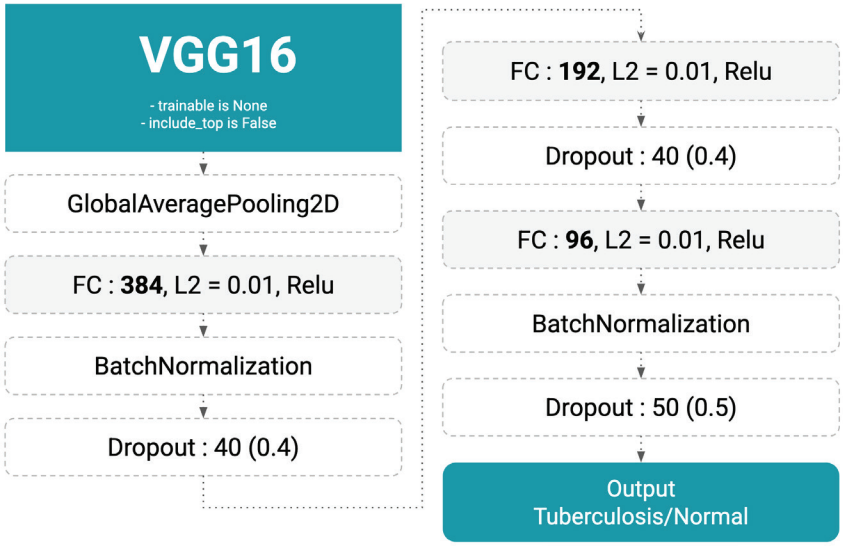

Figure 6 Best Model Architecture (VGG16 + FC Pattern 11)

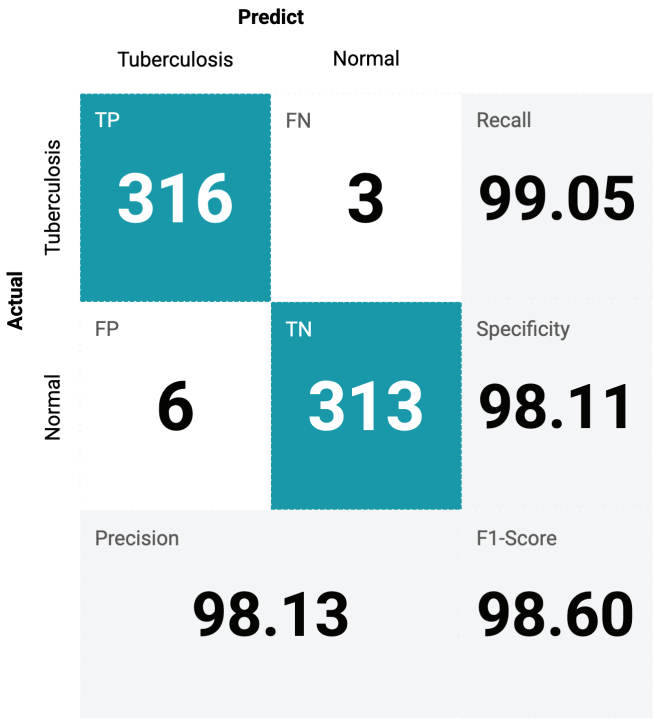

Figure 7 Confusion Matrix Best Model (VGG16 + FC Pattern 11)

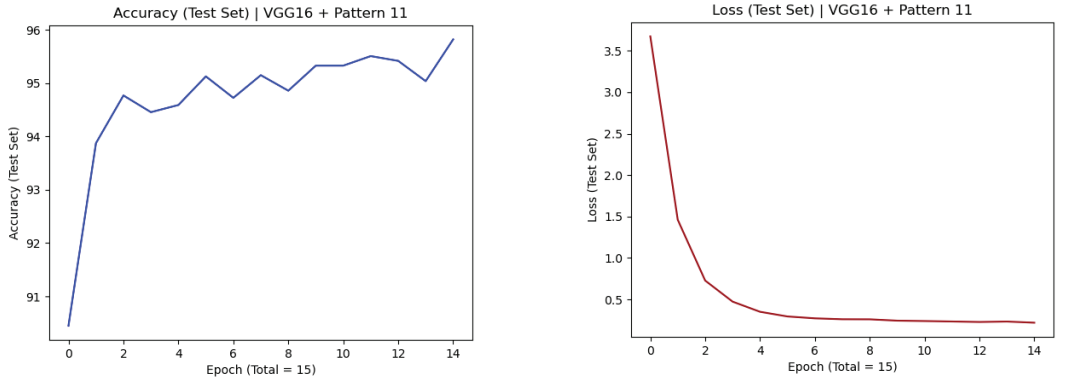

Figure 8 Accuracy/Loss Best Model (VGG16 + FC Pattern 11)

ผลการทดลองแบ่งตามกลุ่มสถาปัตยกรรมออกเป็น 5 กลุ่ม จำนวน 15 สถาปัตยกรรม ดังนี้

กลุ่มสถาปัตยกรรม EfficientNet ร่วมกับ 15 รูปแบบ ในชั้นการเชื่อมโยงแบบสมบูรณ์พบว่าเป็นกลุ่มที่ประสิทธิภาพน้อยที่สุดไม่สามารถจำแนกชุดข้อมูลที่นำมาใช้ในการทดลองได้ทุกสถาปัตยกรรม EfficientNet B0 – B7 ให้ค่าความถูกต้องร้อยละ 50.00 ทุกแบบจำลอง

กลุ่มสถาปัตยกรรม VGG ได้แก่ VGG16 VGG19 ร่วมกับ 15 รูปแบบ ในชั้นการเชื่อมโยงแบบสมบูรณ์พบว่าเป็นกลุ่มที่ให้ประสิทธิภาพมากที่สุดในการจำแนกวัณโรค คือแบบจำลอง VGG16 ร่วมกับรูปแบบที่ 11

ให้ประสิทธิภาพดีที่สุดมีค่าความถูกต้องคือร้อยละ 98.58 และมีค่าความสูญเสียร้อยละ 0.15 และ VGG19 ร่วมกับรูปแบบที่ 4 ให้ประสิทธิภาพดีที่สุดมีค่าความถูกต้องคือร้อยละ 97.80 และมีค่าความสูญเสียร้อยละ 0.25 แบบจำลองที่ให้ค่าความถูกต้องน้อยที่สุดมีจำนวน 13 แบบจำลองมีค่าความถูกต้องคือร้อยละ 50.00 โดยแสดงผลการทดลองที่ให้ค่าความถูกต้องมากที่สุด 5 ลำดับดังแสดงรายละเอียดในตาราง Table 3 ลำดับที่ 1 - 5

กลุ่มสถาปัตยกรรม ResNet ได้แก่ ResNet50 ResNet101 ResNet152 ร่วมกับ 15 รูปแบบในชั้นการเชื่อมโยงแบบสมบูรณ์พบว่าสถาปัตยกรรม ResNet101 ร่วมกับรูปแบบที่ 11 ให้ประสิทธิภาพดีที่สุดมีค่าความถูกต้องคือร้อยละ 73.04 และมีค่าความสูญเสียร้อยละ 0.70 เป็นกลุ่มที่ไม่สามารถจำแนกชุดข้อมูลที่นำมาใช้ในการทดลองได้ทุกสถาปัตยกรรมมีค่าความถูกต้องร้อยละ 50.00 ทุกแบบจำลอง

สถาปัตยกรรม Inception-ResNet-v2 ร่วมกับ 15 รูปแบบ ในชั้นการเชื่อมโยงแบบสมบูรณ์พบว่าสถาปัตยกรรม Inception-ResNet-v2 ร่วมกับรูปแบบที่ 15 ให้ประสิทธิภาพดีที่สุดมีค่าความถูกต้องคือร้อยละ 96.70 และมีค่าความสูญเสียร้อยละ 0.21 อีกทั้งมีแบบจำลองที่ไม่สามารถจำแนกภาพได้จำนวน 3 แบบจำลองค่าความถูกต้องร้อยละ 50.00 ได้แสดงผลการทดลองที่ให้ค่าความถูกต้องมากที่สุด 5 ลำดับดังแสดงรายละเอียดในตาราง Table 3 ลำดับที่ 6 - 10

สถาปัตยกรรม GoogLeNet ร่วมกับ 15 รูปแบบ ในชั้นการเชื่อมโยงแบบสมบูรณ์พบว่าสถาปัตยกรรม GoogLeNet ร่วมกับรูปแบบที่ 12 ให้ประสิทธิภาพดีที่สุดมีค่าความถูกต้องคือร้อยละ 57.99 และมีค่าความสูญเสียร้อยละ 1.20 เป็นกลุ่มที่ไม่สามารถจำแนกชุดข้อมูลที่นำมาใช้ในการทดลองได้ทุกสถาปัตยกรรมมีค่าความถูกต้องร้อยละ 50.00 ทุกแบบจำลอง

เปรียบเทียบแบบจำลองแต่ละสถาปัตยกรรมที่ได้ความถูกต้องมากที่สุดดังแสดงใน Table 4 จำนวน 15 สถาปัตยกรรมสามารถแบ่งออกเป็น 2 กลุ่ม กลุ่มที่สามารถจำแนกวัณโรคได้ ได้แก่ แบบจำลองในกลุ่ม VGG16, VGG19, Inception-ResNet-v2, ResNet101, ResNet152 มีค่าความถูกต้องในช่วงร้อยละ 64.26 - 99.00 แสดงให้เห็นว่าโครงสร้างของแบบจำลองสอดคล้องกับข้อมูลที่สามารถนำไปใช้งานได้และกลุ่มที่ไม่สามารถจำแนกได้เลยได้แก่ EfficientNet B0 - B1, ResNet50, GoogLeNet มีค่าความถูกต้องในช่วงร้อยละ 50.00 - 58.00 เช่น แสดงให้เห็นถึงการนำสถาปัตยกรรมที่มีขนาดใหญ่และมีโครงสร้างที่ซับซ้อนของโครงข่ายมากเกินไป มีจำนวนพารามิเตอร์เกินความจำเป็น ซึ่งไม่สอดคล้องกับลักษณะเฉพาะของชุดข้อมูลภาพรังสีทรวงอกที่มีความจำกัดและชั้นการเชื่อมโยงแบบสมบูรณ์ 15 รูปแบบที่ออกแบบนั้นอาจจะไม่เหมาะกับข้อมูลที่นำมาใช้ในการทดลองหรือไม่เหมาะกับสถาปัตยกรรมเหล่านี้

โครงสร้างของแบบจำลอง (VGG16 + FC Pattern 11) มีชั้นเชื่อมโยงที่มีการลดจำนวนโหนดเป็นลำดับ [384 -> 192 -> 96] ร่วมกับการใช้ Batch Normalization 2 ชั้นที่ช่วยให้การปรับค่าพารามิเตอร์ในระหว่างการเรียนรู้เป็นไปอย่างสมบูรณ์มากยิ่งขึ้น นอกจากนั้นมีการใช้งานเทคนิค L2 Regularization และ Dropout (0.4 - 0.5) ที่เป็นปัจจัยสำคัญที่ช่วยควบคุมไม่ให้แบบจำลองเกิดการเรียนรู้ที่จดจำเฉพาะกับข้อมูลตัวอย่างมากเกินไป (Overfitting) ซึ่งได้แสดงให้เห็นในค่าส่วนเบี่ยงเบนมาตรฐานคือ 0.01 ใน Table 3 รวม

ทั้งแสดงค่าความสูญเสียของทั้งชุดการเรียนรู้และชุดการประเมินมีแนวโน้มลดลงอย่างต่อเนื่องและคงที่ อีกทั้งเส้นกราฟทั้งสองวิ่งเข้าหาจุดเดียวกันไม่มีช่องว่างที่กว้างผิดปกติ (Gap) ใน Figure 9 ส่งผลให้แบบจำลองนี้สามารถคัดเลือกคุณลักษณะเด่นของภาพรังสีทรวงอกและคัดแยกได้อย่างมีประสิทธิภาพมากที่สุด

นำแบบจำลองที่มีประสิทธิภาพมากที่สุดในการทดลอง (VGG16 + FC Pattern 11) มาเปรียบเทียบค่าความถูกต้องและค่าความสูญเสียกับชุดข้อมูล K-fold ในระหว่างการเรียนรู้ของแบบจำลองดังแสดงใน Figure 9 แสดงถึงค่าความถูกต้องมีแนวโน้มเพิ่มขึ้นอย่างรวดเร็วในช่วงต้นและรักษาระดับความถูกต้องไว้ที่ประมาณร้อยละ 94.00 - 95.80 ชุดข้อมูลทดสอบมีความเสถียรในระดับสูง (สูงสุดประมาณร้อยละ 95.80 รอบที่ 14) ในเส้น Folds Line ได้แสดงถึงการกระจายตัวของข้อมูลที่สม่ำเสมอ แม้จะมีความแตกต่างกันบ้างในแต่ละชุดข้อมูลย่อย แต่เส้นเหล่านี้มีแนวโน้มวิ่งเข้าหาจุดเดียวกันในช่วงท้าย ซึ่งช่วยยืนยันว่าชุดข้อมูลทดสอบมีความสอดคล้องสูงกว่าชุดข้อมูล Validation Fold ทำให้แบบจำลองมีความเสถียรสูง (SD = 0.01) รวมทั้งเส้น Validation Line ที่มีทิศทางไปในทางเดียวกับเส้น Training แต่มีความผันผวนมากกว่าเล็กน้อยในช่วงต้น รวมทั้งกราฟแสดง Loss สอดคล้องกันทั้งเส้น Training และ Validation ลดลงอย่างต่อเนื่องและคงที่ตลอดการเรียนรู้ แสดงให้เห็นว่าแบบจำลองไม่เกิดปัญหาการ Overfitting และมีความผิดพลาด (Error/Loss) ต่ำในทุกส่วนของชุดข้อมูล

จากการวิเคราะห์ความซับซ้อนและประสิทธิภาพ (Trade-off Analysis) พบว่ามีความซับซ้อนเชิงคำนวณที่มากเกินไปเป็นปัญหาต่อการจำแนกวัณโรคในชุดข้อมูลที่มีจำกัด เช่น สถาปัตยกรรม EfficientNet B7 ที่มีชั้นความลึก 438 ชั้น และ Inception-ResNet-V2 ที่มีชั้นความลึก 449 ชั้น แม้ว่าจะมีศักยภาพสูงในเชิงทฤษฎี แต่ให้ผลลัพธ์ที่ไม่เสถียรในการเรียนรู้ มีค่าความถูกต้องร้อยละระหว่าง 50.00 - 64.26 เนื่องจากมีพารามิเตอร์เกินความจำเป็นเมื่อเทียบกับปริมาณข้อมูล ส่วนสถาปัตยกรรม VGG16 ซึ่งมีความเป็น Lightweight ในเชิงโครงสร้างด้วยชั้นความลึก 16 ชั้น เมื่อทำงานร่วมกับชั้นเชื่อมโยงแบบสมบูรณ์รูปแบบที่ 11 (FC Pattern 11) ที่ใช้ลดจำนวนโหนดแบบลำดับชั้น สามารถให้ค่าประสิทธิภาพ (Effectiveness) สูงสุดที่ร้อยละ 98.58 ซึ่งผลการทดลองนี้ยืนยันว่าการใช้สถาปัตยกรรมที่ไม่ซับซ้อนเกินไป (Shallow architecture) ร่วมกับชั้นเชื่อมโยงแบบสมบูรณ์ที่เหมาะสมเป็นแนวทางที่มีประสิทธิภาพที่สุดสำหรับระบบสนับสนุนการตัดสินใจทางคลินิกในสภาวะทรัพยากรจำกัด

Table 3 Summary of Experimental Results (Top 10 Models)

No.	Architecture	Pattern	Accuracy (Acc.)			Loss		
			Acc.	SD.	Average	Loss	SD.	Average
1	VGG 16 (Best Model)	11	98.58	0.01	97.52	0.15	0.02	0.16
2	VGG 16	14	98.43	0.01	96.64	0.15	0.02	0.18
3	VGG 16	7	98.27	0.01	96.96	0.18	0.05	0.19
4	VGG 19	4	97.80	0.01	97.07	0.25	0.03	0.26
5	VGG 16	2	97.64	0.01	96.04	0.20	0.03	0.27
6	Inception-ResNet-v2	15	96.70	0.01	95.95	0.21	0.02	0.21
7	Inception-ResNet-v2	13	96.08	0.009	95.55	0.18	0.02	0.18
8	Inception-ResNet-v2	5	95.76	0.01	96.70	0.24	0.02	0.21
9	Inception-ResNet-v2	2	95.61	0.006	96.52	0.23	0.01	0.25
10	Inception-ResNet-v2	14	95.29	0.01	95.94	0.21	0.05	0.25

Table 4 Best-Performing Model Results per Architecture

No.	Architecture	Pattern	Accuracy (Acc.)			Loss		
			Acc.	SD.	Average	Loss	SD.	Average
1	VGG16 (Best Model)	11	98.58	0.01	97.52	0.15	0.02	0.16
2	VGG19	4	97.80	0.01	97.07	0.25	0.03	0.26
3	Inception-ResNet-v2	15	96.70	0.01	95.95	0.21	0.02	0.21
4	ResNet101	11	73.04	0.13	64.06	0.70	1.22	1.56
5	ResNet152	1	64.26	0.17	61.85	0.64	0.99	1.49
6	GoogLeNet	12	57.99	0.00	50.00	1.2	0.0009	1.39
7	ResNet50	11	50.15	0.16	57.49	4.23	1.25	5.4
8	EfficientNet B0 - B1	-	50.00	-	-	0.69	-	-

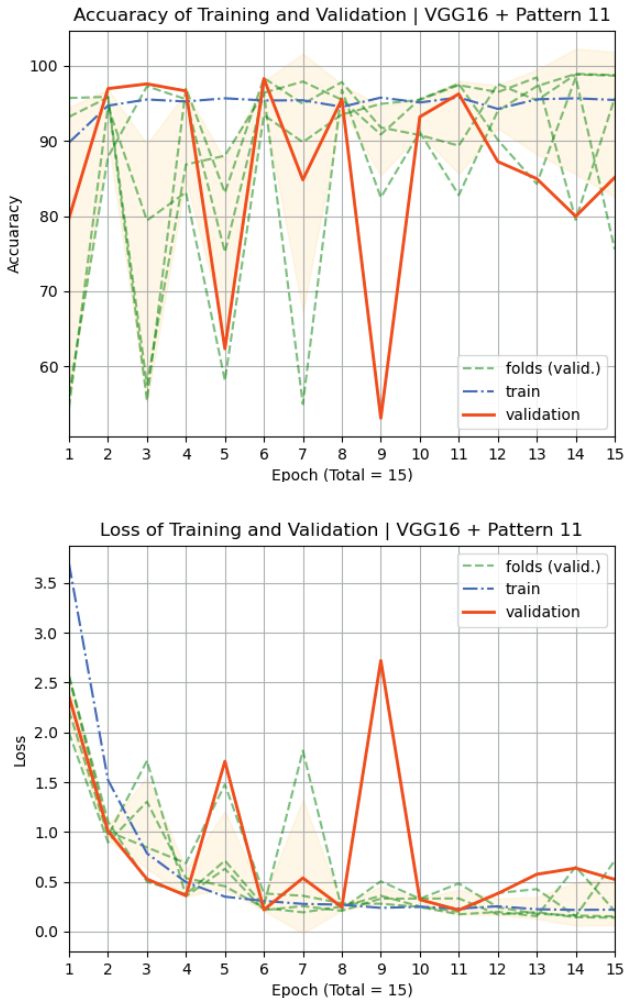

Figure 9 Accuracy/Loss Best Model Compared K-fold. (VGG16 + FC Pattern 11)

Conclusion

งานวิจัยได้นี้นำเสนอการพัฒนาแบบจำลองในการจำแนกวัณโรค โดยใช้สถาปัตยกรรมที่ได้รับความนิยมจำนวน 15 สถาปัตยกรรมในการสกัดคุณลักษณะเด่นของข้อมูลและออกแบบชั้นเชื่อมโยงแบบสมบูรณ์จำนวน 15 รูปแบบสำหรับการจำแนกข้อมูล รวมทั้งสิ้น 225 แบบจำลอง ผลการทดลองได้แสดงให้เห็นถึงความสำเร็จในการเปรียบเทียบและคัดเลือกสถาปัตยกรรมที่เหมาะสมที่สุดสำหรับการจำแนกวัณโรค พบว่าแบบจำลอง VGG16 ร่วมกับชั้นเชื่อมโยงแบบสมบูรณ์รูปแบบที่ 11 (VGG16 + FC Pattern 11) ให้ประสิทธิภาพสูงสุดมีค่าความถูกต้องแม่นยำที่ร้อยละ 98.58 โดยโครงสร้างมีการลดจำนวนโหนดเป็นลำดับ [384 -> 192 -> 96] ร่วมกับเทคนิค Batch Normalization และ L2 Regularization ช่วยลดความซับซ้อนของพารามิเตอร์ที่ล้นเกินได้อย่างมีประสิทธิภาพ ส่งผลให้แบบจำลองมีความประหยัดทรัพยากรเชิงคำนวณ (Lightweight) ซึ่งมีโครงสร้างสถาปัตยกรรมที่ไม่ซับซ้อนมีความลึก (Depth) 20 ชั้น เหมาะกับบริบท

โรงพยาบาลที่ขาดแคลนเครื่องมือสมรรถนะสูงซึ่งตอบโจทย์ข้อจำกัดด้านเทคโนโลยี อีกทั้งในข้อจำกัดด้านปริมาณข้อมูลที่นำมาใช้ในการเรียนรู้ของแบบจำลอง จำนวน 6,388 ภาพ โดยไม่ใช้เทคนิคการเพิ่มปริมาณข้อมูล (Data Augmentation) แบบจำลองยังสามารถมีค่าความถูกต้องมากกว่าร้อยละ 98.00

นอกจากนั้นผลการทดลองยังแสดงให้เห็นถึงศักยภาพเชิงทฤษฎีของแบบจำลองที่พัฒนาขึ้นในการนำไปใช้เป็นเครื่องมือคัดกรองเบื้องต้น (Preliminary Screening) เมื่อพิจารณาจากค่าความไว (Sensitivity) ร้อยละ 99.05 ซึ่งมีความสำคัญอย่างยิ่งในการลดความเสี่ยงที่ผู้ป่วยจะหลุดรอดจากการคัดกรอง แบบจำลองที่พัฒนาขึ้นนี้จึงมีความพร้อมในการพัฒนาต่อยอดเป็นระบบสนับสนุนการตัดสินใจทางคลินิก (CDSS) เพื่อช่วยรังสีแพทย์ในการคัดกรองเคสปกติออกจากเคสที่สงสัยในขั้นแรกก่อนส่งให้ผู้เชี่ยวชาญดำเนินการในขั้นตอนถัดไป

Future Research

ในระยะถัดไปควรเน้นการเพิ่มความหลากหลายของข้อมูล (Data Diversity) จากแหล่งที่มาที่ต่างกัน เช่น โรงพยาบาลในประเทศไทยที่เปิดข้อมูลแบบสาธารณะหรือทำงานร่วมกับโรงพยาบาล เพื่อให้มีปริมาณจำนวนเคส และรูปแบบของข้อมูลมากขึ้นซึ่งจะช่วยยืนยันความแม่นยำเมื่อนำไปปรับใช้ในบริบทหน้างานจริงได้อย่างมีประสิทธิภาพ

สำหรับประสิทธิภาพเชิงประจักษ์ของแบบจำลองในการนำมาใช้งานจริง เช่น การเพิ่มความรวดเร็วในกระบวนการวินิจฉัยหรือการลดภาระงานของบุคลากรทางการแพทย์ในพื้นที่ขาดแคลน ซึ่งควรเน้นการศึกษาความเป็นไปได้ (Feasibility Study) ในสถานการณ์จริงร่วมกับสถานพยาบาลในประเทศไทย โดยจะมีการเก็บข้อมูลเชิงตัวเลขเพื่อวัดผลกระทบที่เป็นรูปธรรม เช่น ระยะเวลาที่ลดลงในการวินิจฉัย และความพึงพอใจของผู้ใช้งาน

นอกจากนี้ในเชิงเทคนิคควรมีการพัฒนาแบบจำลองให้รองรับการจำแนกหลายกลุ่มอาการ (Multi-class Classification) เพื่อให้ระบบสามารถแยกแยะวัณโรคออกจากโรคทรวงอกอื่นที่มีลักษณะใกล้เคียงกัน เช่น ปอดอักเสบ หรือลมในช่องปอด และศึกษาเทคนิคการสกัดคุณลักษณะเชิงลึกในระดับ Patch-level เช่น การใช้ Deformable Convolution รวมถึงการบูรณาการตรรกะคลุมเครือ (Fuzzy Logic) เพื่อจัดการกับความซับซ้อนของค่าความเข้มสีในภาพ

และศึกษาแนวทางการนำแบบจำลองไปติดตั้งบนอุปกรณ์เคลื่อนที่ (Mobile Health) หรือพัฒนาแพลตฟอร์มการแพทย์ทางไกล (Telemedicine) เพื่อพิสูจน์ความเป็นแบบจำลองที่ประหยัดทรัพยากร (Lightweight) ที่สามารถใช้งานได้จริงในพื้นที่ห่างไกลที่ขาดแคลนทั้งงบประมาณและบุคลากรผู้เชี่ยวชาญ

References

1. World Health Organization (WHO). Tuberculosis [Online]. Available: https://www.who.int/health-topics/tuberculosis. [1 October 2025]
2. The Government Public Relations Department (PRD). Warning... Be Alert to the Danger of Tuberculosis (TB) [Online]. Available: https://www.prd.go.th/th/content/ category/detail/id/39/iid/376117. [1 October 2025] (In Thai)
3. Wajgi, R., Yenurkar, G., Nyangaresi, V., Wanjari, B., Verma, S., Deshmukh, A. and Mallewar, S. 2024. Optimized tuberculosis classification system for chest X - ray images: Fusing hyperparameter tuning with transfer learning approaches. *Engineering Reports*, 6, e12906. https://doi.org/10.1002/eng2.12906
4. Rim, B., Jang, H., Lee, H. and Jeon, W. 2025. Active and Inactive Tuberculosis Classification Using Convolutional Neural Networks with MLP-Mixer. *Bioengineering (Basel)*, 12, 630. https://doi.org/10.3390/bioengineering12060630
5. Mirugwe, A., Tamale, L. and Nyirenda, J. 2025. Improving Tuberculosis Detection in Chest X-Ray Images Through Transfer Learning and Deep Learning: Comparative Study of Convolutional Neural Network Architectures. *JMIRx Med*, 6, e66029. https://doi.org/10.2196/66029
6. Haque, B., Siddiqui, A. and Gupta, A. 2025. Optimized tuberculosis image classification in chest X-rays using transfer learning and hyper parameters. *IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, 18-19 January 2025, Bhopal, India, 1-6. https://doi.org/10.1109/SCEECS64059.2025.10940331
7. Fati, M., Senan, M. and ElHakim, N. 2022. Deep and hybrid learning technique for early detection of tuberculosis based on X-ray images using feature fusion. *Applied Sciences*, 12, 7092. https://doi.org/10.3390/app12147092
8. Genitha, C., Kalaivani, I., Ajibah, H., Jalagandeswaran, S. and Balamurugan, K. 2023. Automated Framework for the Tuberculosis Detection and Classification in X-Ray Images Using Deep Learning Algorithm. *International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS)*, 18-20 October 2023, Erode, India, 215-220. https://doi.org/10.1109/ICSSAS57918.2023.10331715

9. Hossain, M., Mural, H., Monowar, and Hridoy, W. 2026. Automated Tuberculosis Detection in Chest Radiographs: A Hybrid Deep Learning Framework for Clinical Decision Support. *Applied AI Letters*, 7, e70022. https://doi.org/10.1002/ail2.70022
10. Kiran, S. and Jabeen, I. 2024. Dataset of Tuberculosis Chest X-rays Images [Online]. Available: https://data.mendeley.com/datasets/8j2g3csprk/2 [1 October 2025]
11. Rahman, T., Chowdhury, M. and Khandakar, A. 2021. Tuberculosis (TB) Chest X-ray Database [Online]. Available: https://www.kaggle.com/datasets/tawsifurrahman/tuberculosis-tb-chest-xray-dataset [1 October 2025]