

การคัดกรองวัณโรคด้วยวิธีการถ่ายภาพการเรียนรู้เชิงลึก



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2564

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Screening Tuberculosis Using Deep Transfer Learning



A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Computer Science

Department of Computer Engineering

FACULTY OF ENGINEERING

Chulalongkorn University

Academic Year 2021

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การคัดกรองไวรัสโรคด้วยวิธีการถ่ายทอดการเรียนรู้เชิงลึก
โดย	นายชัยสิทธิ์ พัฒนาสุวรรณ
สาขาวิชา	วิทยาศาสตร์คอมพิวเตอร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ศาสตราจารย์ ดร.ประภาส จงสฤษดิ์วัฒนา

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้หัวข้อวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่ง
ของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

----- คณบดีคณะวิศวกรรมศาสตร์
(ศาสตราจารย์ ดร.สุพจน์ เตชวรสินสกุล)

คณะกรรมการสอบวิทยานิพนธ์

----- ประธานกรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.สุกรี สินธุภิญโญ)

----- อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ศาสตราจารย์ ดร.ประภาส จงสฤษดิ์วัฒนา)

----- กรรมการภายนอกมหาวิทยาลัย
(ผู้ช่วยศาสตราจารย์ ดร.ธนันท์ เพ็ชรตระกูล)

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ชัยสิทธิ์ พัฒนาสุวรรณ : การคัดกรองวัณโรคด้วยวิธีการถ่ายถอดการเรียนรู้เชิงลึก. (Screening Tuberculosis Using Deep Transfer Learning) อ.ที่ปรึกษาหลัก : ศ. ดร. ประภาส จงสถิตย์วัฒนา

วัณโรคเป็นปัญหาสาธารณสุขที่สำคัญ ควรมีการคัดกรองเชิงรุกเพื่อหาผู้ที่ติดเชื้อแยก ออกมารักษาเพื่อไม่ให้โรคแพร่กระจาย โดยการพัฒนารูปแบบการเรียนรู้ของเครื่องเพื่อคัดกรองผู้ที่มีความเสี่ยง ซึ่งจะเป็นประโยชน์ต่อทั้งผู้ให้บริการด้านการดูแลสุขภาพและผู้ป่วย การประยุกต์ใช้เทคนิคการเรียนรู้เชิงลึกสำหรับการจำแนกภาพทางการแพทย์ได้รับการพัฒนาและเติบโตอย่างรวดเร็วในช่วงไม่กี่ปีที่ผ่านมา คณะวิจัยได้เลือกใช้แบบจำลอง Convolution Neural Network (CNN) เนื่องจากเป็นหนึ่งในโมเดลที่มีชื่อเสียงและมีประสิทธิภาพสูงสำหรับงานในการแยกประเภทรูปภาพานวิจัยนี้นำเสนอโครงข่ายประสาทเทียมเพื่อจำแนกภาพทรวงอกออกเป็น 2 ประเภท คือ ภาพเอกซเรย์ปอดปกติและปอดที่ติดเชื้อวัณโรค โดยมีแหล่งชุดข้อมูลภาพเอกซเรย์ทรวงอก 3 ชุด ได้แก่ Montgomery, Shenzhen และกรองวัณโรค กรมควบคุมโรค นักวิจัยได้เปรียบเทียบสถาปัตยกรรมแบบจำลองการประมวลผลของ CNN ทั้ง 4 แบบเพื่อหาแบบจำลองที่ดีที่สุดที่เหมาะสมสำหรับการเอกซเรย์ทรวงอก โดยประเมินประสิทธิภาพแบบจำลองด้วยการใช้ Accuracy, Precision, Recall และ AUC ผลการศึกษานี้แสดงให้เห็นว่าแบบจำลอง DenseNet มีความแม่นยำมากกว่ารุ่นอื่นๆ และเราปรับแต่งแบบจำลองสำหรับเกณฑ์ที่ดีที่สุด และนำมาฝึกเพิ่มกับภาพเอกซเรย์ทรวงอกของกองวัณโรคของไทย เพื่อให้เหมาะสมกับงานคัดกรองวัณโรคสำหรับคนไทย ความแม่นยำในการทำนายภาพเอกซเรย์ของ ปอดปกติและปอดที่ติดเชื้อวัณโรคในแบบจำลองที่ดีที่สุดคือ 91% และ AUC คือ 95% แบบจำลองที่ได้พัฒนาขึ้นมานี้จะช่วยผู้ให้บริการด้านสุขภาพในการคัดกรองวัณโรค สำหรับประชากรจำนวนมากในประเทศไทย

สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์

ลายมือชื่อนิสิต

ปีการศึกษา 2564

ลายมือชื่อ อ.ที่ปรึกษาหลัก

6272027221 : MAJOR COMPUTER SCIENCE

KEYWORD:

Chaiyasit Pattanasuwan : Screening Tuberculosis Using Deep Transfer Learning. Advisor: Prof. PRABHAS CHONGSTITVATANA, Ph.D.

Tuberculosis is a major public health problem and has to be proactive screening for quarantine by means of developing machine learning model to screen suspected case. This can be mutually beneficial to healthcare providers and patients. The application of deep learning technique for medical image classification has been developed and grown exponentially over the past few years. We propose Convolution Neural Network (CNN) model because it is one of several well-known and high performance models for image classification. This research presents neural network to classify chest imaging into 2 classes: normal and tuberculosis. We collect 3 datasets of chest X-ray image: Montgomery, Shenzhen and Bureau of tuberculosis. The researchers compared 4 CNN classification models to find out the best model that is suitable for chest X-ray. Performance was measured by using metrics: accuracy, precision, recall and AUC. The result of this study shows that DenseNet model is more accurate than others and we tune the model for the best threshold and train it with Thai Bureau of tuberculosis chest image for screening TB for Thai people. The accuracy for discrimination normal lung and TB-infected lung in the best model is 91% and AUC is 95%. This model would aided healthcare providers for TB screening large population in Thailand.

Field of Study: Computer Science

Student's Signature

Academic Year: 2021

Advisor's Signature

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความอนุเคราะห์อย่างยิ่งของศาสตราจารย์ ดร. ประภาส จงสถิตย์วัฒนา อาจารย์ที่ปรึกษาวิทยานิพนธ์ ซึ่งท่านได้กรุณาสละเวลาให้ความรู้ ให้คำปรึกษา ตรวจสอบ ให้คำแนะนำแนวทางการวิจัย และสนับสนุนจนทำให้การวิจัยในครั้งนี้สำเร็จลุล่วงด้วยดี ข้าพเจ้าจึงขอกราบระลึกพระคุณศาสตราจารย์ ดร.ประภาส จงสถิตย์วัฒนา ไว้ ณ ที่นี้

ข้าพเจ้าขอกราบขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.สุกรี สິนธูภิญโญ และ ผู้ช่วยศาสตราจารย์ ดร.ธনী เพียรตระกูล กรรมการสอบวิทยานิพนธ์ ที่ได้กรุณาสละเวลา ให้คำแนะนำ ตรวจสอบ และแก้ไขวิทยานิพนธ์ฉบับนี้ให้ถูกต้องสมบูรณ์ยิ่งขึ้น

ข้าพเจ้าขอขอบพระคุณกองวินโรค กรมควบคุมโรค กระทรวงสาธารณสุขเป็นอย่างยิ่ง ที่ได้ให้คำแนะนำทางวิชาการ และแหล่งข้อมูลที่สำคัญในการวิจัยครั้งนี้

ท้ายนี้ข้าพเจ้าขอกราบขอบพระคุณ คุณพ่อ คุณแม่ และครอบครัวสำหรับกำลังใจที่มีค่ายิ่ง รวมถึงขอขอบพระคุณผู้บังคับบัญชาในสายงาน เพื่อนร่วมงาน และมิตรสหาย ที่คอยติดตามให้กำลังใจ ให้การสนับสนุนและความช่วยเหลือในด้านต่าง ๆ และท่านอื่น ๆ ที่มีได้กล่าวชื่อไว้ ณ ที่นี้ที่มีส่วนช่วยให้วิทยานิพนธ์ของข้าพเจ้าสำเร็จไปได้ด้วยดี

ชัยสิทธิ์ พัฒนาสุวรรณ

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

สารบัญ

	หน้า
.....	ค
บทคัดย่อภาษาไทย.....	ค
.....	ง
บทคัดย่อภาษาอังกฤษ	ง
กิตติกรรมประกาศ	จ
สารบัญ.....	ฉ
สารบัญตาราง.....	ฉ
สารบัญรูปภาพ.....	ญ
บทที่ 1 บทนำ.....	1
1.1 ความสำคัญและที่มาของงานวิจัย	1
1.2 วัตถุประสงค์ของงานวิจัย.....	3
1.3 ขอบเขตของงานวิจัย.....	3
1.4 แนวทางในการดำเนินงาน.....	3
1.5 วิธีดำเนินการวิจัย.....	4
1.6 ประโยชน์ที่คาดว่าจะได้รับ.....	4
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	5
2.1 ทฤษฎีที่เกี่ยวข้อง	5
2.1.1 ภาพรังสีเอกซเรย์ทรวงอก (Chest X-ray imaging).....	5
2.1.2 การเรียนรู้เชิงลึก (Deep learning).....	6
2.1.3 โครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional neural networks).....	7
2.1.4 การถ่ายทอดการเรียนรู้ (Transfer learning)	9

2.1.5 วัณโรค (Tuberculosis).....	10
2.2 งานวิจัยที่เกี่ยวข้อง	10
บทที่ 3 แนวคิดและวิธีวิจัย	12
3.1 การเก็บรวบรวมข้อมูล.....	12
3.2 การเตรียมชุดข้อมูล	12
3.3 การเลือกสถาปัตยกรรมแบบจำลองที่นำมาใช้เพื่อทำนายวัณโรคปอด	13
3.3.1 VGG16 13	
3.3.2 ResNet50	14
3.3.3 DenseNet121	15
3.3.4 EfficientNetB0.....	17
3.4 การประเมินแบบจำลอง	18
บทที่ 4 การทดลองและผลการทดลอง.....	21
4.1 ระบบปฏิบัติการและภาษาที่ใช้ในการทดลอง	21
4.2 ชุดข้อมูลที่ใช้ในการทดลอง	21
4.3 การเตรียมชุดข้อมูล	21
4.5 การแบ่งข้อมูลเพื่อให้โมเดลเรียนรู้.....	22
4.6 การดำเนินการทดลองและผลที่ได้ในแต่ละแบบจำลอง	22
4.7 ผลการทดลองของแบบจำลองต่าง ๆ.....	26
4.8 การนำแบบจำลอง DenseNet121 มาทดสอบกับฐานข้อมูล.....	27
4.9 การกำหนดค่าความไวให้เป็น 100%	30
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ	32
5.1 สรุปผลการวิจัย.....	32
5.2 ข้อเสนอแนะ	33

บรรณานุกรม.....	35
ประวัติผู้เขียน.....	38



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

สารบัญตาราง

	หน้า
ตารางที่ 1 โครงข่ายประสาทเทียม EfficientNetB0 baseline.....	18
ตารางที่ 2 รายละเอียดของตาราง Confusion matrix.....	19
ตารางที่ 3 รายละเอียดของภาพเอกซเรย์ทรวงอกแยกตามฐานข้อมูล.....	21
ตารางที่ 4 แสดง Confusion matrix ของแบบจำลอง VGG16.....	23
ตารางที่ 5 แสดง Confusion matrix ของแบบจำลอง ResNet50.....	24
ตารางที่ 6 แสดง Confusion matrix ของแบบจำลอง DenseNet121.....	24
ตารางที่ 7 แสดง Confusion matrix ของแบบจำลอง EfficientNetB0.....	25
ตารางที่ 8 แสดงประสิทธิภาพของสถาปัตยกรรมแบบจำลองต่าง ๆ.....	25
ตารางที่ 9 แสดงประสิทธิภาพของ DenseNet ที่จำนวนชั้นต่าง ๆ กัน.....	26
ตารางที่ 10 แสดงประสิทธิภาพของแบบจำลองที่ไม่ได้ทำ transfer learning.....	27
ตารางที่ 11 แสดง Confusion matrix ของแบบจำลอง DenseNet121.....	29
ตารางที่ 12 แสดงประสิทธิภาพของ DenseNet121 บนฐานข้อมูลแต่ละฐาน.....	30
ตารางที่ 13 แสดงประสิทธิภาพของแบบจำลองที่กำหนดค่า Recall 100%.....	31

สารบัญรูปภาพ

	หน้า
ภาพที่ 1 Normal and TB Chest film	5
ภาพที่ 2 Neural Network	6
ภาพที่ 3 Convolutional layer.....	8
ภาพที่ 4 Pooling layer	8
ภาพที่ 5 Convolutional Neural Network[11].....	9
ภาพที่ 6 แสดงสถาปัตยกรรมของ VGG16.....	14
ภาพที่ 7 แสดงสถาปัตยกรรมของ ResNet.....	15
ภาพที่ 8 แสดงสถาปัตยกรรมของ DenseNet.....	17
ภาพที่ 9 แสดงการเพิ่มประสิทธิภาพโมเดล EfficientNet โดยการทำให้ Scaling แบบต่าง ๆ.....	18
ภาพที่ 10 ROC Curve	20
ภาพที่ 11 แสดงโครงสร้างของแบบจำลอง DenseNet121.....	28
ภาพที่ 12 ROC Curve ของ DenseNet121 บน 3 dataset	29

บทที่ 1

บทนำ

1.1 ความสำคัญและที่มาของงานวิจัย

- i) วัณโรคเป็นโรคติดต่อที่สำคัญและเป็นปัญหาทางด้านสาธารณสุขทั่วโลก องค์การอนามัยโลกได้ประกาศให้วัณโรคเป็นปัญหาฉุกเฉินด้านสาธารณสุขทั่วโลก ข้อมูลทางระบาดวิทยาพบว่าปัญหาที่สำคัญสามประเด็นของวัณโรคคือ จำนวนผู้ป่วยวัณโรคที่มีจำนวนมากและมีแนวโน้มจำนวนลดลงช้ากว่าที่จะทำให้สามารถควบคุมวัณโรคได้ในระยะเวลาอันใกล้ ประเด็นถัดมาคือวัณโรคในผู้ติดเชื้อเอชไอวีที่ทำให้การดูแลรักษายากลำบากมากขึ้น และประเด็นสุดท้ายคือวัณโรคคือยาหลายขนาดที่จำนวนผู้ป่วยมากขึ้นและการรักษาด้วยระบบยาแนวที่สองมีผลสำเร็จที่ต่ำอีกทั้งมีราคาสูง องค์การอนามัยโลกได้จัดให้ประเทศไทยมีปัญหาวัณโรครุนแรงทั้งสามด้าน โดยทุกปีตั้งแต่ปีค.ศ. 1997 องค์การอนามัยโลก (WHO) ได้ตีพิมพ์รายงานวัณโรคทั่วโลก (Global Tuberculosis) ซึ่งให้ข้อมูลการประเมินสถานการณ์วัณโรคทั่วโลกที่เป็นปัจจุบัน และสรุปความก้าวหน้าและความพยายามในการป้องกัน การวินิจฉัย และการรักษาโรคในระดับประเทศ ภูมิภาค และระดับโลก รายงาน Global TB ปีค.ศ. 2020 เผยแพร่เมื่อวันที่ 14 ตุลาคม ค.ศ. 2020 และรวบรวมในบริบทของกลยุทธ์การควบคุมวัณโรคทั่วโลกและเป้าหมายของสหประชาชาติ (UN)

ในขณะที่ความคืบหน้าในการดำเนินการเป็นไปอย่างล่าช้า โดยดูจากเป้าหมายของการลดอุบัติการณ์วัณโรคลงตั้งเป้าไว้ที่ 20% ระหว่างปีค.ศ. 2015 ถึง 2020 แต่จากรายงานวัณโรคทั่วโลกปี ค.ศ. 2020 ระบุว่ามีการลดอุบัติการณ์วัณโรคลงเพียง 9% โดยลดลงเฉลี่ย 2% ต่อปีเท่านั้น ในทำนองเดียวกัน เป้าหมายการลดอัตราการตายซึ่งตั้งไว้ที่ 35% ยังไม่บรรลุผล โดยพบว่ามีอัตราการลดการเสียชีวิตได้เพียง 14% ระหว่างปีค.ศ. 2015 ถึงปี 2020 โดยพบว่าจาก 10 ล้านคนที่คาดว่าจะเป็วัณโรคในปี 2019 มีการระบุผู้ติดเชื้อและรายงานอยู่ที่ 7.1 ล้านคน (71%) ไปยังโครงการวัณโรคระดับประเทศทั่วโลก โดยเหลือผู้ป่วยที่ได้รับการวินิจฉัยว่าเป็นวัณโรคแต่ไม่ได้รายงานต่อหน่วยงานด้านสาธารณสุข (รวมถึงผู้ที่ไม่ได้รับรายงานจากภาคเอกชน) และผู้ที่ไม่ได้รับการรักษาอีกเป็นจำนวน 2.9 ล้านคน (29%) จึงมีความจำเป็นต้องได้รับความร่วมมือจากหลายฝ่ายอย่างต่อเนื่องและมุ่งเน้นที่เป็นสากลในการระบุและรักษาผู้ป่วยวัณโรค องค์การอนามัยโลกรายงานว่า จำนวนผู้ป่วยวัณโรคที่ตรวจพบลดลง 50% อาจส่งผลให้มีผู้เสียชีวิตจากวัณโรคเพิ่มขึ้นถึง 400,000 คนในหนึ่งปี และคาดว่าจะไม่สามารถยุติวัณโรคในฐานะภัยคุกคามด้านสาธารณสุขทั่วโลกภายในปีค.ศ. 2035 ตามที่กำหนดไว้ในยุทธศาสตร์การยุติวัณโรค ข้อจำกัดที่สำคัญที่ทำให้ยังไม่บรรลุถึงเป้าหมาย คือ ความไม่เพียงพอ

ของบุคลากรและระบบในการหาผู้ติดเชื้อรวมถึงเด็กที่เป็นวัณโรค โดยเฉพาะในกลุ่มวัณโรคดื้อยา การป้องกันวัณโรคและการจัดหาเงินทุนเพื่อตอบสนองต่อวัณโรค รวมทั้งการวิจัยวัณโรคที่จำเป็น[1]

รัฐบาลของประเทศที่มีวัณโรคเฉพาะถิ่นสูงจำเป็นต้องทำให้แน่ใจว่ามีบริการวินิจฉัยวัณโรคอย่างรวดเร็วในสถานบริการสุขภาพทุกแห่ง ด้วยเหตุที่ COVID-19 ทำให้เกิดความขัดข้องในการบริการด้านสุขภาพ หลายประเทศได้รับรายงานว่าใช้เครื่อง GeneXpert สำหรับการทดสอบ COVID-19 และอีกหลายประเทศได้มอบหมายเจ้าหน้าที่โครงการ TB ใหม่ให้กับ COVID-19 ทำให้เกิดการขาดแคลนทรัพยากรการวินิจฉัยและการรักษา TB ที่ขาดแคลนอยู่แล้ว จำเป็นต้องมีแผนนวัตกรรมเพื่อรักษาบริการวินิจฉัยวัณโรคหลังการระบาดใหญ่ของ COVID-19 นวัตกรรมในการปรับแพลตฟอร์มการวินิจฉัยวัณโรคเพื่อคัดกรองทั้งวัณโรคและโควิด-19 การหาเครื่องมือใหม่ในการตรวจคัดกรองเพิ่มเติม และการลงทุนในการพัฒนาชุดตรวจวินิจฉัยอย่างรวดเร็วและประหยัดสำหรับการติดเชื้อทั้งสองนั้นมีความสำคัญและจำเป็นเร่งด่วน[2]

โดยให้ดำเนินการจัดการกับความท้าทายนี้ต่อไป โดยประสานแนวทางสำหรับการประเมินภาระงานสำหรับการดูแลรักษาวัณโรคให้มีประสิทธิภาพมากขึ้น การพัฒนาแนวทางและเครื่องมือที่มีประสิทธิภาพและมีความไวสำหรับการตรวจหาผู้ติดเชื้อวัณโรคสูงโดยเฉพาะ (เช่น Xpert MTB/CAD4TB) และพัฒนาแนวทางในการติดตามการรักษาของผู้ป่วยวัณโรค[3] การเชื่อมโยงเข้ากับการดูแลและการรักษาทุกคนที่สงสัยว่าจะเป็นวัณโรค มาตรการเพื่อให้มั่นใจว่าผู้ป่วยจะเข้ารับการรักษาและอยู่ในการดูแลจนกว่าจะหาย การมีส่วนร่วมกับชุมชนและการพัฒนาความร่วมมือกับภาคเอกชน และมีแผนการนโยบายว่าวัณโรคเป็นองค์ประกอบหนึ่งของระบบการดูแลสุขภาพถ้วนหน้า

จึงเป็นงานที่สำคัญที่จะต้องดำเนินการในการตรวจเชิงรุกเพื่อหาผู้ติดเชื้อวัณโรค นำมาเข้ารับการรักษาเพื่อลดการติดเชื้อจากคนปกติ เพราะเป็นโรคติดเชื้อที่ใช้เวลาในการรักษานานและต่อเนื่องจากการใช้ X-ray พบว่าสามารถหาตรวจได้ทั่วไปและมีความคุ้มค่า การเอกซเรย์จึงถือว่าเป็นเครื่องมือสำคัญที่ใช้ในการคัดกรองและวินิจฉัยวัณโรคร่วมกับการตรวจเสมหะ ผู้วิจัยจึงได้มีการพัฒนาการเรียนรู้ของเครื่องที่จะช่วยให้บุคลากรทางการแพทย์ลดเวลาและค่าใช้จ่ายในการตรวจคัดกรองจำนวนมากโดยใช้ภาพถ่ายเอกซเรย์ การพัฒนาเริ่มจากใช้โครงข่ายประสาทเทียมแบบสังวัตนาการนำมาสอนด้วยวิธีการถ่ายทอดความรู้ กับชุดข้อมูลพื้นฐานและชุดข้อมูลของกองวัณโรค ประเทศไทย เพื่อให้มีความเข้ากันได้กับการใช้งานจริง

1.2 วัตถุประสงค์ของงานวิจัย

ศึกษาและพัฒนาแบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการสำหรับการตรวจหาโรคของวินโรคปอดจากภาพเอกซเรย์ทรวงอก ประเมินเปรียบเทียบประสิทธิภาพของแบบจำลองและผลลัพธ์ที่ได้จากแหล่งข้อมูลต่าง ๆ เพื่อหาโมเดลที่เหมาะสมมาใช้ในการคัดกรองวินโรค

1.3 ขอบเขตของงานวิจัย

- 1) ข้อมูลชุดที่ใช้สอนเป็นข้อมูล opensource ที่ได้มาจากฐานข้อมูลของ Shenzhen, Montgomery และชุดข้อมูลจากกรมวินโรค กองควบคุมโรค
- 2) สร้างโมเดลปัญญาประดิษฐ์ที่พัฒนาขึ้นมาด้วย CNN โดยใช้ภาษา python ในการเขียนโปรแกรมและไลบรารีของ tensorflow และ keras ในการจัดทำโมเดลและใช้การประมวลผลบน GPU
- 3) เปรียบเทียบประสิทธิภาพและผลลัพธ์ที่ได้จากโครงข่ายประสาทเทียมแบบสังวัตนาการ ด้วยวิธีการถ่ายทอดความรู้ในโมเดลแบบต่าง ๆ โดยใช้ชุดข้อมูลเปิดที่ใช้กันอย่างแพร่หลายทั้งการสอนและการทดสอบ
- 4) เปรียบเทียบประสิทธิภาพจากการสอนโมเดลด้วยชุดข้อมูลแต่ละแบบ และรวมทุกฐานข้อมูลนำมาทดสอบกับชุดทดสอบ เพื่อปรับปรุงโมเดลเพื่อให้เหมาะกับการตรวจคัดกรอง

1.4 แนวทางในการดำเนินงาน

- 1) ทบทวนวรรณกรรมและงานวิจัยที่เกี่ยวข้อง
- 2) ศึกษาวิธีการเลือกใช้เครื่องมือเพื่อสร้างแบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ
- 3) รวบรวมและจัดเก็บข้อมูลภาพเอกซเรย์ทรวงอกจากแหล่งข้อมูลต่าง ๆ
- 4) เตรียมชุดข้อมูลภาพเอกซเรย์เพื่อใช้ในการสอนและทดสอบโมเดล
- 5) พัฒนาและเปรียบเทียบโมเดลต่าง ๆ ด้วยวิธีการถ่ายทอดความรู้
- 6) วิเคราะห์ข้อมูลและประเมินผลการวิจัย
- 7) เรียบเรียงและตีพิมพ์ผลงานทางวิชาการ
- 8) จัดทำวิทยานิพนธ์

1.5 วิธีดำเนินการวิจัย

- 1) การจัดการรูป
- 2) การแบ่งคลาส
- 3) เทรนโมเดล
- 4) การวิเคราะห์ความแม่นยำและ AUC
- 5) การปรับจูนโมเดลเพื่อเพิ่มประสิทธิภาพ

1.6 ประโยชน์ที่คาดว่าจะได้รับ

- 1) ใช้ในการตรวจคัดกรองผู้ที่สงสัยว่าจะติดเชื้อไวรัสโคโรนาได้เป็นจำนวนมาก
- 2) ลดโอกาสการแพร่กระจายโรคไปยังประชากรที่ปกติไม่ได้ติดเชื้อ
- 3) ช่วยลดเวลาการทำงานในการอ่านฟิล์มเพื่อคัดกรองคนเป็นจำนวนมาก และเพิ่มขีดความสามารถของบุคลากรทางการแพทย์หน้างานเพื่อแบ่งเบาภาระและมีเวลาไปดูแลผู้ป่วยได้เพิ่มมากขึ้น
- 4) เป็นตัวช่วยในการตัดสินใจของบุคลากรทางการแพทย์เพื่อช่วยในการวินิจฉัยหรือการส่งตรวจเพิ่มเติม

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

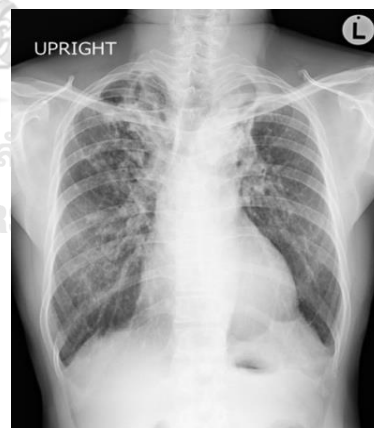
2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 ภาพรังสีเอกซเรย์ทรวงอก (Chest X-ray imaging)

การถ่ายภาพรังสีเอกซเรย์ทรวงอก เป็นการใช้รังสีเอกซ์ซึ่งมีคุณสมบัติที่สามารถทะลุทะลวงผ่านวัตถุต่าง ๆ ได้ โดยฟิล์มเอกซเรย์จะมีความเข้มข้นของการทะลุผ่านวัตถุแตกต่างกันเป็นระดับ ถ้ารังสีผ่านพวกความหนาแน่นต่ำเช่น อากาศ จะเห็นเป็นสีดำ ส่วนเนื้อเยื่อ น้ำ จะเห็นเป็นสีเทา ส่วนเส้นเลือด หัวใจ และกระดูกจะเห็นเป็นลักษณะสีขาวเนื่องจากมีความหนาแน่นสูงกว่า ทำให้แพทย์อ่านผลจากฟิล์มเอกซเรย์เห็นเป็นรูปร่างอวัยวะ และเห็นลักษณะผิดปกติได้ โดยภาพฟิล์มเอกซเรย์ทรวงอกจะเห็นอวัยวะปอดมีลักษณะสีดำ เนื่องจากปอดมีลักษณะที่โปร่งมีลมหรืออากาศอยู่มาก แต่เมื่อผู้ป่วยเป็นวัณโรคปอดจะเกิดการอักเสบทำให้เห็นบริเวณติดเชื่อมมีสีขาวกว่าปอดปกติ มีรูปแบบที่เปลี่ยนไปช่วยให้แพทย์วินิจฉัยโรคได้[4]



Normal film

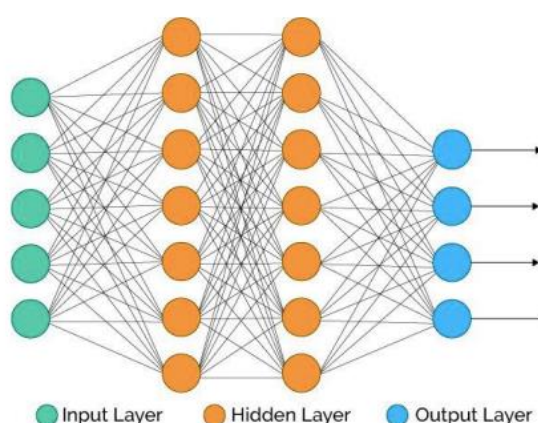


Tuberculosis film

ภาพที่ 1 Normal and TB Chest film

2.1.2 การเรียนรู้เชิงลึก (Deep learning)

การเรียนรู้เชิงลึก (deep learning) เป็นวิธีการเรียนรู้แบบอัตโนมัติโดยเลียนแบบรูปแบบการประมวลผลของสมองมนุษย์โดยใช้เครือข่ายคล้ายเซลล์ประสาท โดยนำระบบโครงข่ายประสาทมาเชื่อมโยงเข้ากันหลายชั้นและให้เรียนรู้ข้อมูลตัวอย่าง ซึ่งข้อมูลดังกล่าวจะนำไปใช้เพื่อวิเคราะห์และความแตกต่างของข้อมูลในเชิงลึก โมเดลจะรวมข้อมูลและเลือกผลความน่าจะเป็นที่ดีที่สุดออกมาเป็นผลลัพธ์ พื้นฐานของการเรียนรู้เชิงลึกคือ อัลกอริทึมที่พยายามจะสร้างแบบจำลองเพื่อแทนความหมายของข้อมูลในระดับสูงโดยการสร้างสถาปัตยกรรมข้อมูลขึ้นมาที่ประกอบไปด้วยโครงสร้างย่อย ๆ หลายอัน ประกอบด้วยชั้น Input Layer ชั้น Hidden Layer และชั้น Output Layer (ดังภาพที่ 2) การเรียนรู้เชิงลึกนั้นอาจมองได้ว่าเป็นวิธีการหนึ่งของการเรียนรู้ของเครื่องที่พยายามเรียนรู้วิธีการแทนข้อมูลอย่างมีประสิทธิภาพ ตัวอย่างเช่น รูปภาพภาพหนึ่ง สามารถแทนได้เป็นเวกเตอร์ของความสว่างต่อจุดพิกเซล หรือมองในระดับสูงขึ้นเป็นเซตของขอบของวัตถุต่าง ๆ หรือมองว่าเป็นพื้นที่ของรูปร่างใด ๆ ก็ได้ การแทนความหมายดังกล่าวจะทำให้การเรียนรู้ที่จะทำงานต่าง ๆ ทำได้ง่ายขึ้น โดยรูปแบบการเรียนรู้เชิงลึกนี้มีหลายแบบได้แก่ โครงข่ายประสาทเทียมแบบลึก (Deep Artificial Neural Networks) โครงข่ายประสาทเทียมแบบวนซ้ำ (Recurrent Neural Network) และโครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Networks) ซึ่งมีการนำมาใช้งานอย่างแพร่หลายในทางคอมพิวเตอร์วิทัศน์ (Computer vision) การรู้จำเสียง การประมวลผลภาษาธรรมชาติ (NLP) ชีวสารสนเทศ เป็นต้น[5]



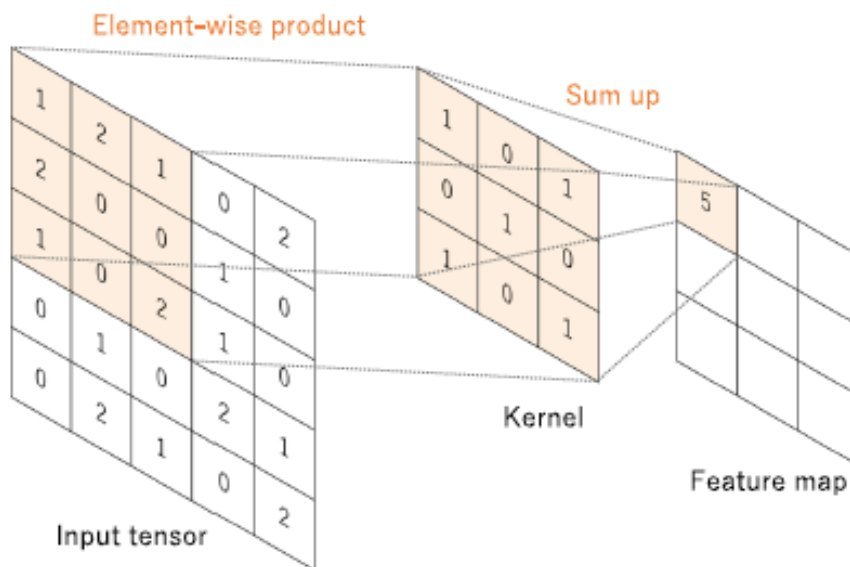
ภาพที่ 2 Neural Network

2.1.3 โครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional neural networks)

โครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Networks, CNN) เป็นโครงสร้างโครงข่ายประสาทเทียมที่สามารถจำแนกข้อมูลภาพได้ดีกว่าโครงข่ายประสาทเทียมแบบอื่นๆ ได้รับแรงบันดาลใจมาจากโครงสร้างประสาทของสมองในส่วนการมองเห็น และออกแบบมาเพื่อเรียนรู้ลำดับชั้นเชิงพื้นที่ โดยมีการนำมาใช้ในงานการจำแนกรูปภาพ การจดจำวัตถุ การจดจำใบหน้า เป็นต้น เนื่องจาก CNN เป็นรูปแบบการเรียนรู้เชิงลึกประเภทหนึ่งที่ทำให้การประมวลผลข้อมูลที่มีรูปแบบกริดจึงเหมาะกับรูปภาพ และอีกเหตุผลที่ทำให้ CNN มีประสิทธิภาพสูงสำหรับการประมวลผลภาพก็คือจะมี kernel ซึ่งเป็นตัวแยกคุณลักษณะในแต่ละตำแหน่งภาพ เนื่องจากคุณลักษณะเฉพาะนั้นอาจเกิดขึ้นที่ใดก็ได้ในภาพ เมื่อชั้นหนึ่งดึงเอาเอาต์พุตออกไปยังชั้นถัดไป คุณลักษณะที่แยกออกมาจะมีความซับซ้อนมากขึ้นตามลำดับชั้นที่เพิ่มขึ้น ทำให้การมี kernel จึงดึงคุณลักษณะที่สำคัญออกมาจากภาพได้ไม่ว่าจะอยู่ในส่วนใดของภาพก็ตาม[6] โดย CNN ประกอบไปด้วยชั้นหลัก ๆ 3 ชั้นคือ Convolutional layer, Pooling layer และ Fully connected layer โดยสองชั้นแรก Convolutional layer, Pooling layer จะทำการแยกคุณลักษณะ ในขณะที่ชั้นที่สาม Fully connected layer จะทำการคำนวณคุณลักษณะที่แยกออกมาเป็นผลลัพธ์สุดท้าย

Convolutional layer

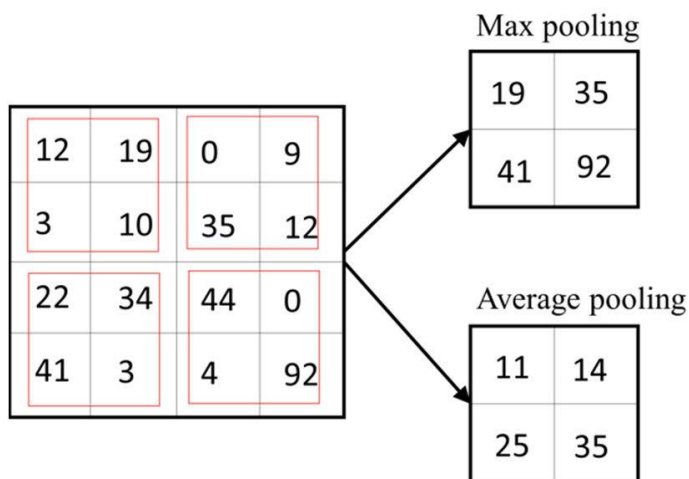
เป็นชั้นที่สำคัญของโครงสร้าง CNN โดยใช้กระบวนการเชิงเส้นแบบพิเศษ ที่สามารถทำการแยกคุณลักษณะส่วนต่าง ๆ ของภาพโดยใช้ตัวตรวจจับคุณลักษณะที่เป็นอาร์เรย์ของตัวเลขขนาดเล็กที่เรียกว่า kernel โดยเพื่อให้โมเดลสามารถเรียนรู้ลักษณะของภาพได้อย่างมีประสิทธิภาพและแม่นยำ kernel จะถูกนำไปใช้กับอินพุตซึ่งเป็นอาร์เรย์ของตัวเลข ที่เรียกว่า tensor โดยใช้การคูณแบบ element-wise product ระหว่าง kernel และ tensor ที่แต่ละตำแหน่งของ tensor แล้วรวมเข้าด้วยกันเพื่อให้ได้ผลลัพธ์ออกมา จะได้ชุดอาร์เรย์ ที่เรียกว่าแผนผังคุณลักษณะ (Feature map) กระบวนการนี้จะทำซ้ำวนไปเรื่อย ๆ โดย kernel หลายรูปแบบ เพราะ kernel แต่ละแบบจะเป็นตัวแยกคุณลักษณะที่แตกต่างกัน[7] (รูปที่ 3) โดยมีไฮเปอร์พารามิเตอร์ที่ปรับได้คือ ขนาดและจำนวนของ kernel, stride and padding เป็นต้น



ภาพที่ 3 Convolutional layer

Pooling layer

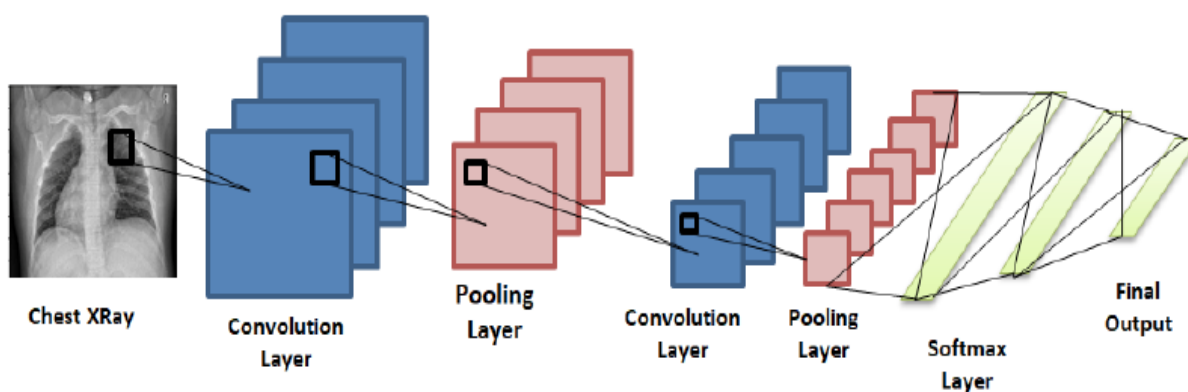
เป็นชั้นของ CNN ที่ช่วยในการลดขนาดของข้อมูล โดยลดมิติของแผนที่คุณลักษณะ จึงทำให้เหลือเฉพาะคุณลักษณะที่สำคัญไฮเปอร์พารามิเตอร์คล้ายกับที่ใช้ใน Convolutional layer โดยจะมีการใช้งานใน 2 แบบคือ การเลือกค่ามากที่สุด (Max Pooling) เป็นกระบวนการทำ pooling ที่ได้รับความนิยมมากที่สุดคือการหาค่าสูงที่สุดจากขนาดช่วงเมทริกซ์ที่กำหนด แล้วส่งออกค่าสูงที่สุดในแต่ละช่วงที่ได้ ในทางปฏิบัติมักใช้การรวมค่าสูงที่สุดที่มี filter ขนาด 2×2 ที่มี stride เท่ากับ 2 ซึ่งจะลดขนาดตัวอย่างในระนาบของแผนที่คุณลักษณะลงได้ 2 เท่า [8] และแบบที่ 2 คือการหาค่าเฉลี่ย (Average pooling) จากขนาดช่วงเมทริกซ์ที่กำหนด (รูปที่ 4)



ภาพที่ 4 Pooling layer

Fully connected layer

เป็นชั้นสุดท้ายของโครงข่ายประสาทสังวัตนาการ โดยจะรับผล feature map ที่ได้จากชั้นสุดท้ายของ convolution หรือ pooling layer ซึ่งส่วนใหญ่จะอยู่ในรูป 1-D หรือในรูป vector ซึ่งผลลัพธ์นี้จะมาเป็น input เชื่อมต่อเข้ากับ fully connected layer และการคำนวณในชั้นนี้มักจะใช้ non-linear function เช่น ReLU, Softmax เป็นต้น โดยผลลัพธ์สุดท้ายจากโครงข่ายประสาทสังวัตนาการ จะได้เป็นค่าความน่าจะเป็นของแต่ละคลาสในการจำแนก โดยชั้นสุดท้ายของ fully connected layer มักจะมีจำนวน node เท่ากับจำนวนคลาสที่ต้องการจำแนก



ภาพที่ 5 Convolutional Neural Network[11]

2.1.4 การถ่ายทอดการเรียนรู้ (Transfer learning)

Transfer Learning คือ กระบวนการถ่ายทอดการเรียนรู้จากการเรียนรู้ครั้งก่อน ๆ นำไปใช้ในการเรียนรู้ใหม่ เพื่อให้เวลาที่ใช้ในการสอนโมเดลเร็วขึ้น ซึ่งแตกต่างจากกระบวนการฝึกอบรมการเรียนรู้ของเครื่องแบบเดิม ๆ ที่เรียนรู้ข้อมูลอินพุตตั้งแต่เริ่มต้น ทำให้ต้องใช้เวลาในการคำนวณงานการทำงานของ การถ่ายทอดการเรียนรู้คือการรักษาพารามิเตอร์ในชั้นก่อนหน้าและลบชั้นสุดท้ายของโมเดล ความรู้จากตัวอย่างที่มีแหล่งที่มาต่าง ๆ การถ่ายทอดการเรียนรู้ทำให้โมเดลสามารถเรียนรู้ตัวอย่างในสภาพแวดล้อมที่มีข้อมูลจำกัดได้[9] โดยที่สามารถทำได้ผ่านโมเดล CNN ที่นิยมใช้กัน เช่น AlexNet, DenseNet, VGG, ResNet และ Inception EfficientNet เป็นต้น

การถ่ายภาพรังสีเอกซเรย์ทรวงอก เป็นการใช่วัตถุที่มีคุณสมบัติที่สามารถทะลุทะลวงผ่านวัตถุต่าง ๆ ได้ โดยฟิล์มเอกซเรย์จะมีความเข้มข้นของการทะลุผ่านวัตถุแตกต่างกันเป็นระดับ ถ้าเป็นวัตถุพวกความหนาแน่นสูง เช่น กระดูกหรือโลหะ จะเห็นเป็นสีขาวในภาพฟิล์มเอกซเรย์ แต่ถ้า

รังสีผ่านพวกความหนาแน่นต่ำ เช่น อากาศ จะเห็นเป็นสีดำ ส่วนเนื้อเยื่อ น้ำ จะเห็นเป็นสีเทา ทำให้แพทย์อ่านผลจากฟิล์มเอกซเรย์เห็นเป็นรูปร่างอวัยวะและเห็นลักษณะผิดปกติได้ ในภาพฟิล์มเอกซเรย์ทรวงอกจะเห็นอวัยวะปอดมีลักษณะสีดำ เนื่องจากปอดมีลักษณะที่โปร่งมีลมหรืออากาศอยู่มาก ส่วนเส้นเลือดและหัวใจจะเห็นเป็นลักษณะสีขาวเป็นมีความหนาแน่นมากกว่า แต่เมื่อผู้ป่วยเป็นวัณโรคปอดจะเห็นลักษณะสำคัญที่ใช้วินิจฉัยโรคได้ เมื่อมีการติดเชื้อจะเกิดการอักเสบทำให้เห็นบริเวณติดเชื้อมีสีขาวกว่าปอดปกติ ลักษณะเช่นนี้ทำให้แพทย์วินิจฉัยโรคได้

2.1.5 วัณโรค (Tuberculosis)

วัณโรคเป็นเชื้อแบคทีเรียชนิดหนึ่ง มีชื่อเป็นทางการว่า *Mycobacterium tuberculosis* เชื้อนี้สามารถทนอยู่ในอากาศ และสิ่งแวดล้อมได้นาน มักเข้าสู่คนโดยการหายใจเอาเชื้อนี้เข้าไป แต่ก็อาจเข้าทางอื่น เช่น บาดแผลได้ด้วย โดยปกติเมื่อร่างกายได้รับเชื้อวัณโรค ร่างกายจะสามารถควบคุมเชื้อนี้ได้ระดับหนึ่งทำให้ไม่เป็นโรค เชื้อจะซ่อนอยู่ในร่างกายที่ติดเชื้อวัณโรคมักไม่แสดงอาการใด ๆ แต่ตัวโรคจะดำเนินต่อไป ผู้ที่ติดเชื้อวัณโรคส่วนใหญ่จะอยู่ในระยะแฝง จะยังมีเชื้อวัณโรคอยู่ในร่างกาย แต่ตัวเชื้อจะอยู่นิ่ง ไม่ทำปฏิกิริยาใด ๆ ต่อร่างกาย ผู้ที่ติดเชื้อในระยะนี้จึงไม่มีอาการใด ๆ และไม่แพร่เชื้อไปยังผู้อื่น หากร่างกายอ่อนแอ เชื้อก็จะกำเริบก่อให้เกิดโรคได้ เช่น เด็กเล็ก ๆ คนที่เป็นโรคเบาหวาน คนที่ได้รับยากดภูมิคุ้มกัน คนใช้เฮดส์

วัณโรคปอด สามารถตรวจได้ค่อนข้างง่าย โดยเอกซเรย์ปอดและตรวจเสมหะหาเชื้อ ซึ่งการใช้เอกซเรย์ปอดในการตรวจนี้มักจะพบความผิดปกติอยู่ที่ปอดด้านบนทั้งสองข้างก่อนลักษณะเป็นฝ้าขาวๆกว่าบริเวณอื่น

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

2.2 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่มีการใช้ปัญญาประดิษฐ์ในทางการแพทย์ปัจจุบันมีจำนวนงานเพิ่มขึ้นอย่างต่อเนื่อง และมีมากขึ้นในหลายสาขา Pesapane et al.[10] พบว่างานวิจัยเหล่านี้เพิ่มขึ้นจาก 100-150 งานวิจัยในช่วงปี 2010 เป็น 700-800 กว่างานวิจัยในช่วงปี 2020 โดยงานวิจัยปัญญาประดิษฐ์ทางการแพทย์ส่วนใหญ่จะเกี่ยวข้องกับงานทางรังสีวิทยา โดยจะเป็นภาพที่ได้จากการทำ X-ray, Ultrasound, CT, MRI โดยมีตั้งแต่สมอง กระดูก ปอด ทางเดินอาหาร เป็นต้น โดยภาพเอกซเรย์ปอดจะผ่านการจำแนกโรคโดยใช้โครงข่ายประสาทเทียมเป็นหลักโดย Rahib et al.[11] พบว่าโครงข่าย

ประสาทเทียมที่จะนำมาตรวจหาความผิดปกติโดยการใช้ CNN จะดีกว่าโครงข่ายประสาทเทียมแบบอื่นๆ

หลังจากนั้นจึงมีการใช้โมเดลที่มีการพัฒนาขึ้นมาเป็นอย่างดี ซึ่งมีโครงสร้างแบบต่าง ๆ โดยพัฒนามาเป็นลำดับโดยคณบดีนักวิจัย Krizhevsky et al.[12] ได้ศึกษาและพัฒนา deep CNN AlexNet ในปี 2012 Simonyan et al. [13] ได้ศึกษาและพัฒนา deep CNN VGG ในปี 2014 He et al.[14] ได้ศึกษาและพัฒนา deep CNN ResNet ในปี 2016 Huang et al.[15] ได้ศึกษาและพัฒนา deep CNN DenseNet ในปี 2017 Tan et al.[16] ได้ศึกษาและพัฒนา deep CNN EfficientNet ในปี 2019 โดยงานวิจัยเหล่านี้ได้เสนอการใช้แบบจำลองที่มีการปรับปรุงและเสริมการทำงานบนโครงสร้างของ CNN เพื่อการจำแนกภาพในรูปแบบต่าง ๆ ได้ประยุกต์ใช้เทคนิค deep transfer learning มีการพรีเทรน (Pretrained) ที่เลเยอร์ชั้นบน และการประเมินผลใช้ค่าพื้นที่ใต้เส้นโค้ง (Area Under Curve, AUC)

โดย Jaeger et al.[17] นั้นได้พัฒนาโครงสร้างของ CNN ขึ้นมาแล้วใช้ชุดข้อมูลเปิดจากฐานข้อมูลของ Montgomery และ Shenzhen พบว่าได้ค่า AUC 87% และ 90% ตามลำดับและความแม่นยำอยู่ที่ 78% และ 84% ตามลำดับ Hwang et al.[18] ได้ทำการศึกษานบนโครงสร้างของ AlexNet และได้ทำการทดสอบบนข้อมูลผู้ป่วยวัณโรคของชุดข้อมูลที่คณะวิจัยมีได้ AUC 93% ต่อมา Hashmi et al.[19] ได้ทำการศึกษาคัดเลือกในปอดโดยการถ่ายทอดการเรียนรู้บนโครงสร้าง Resnet18, DenseNet121, MobileNetV2, InceptionV3 และ Xception โดยจากการศึกษาพบว่า DenseNet121 และ Resnet18 ได้ความแม่นยำที่สูงกว่าอีก 3 แบบ ทางคณะวิจัยจึงเห็นว่าควรจะนำโครงสร้างของ CNN เหล่านี้มาพัฒนาใช้ในชุดข้อมูลเปิดและชุดข้อมูลของกองวัณโรค เพื่อเพิ่มข้อมูลการเรียนรู้และประสิทธิภาพ ให้สอดคล้องการใช้งานกับรูปแบบความผิดปกติที่พบได้บ่อยได้ประชากรไทย โดยเปรียบเทียบชุดโมเดลแต่ละแบบที่มีประสิทธิภาพสูงและหาโมเดลที่ทำให้ได้ค่าความแม่นยำและค่าความไวในการตรวจคัดกรองสูงที่สุด

บทที่ 3

แนวคิดและวิธีวิจัย

งานวิจัยนี้เป็นการหาแบบจำลองที่มีประสิทธิภาพสูงที่สุด ในการทำนายผลที่เหมาะสมกับภาพเอกซเรย์ทรวงอก เพราะเนื่องจากในแต่ละสถาปัตยกรรมจะให้ผลที่ดีกับภาพในแต่ละแบบที่ไม่เหมือนกัน ดังนั้นเราจึงต้องทดสอบภาพเอกซเรย์ปอดซึ่งเป็นสีขาว ดำ เทาเป็นส่วนใหญ่ และมีการใช้การถ่ายทอดการเรียนรู้มาเพื่อช่วยลดระยะเวลาในการเรียนรู้ของแต่ละแบบจำลอง โดยเราเลือกใช้แบบจำลองซึ่งเป็นที่นิยมได้แก่ VGG16, DenseNet121, ResNet50 และ EfficientNetB0 เพื่อเปรียบเทียบว่าสถาปัตยกรรมของแบบจำลองใดที่จะเหมาะสมที่สุดในการนำมาแยกภาพเอกซเรย์ โดยมีการแบ่งคลาสของรูปที่จะนำมาใช้กับแบบจำลองเป็น 2 ประเภทคือ ปอดปกติ กับปอดที่ติดเชื้อวัณโรค ด้วยวิธีการเรียนรู้เชิงลึกแบบโครงข่ายประสาทเทียมแบบสังวัตนาการ และทำการทดลองต่อโดยการนำเอาแบบจำลองที่ประเมินแล้วว่าดีที่สุด มาทำการเรียนรู้จากชุดข้อมูลจาก 3 แหล่งคือ Montgomery Shenzen และกองวัณโรค ให้แบบจำลองได้ทำการเรียนรู้และทดสอบดูประสิทธิภาพว่า เมื่อเพิ่มชุดข้อมูลเข้าไปแล้วจะทำให้แบบจำลองสามารถทำงานได้ดีขึ้นหรือไม่ แล้วหลังจากนั้นจะนำแบบจำลองที่เรียนรู้จากข้อมูลทั้ง 3 แหล่งมาทดสอบแยกแต่ละฐานข้อมูล วิธีวิจัยจะแบ่งเป็น 4 ขั้นตอนหลักคือ 1) การเก็บรวบรวมข้อมูล 2) การเตรียมชุดข้อมูล 3) การเลือกสถาปัตยกรรมแบบจำลองที่นำมาใช้เพื่อทำนายวัณโรคปอด 4) การประเมินแบบจำลอง

3.1 การเก็บรวบรวมข้อมูล

นักวิจัยเก็บรวบรวมข้อมูลจากฐานข้อมูลเปิด 2 ชุด คือ Montgomery 138 ภาพ Shenzen 662 ภาพ และภาพของกองวัณโรค กรมควบคุมโรคอีก 943 ภาพ (โดยได้ทำการนิรนามข้อมูลแล้ว) ย่อขนาดภาพเอกซเรย์เพื่อให้มีขนาดที่เท่ากัน เพื่อช่วยลดเวลาที่ใช้ในการประมวลผลภาพในแบบจำลองต่าง ๆ

3.2 การเตรียมชุดข้อมูล

จัดเตรียมชุดข้อมูลสอนภาพเอกซเรย์ทรวงอกจากฐานข้อมูลทั้ง 3 แหล่ง เนื่องจากข้อมูลภาพเอกซเรย์ทรวงอกมีขนาดที่หลากหลายและมีขนาดที่ใหญ่มาก จำเป็นต้องทำการย่อขนาดภาพเอกซเรย์เพื่อให้มีขนาดที่เท่ากันก่อนนำไปการประมวลผลในแบบจำลอง โดยย่อขนาดลงมาที่ 224 x 224 pixel และทำ Data augmentation โดยการปรับภาพให้เป็นสีขาว-ดำ (Image adjustment to be black and white) การปรับความสว่างของภาพ (Random brightness adjustment) การ

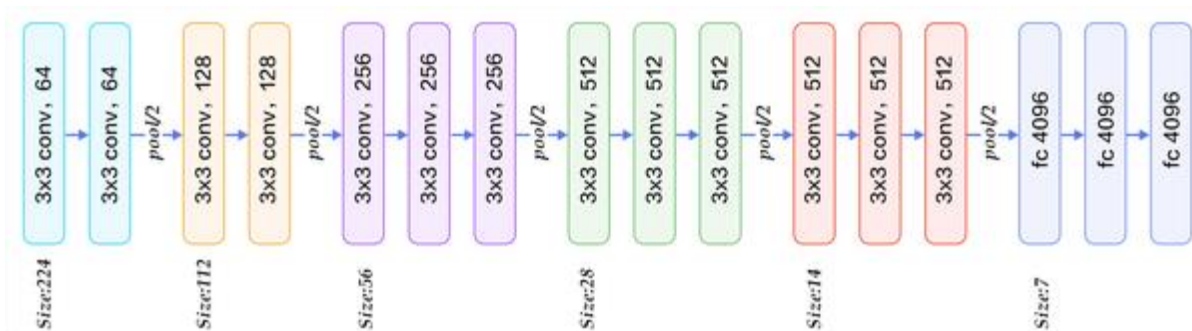
หมุนภาพ (Random rotation range) การหดหรือขยายภาพ (Random scaling zoom range)
การกลับภาพซ้าย-ขวา (Horizontal image flipping)

3.3 การเลือกสถาปัตยกรรมแบบจำลองที่นำมาใช้เพื่อทำนายวัณโรคปอด

3.3.1 VGG16

VGG16 เป็นสถาปัตยกรรมโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) ที่เรียบง่ายและใช้กันอย่างแพร่หลายซึ่งใช้สำหรับ ImageNet ซึ่งเป็นโครงการฐานข้อมูลภาพขนาดใหญ่ที่ใช้ในการวิจัยซอฟต์แวร์การรู้จำภาพ สถาปัตยกรรม VGG16 ได้รับการพัฒนาโดย Karen Simonyan และ Andrew Zisserman จาก University of Oxford ในปี 2014 ผ่านบทความเรื่อง "Very Deep Convolutional Networks for Large-Scale Image Recognition" 'VGG' เป็นตัวย่อของ Visual Geometry Group ซึ่งเป็นกลุ่มนักวิจัยจาก University of Oxford ที่พัฒนาสถาปัตยกรรมนี้ และ '16' บ่งบอกว่าสถาปัตยกรรมนี้มีความลึก 16 ชั้น โดยโมเดล VGG16 ได้รับความแม่นยำในการทดสอบ 5 อันดับสูงสุด 92.7% ใน ImageNet ซึ่งเป็นชุดข้อมูลมากกว่า 14 ล้านภาพที่อยู่ใน 1,000 คลาส เป็นหนึ่งในโมเดลที่มีชื่อเสียงที่ส่งไปยัง ImageNet Large Scale Visual Recognition Challenge (ILSVRC)

โครงสร้าง VGG16 ประกอบไปด้วย 13 Convolution layer และ 3 Fully connected layer ซึ่ง Convolution layer จะทำหน้าที่ดึงคุณลักษณะภาพออกมา (Extract feature) โดยชั้น conv 1-2 จะมี 64 ตัวกรอง ชั้น conv 3-4 จะมี 128 ตัวกรอง ชั้น conv 5-7 จะมี 256 ตัวกรอง และชั้น conv 8-13 มี 512 ตัวกรอง ขนาดตัวกรองที่ใช้ในแต่ละชั้นคือ 3x3 ชั้น conv 2, conv 4, conv 7, conv 10 และ conv13 จะมีการรวมค่ามากที่สุด (Max Pooling Layer) มีขนาด 2x2 และจะมีการทำ Batch Normalize เพื่อที่จะช่วยเพิ่มความเร็วในการเรียนรู้ของโมเดลให้มากยิ่งขึ้น มีการใช้ Activation function ที่ใช้ในแต่ละชั้นคือ Relu function ยกเว้นในชั้น FC 3 ที่จะใช้ Activation function เป็น Softmax function เพื่อที่จะให้ในชั้นสุดท้ายของ fully connected layer นี้เป็นชั้นที่ไว้ทำนายภาพเอกซเรย์ปอดว่าเป็นปกติหรือเป็นวัณโรคปอดหรือไม่ มีการตั้งค่าพารามิเตอร์ต่าง ๆ คือ อัตราการเรียนรู้ (Learning rate) 0.0001 และ Optimizer ที่ใช้คือ Stochastic Gradient Descent (SGD) รายละเอียดของโครงสร้างแต่ละชั้น ดังภาพที่ 6



ภาพที่ 6 แสดงสถาปัตยกรรมของ VGG16

3.3.2 ResNet50

ResNet50 เป็นโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) ที่มีความลึก 50 ชั้น Residual Neural Network (ResNet) เป็นโครงข่ายประสาทเทียมที่วางบล็อกที่เหลื่อมทับกันเพื่อสร้างเครือข่าย โดยการทำให้ residual mapping เพื่อแก้ปัญหาชั้นที่มากขึ้นเรื่อยๆจนทำให้เกิด vanishing/exploding gradients

เนื่องจากเฟรมเวิร์กที่ ResNet50 ได้แก้ไขในจุดนี้มานั้น จึงทำให้สามารถฝึกโครงข่ายประสาทเทียมแบบที่มีชั้นลึกมากๆได้ โดยสามารถมีเลเยอร์ได้หลายร้อยหรือหลายพันเลเยอร์ และตัวโมเดลก็ยังคงมีประสิทธิภาพที่ดี

โครงสร้าง ResNet50 ประกอบไปด้วยชั้น conv1 กับ kernel ขนาด 7×7 64 kernel ที่แตกต่างกันนั้นจะเลื่อนโดยใช้ค่า stride เท่ากับ 2 ในชั้นนี้จะมีทั้งหมด 1 layer และจะมีการรวมค่ามากที่สุด (max pooling) โดยเลื่อนโดยใช้ค่า stride เท่ากับ 2

ในชั้น conv2.x จะมี kernel ขนาด 1×1 64 kernel ตามมาด้วย kernel ขนาด 3×3 64 kernel และ kernel สุดท้ายขนาด 1×1 256 kernel ซึ่งทั้ง 3 เลเยอร์นี้จะมีการทำซ้ำทั้งหมด 3 ครั้ง จึงได้ทั้งหมด 9 เลเยอร์

ในชั้น conv3.x จะมี kernel ขนาด 1×1 128 kernel ตามมาด้วย kernel ขนาด 3×3 128 kernel และ kernel สุดท้ายขนาด 1×1 512 kernel ซึ่งทั้ง 3 เลเยอร์นี้จะมีการทำซ้ำทั้งหมด 4 ครั้ง จึงได้ทั้งหมด 12 เลเยอร์

ในชั้น conv4.x จะมี kernel ขนาด 1×1 256 kernel ตามมาด้วย kernel ขนาด 3×3 256 kernel และ kernel สุดท้ายขนาด 1×1 1024 kernel ซึ่งทั้ง 3 เลเยอร์นี้จะมีการทำซ้ำทั้งหมด 6 ครั้ง จึงได้ทั้งหมด 18 เลเยอร์

ในชั้น conv5.x จะมี kernel ขนาด 1×1 512 kernel ตามมาด้วย kernel ขนาด 3×3 512 kernel และ kernel สุดท้ายขนาด 1×1 2048 kernel ซึ่งทั้ง 3 เลเยอร์นี้จะมีการทำซ้ำทั้งหมด 3 ครั้ง จึงได้ทั้งหมด 9 เลเยอร์

หลังจากนั้นที่เลเยอร์ 1 ชั้นสุดท้ายจะทำการ average pooling เพื่อลดขนาดมิติลง และเข้าสู่ชั้น fully connected layer ที่ประกอบไปด้วย 1000 nodes ผ่าน activation function คือ softmax function เพื่อให้ได้ผลลัพธ์ว่าเป็นคลาสใด โดยสรุปแล้วเมื่อรวมชั้นทั้งหมดจะได้ 50 ชั้นของ Deep Convolutional neural network เป็นไปตามภาพที่ 7

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	$7 \times 7, 64, \text{stride } 2$				
		$3 \times 3 \text{ max pool, stride } 2$				
conv2_x	56×56	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

ภาพที่ 7 แสดงสถาปัตยกรรมของ ResNet

CHULALONGKORN UNIVERSITY

3.3.3 DenseNet121

ในโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) แบบ feed forward แบบดั้งเดิม แต่ละชั้น convolution ยกเว้นชั้นแรก (ซึ่งรับอินพุต) จะได้รับเอาต์พุตของเลเยอร์ convolution จากชั้นก่อนหน้า และสร้างแมปคุณลักษณะเอาต์พุตที่ส่งต่อไปยังชั้นถัดไป ดังนั้นจึงมีการเชื่อมต่อเลเยอร์แต่ละชั้น

อย่างไรก็ตามเมื่อจำนวนเลเยอร์ใน CNN เพิ่มขึ้น เช่น เมื่อชั้นลึกมากขึ้นๆ ปัญหา vanishing gradient ก็จะมีมากขึ้น ซึ่งหมายความว่าเมื่อชั้นของข้อมูลจากอินพุตไปยังเลเยอร์เอาต์พุตมีเพิ่มมากขึ้นก็อาจทำให้ข้อมูลบางอย่างหายไป ซึ่งทำให้การเรียนรู้ของโมเดลไม่มีประสิทธิภาพเท่าที่ควร DenseNet แก้ปัญหานี้โดยการปรับเปลี่ยนสถาปัตยกรรม CNN แบบดั้งเดิม โดยทำให้มีรูปแบบการ

เชื่อมต่อระหว่างเลเยอร์ที่ง่ายขึ้น ในสถาปัตยกรรม DenseNet แต่ละเลเยอร์จะเชื่อมต่อโดยตรงกับทุกเลเยอร์

โครงสร้างหลักของ DenseNet

- Connectivity ในแต่ละเลเยอร์ feature map ของเลเยอร์ก่อนหน้าทั้งหมดจะไม่ถูกรวมเข้าด้วยกัน แต่จะเชื่อมต่อเข้าด้วยกันและใช้เป็นอินพุต ดังนั้น DenseNet จึงต้องการพารามิเตอร์น้อยกว่า CNN แบบเดิม และทำให้สามารถใช้คุณลักษณะซ้ำได้ เนื่องจากมีการนำ feature map ที่ซ้ำซ้อนออก
- DenseBlocks การใช้การเชื่อมต่อข้อมูลไม่สามารถทำได้เมื่อขนาดของ feature map เปลี่ยนไป อย่างไรก็ตามส่วนสำคัญของ CNN คือการ down-sampling เลเยอร์ซึ่งลดขนาดของ feature map ผ่านการลดมิติ เพื่อให้ได้ความเร็วในการคำนวณที่ไวมากขึ้น เพื่อที่จะทำการบรรณการนี้ได้ DenseNets จึงถูกแบ่งออกเป็น DenseBlocks โดยที่ขนาดของ feature map จะยังคงคงที่ภายในบล็อก แต่จำนวนตัวกรองระหว่างนั้นจะเปลี่ยนไป เลเยอร์ระหว่างบล็อกเรียกว่า Transition layer ซึ่งลดจำนวน channel ลงเหลือครึ่งหนึ่งของเดิม
- Growth Rate ขนาดของ feature map จะเพิ่มขึ้นหลังจากผ่านแต่ละ dense layer โดยแต่ละเลเยอร์ที่เพิ่มขึ้นจะมีการเพิ่มคุณลักษณะจากเดิมที่มีอยู่ไปเรื่อย ๆ โดยเราจะเรียกพารามิเตอร์นี้ว่า Growth rate ของเครือข่าย

โครงสร้าง DenseNet121 ประกอบไปด้วย

Convolution kernel ขนาด 7×7 64 kernel ที่เลื่อนโดยใช้ค่า stride เท่ากับ 2

Pooling 3×3 max pooling โดยเลื่อนโดยใช้ค่า stride เท่ากับ 2

Dense Block 1 มี 2 convolution layer มี kernel ขนาด 1×1 และ 3×3 ทำซ้ำทั้งหมด 6 ครั้ง

Transition layer 1 มี 1 convolution layer และ 1 average pooling layer

Dense Block 2 มี 2 convolution layer มี kernel ขนาด 1×1 และ 3×3 ทำซ้ำทั้งหมด 12 ครั้ง

Transition layer 2 มี 1 convolution layer และ 1 average pooling layer

Dense Block 3 มี 2 convolution layer มี kernel ขนาด 1×1 และ 3×3 ทำซ้ำทั้งหมด 24 ครั้ง

Transition layer 3 มี 1 convolution layer และ 1 average pooling layer

Dense Block 4 มี 2 convolution layer มี kernel ขนาด 1×1 และ 3×3 ทำซ้ำทั้งหมด 16 ครั้ง

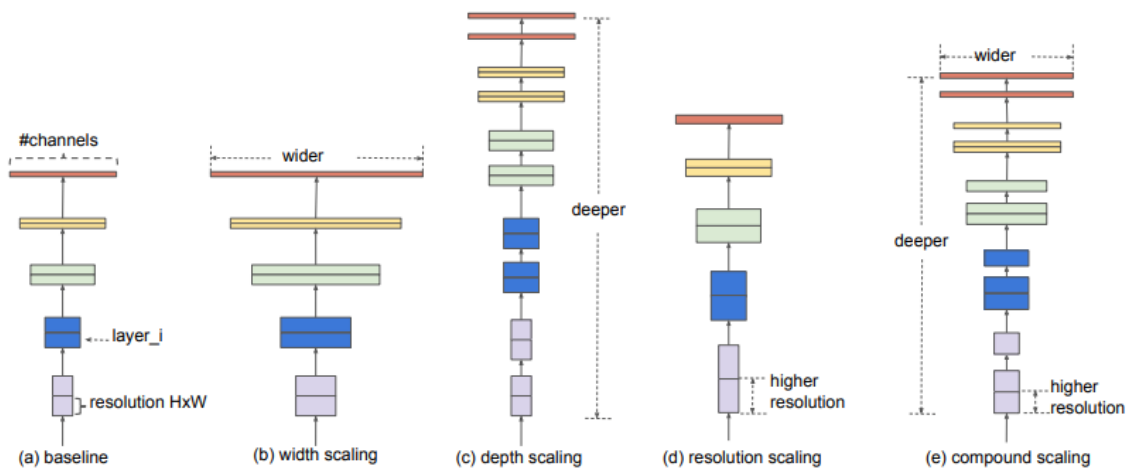
Output layer ใช้ 7×7 Global Average Pooling และเข้าสู่ชั้น fully connected layer ที่ประกอบไปด้วย 1000 nodes ผ่าน activation function คือ softmax function เพื่อให้ได้ผลลัพธ์ว่าเป็นคลาสใด

Layers	Output Size	DenseNet-121	DenseNet-169	DenseNet-201	DenseNet-264
Convolution	112×112	7×7 conv, stride 2			
Pooling	56×56	3×3 max pool, stride 2			
Dense Block (1)	56×56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition Layer (1)	56×56	1×1 conv			
	28×28	2×2 average pool, stride 2			
Dense Block (2)	28×28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition Layer (2)	28×28	1×1 conv			
	14×14	2×2 average pool, stride 2			
Dense Block (3)	14×14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 64$
Transition Layer (3)	14×14	1×1 conv			
	7×7	2×2 average pool, stride 2			
Dense Block (4)	7×7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$
Classification Layer	1×1	7×7 global average pool			
		1000D fully-connected, softmax			

ภาพที่ 8 แสดงสถาปัตยกรรมของ DenseNet

3.3.4 EfficientNetB0

EfficientNet เกิดมาจากทีมพัฒนาที่แต่เดิม CNN นั้นมักถูกพัฒนาและนำไปใช้บนทรัพยากรที่มีจำกัด จึงได้มีการปรับขนาดโมเดลอย่างเป็นระบบ ใช้การทำ scaling model โดยพิจารณาองค์ประกอบทั้ง ความลึก ความกว้าง และความละเอียด ซึ่งช่วยให้ประสิทธิภาพโมเดลดีขึ้น การค้นหาสถาปัตยกรรมโครงข่ายประสาทเทียมเพื่อออกแบบเครือข่ายพื้นฐานแบบใหม่ และทำ scaling เพื่อให้ได้กลุ่มโมเดลที่เรียกว่า EfficientNet ซึ่งมีความแม่นยำและประสิทธิภาพที่ดีกว่าโมเดล CNN รุ่นก่อน โดยจะเน้นไปที่การทำ Compound scaling โดยการ scaling 3 อย่างคือ ความลึก ความกว้าง ความละเอียดของภาพ โดยการใช้ค่าสัมประสิทธิ์เข้ามาช่วยปรับให้ได้จุดที่สมดุลที่สุดในการ scaling ตามภาพที่ 9



ภาพที่ 9 แสดงการเพิ่มประสิทธิภาพโมเดล EfficientNet โดยการทำ Scaling แบบต่าง ๆ

ตารางที่ 1 โครงข่ายประสาทเทียม EfficientNetB0 baseline

Stage i	Operator \mathcal{F}_i	Resolution $H_i \times W_i$	#Channels C_i	#Layers L_i
1	Conv3x3	224×224	32	1
2	MBCConv1, k3x3	112×112	16	1
3	MBCConv6, k3x3	112×112	24	2
4	MBCConv6, k5x5	56×56	40	2
5	MBCConv6, k3x3	28×28	80	3
6	MBCConv6, k5x5	14×14	112	3
7	MBCConv6, k5x5	14×14	192	4
8	MBCConv6, k3x3	7×7	320	1
9	Conv1x1 & Pooling & FC	7×7	1280	1

จากตารางที่ 1 จะเห็นได้ว่าโครงสร้างของ EfficientNetB0 จะเริ่มจาก simple convolution layer และผ่าน 7 block ของ MBCConv และเข้าสู่ชั้น output layer เพื่อให้ได้ผลลัพธ์ว่าเป็นคลาสใด

3.4 การประเมินแบบจำลอง

ในการประเมินแบบจำลองที่ใช้ในการทำนายวัณโรคปอด จะใช้การจำแนกทางสถิติคือ Confusion matrix เข้ามาช่วยหรือที่เรียกว่า error matrix เป็นรูปแบบตารางเฉพาะที่ช่วยให้มองเห็นประสิทธิภาพของอัลกอริทึมหรือแบบจำลองที่เลือกใช้ ซึ่งโดยทั่วไปแล้วจะเป็นการเรียนรู้แบบ

supervised learning แต่ละแถวของเมทริกซ์แสดงตัวอย่างในคลาสที่เป็นการทำนายของโมเดล ในขณะที่แต่ละคอลัมน์แสดงถึงตัวอย่างที่เป็นคลาสจริง

Confusion matrix คือตารางสรุปที่ใช้ในการประเมินประสิทธิภาพของ classification model จำนวนการทำนายที่ถูกและผิดจะสรุปด้วยค่าการนับและแยกตามแต่ละคลาส โดยโครงสร้างของ Confusion matrix จะเป็นตารางขนาด 2x2 โดย

Positive (P): การสังเกตผลเป็นบวก

Negative (N): การสังเกตผลเป็นลบ

True Positive (TP): ผลลัพธ์ที่โมเดลทำนายคลาสบวกได้อย่างถูกต้อง

True Negative (TN): ผลลัพธ์ที่โมเดลทำนายคลาสลบได้อย่างถูกต้อง

False Positive (FP): เรียกอีกอย่างว่า type 1 error ซึ่งเป็นผลลัพธ์ที่แบบจำลองคาดการณ์ผลเป็นคลาสบวกซึ่งไม่ถูกต้อง ทั้งที่ความเป็นจริงต้องเป็นคลาสลบ

False Negative (FN): เรียกอีกอย่างว่า type 2 error ซึ่งเป็นผลลัพธ์ที่แบบจำลองคาดการณ์ผลเป็นคลาสลบซึ่งไม่ถูกต้อง ทั้งที่ความเป็นจริงต้องเป็นคลาสบวก

ตารางที่ 2 รายละเอียดของตาราง Confusion matrix

Confusion matrix		Actual class	
		True	False
Model-predicted class	True	TP	FP
	False	FN	TN

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

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

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

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

Accuracy คือ สัดส่วนของจำนวนคำทำนายทั้งหมดที่ถูกต้อง

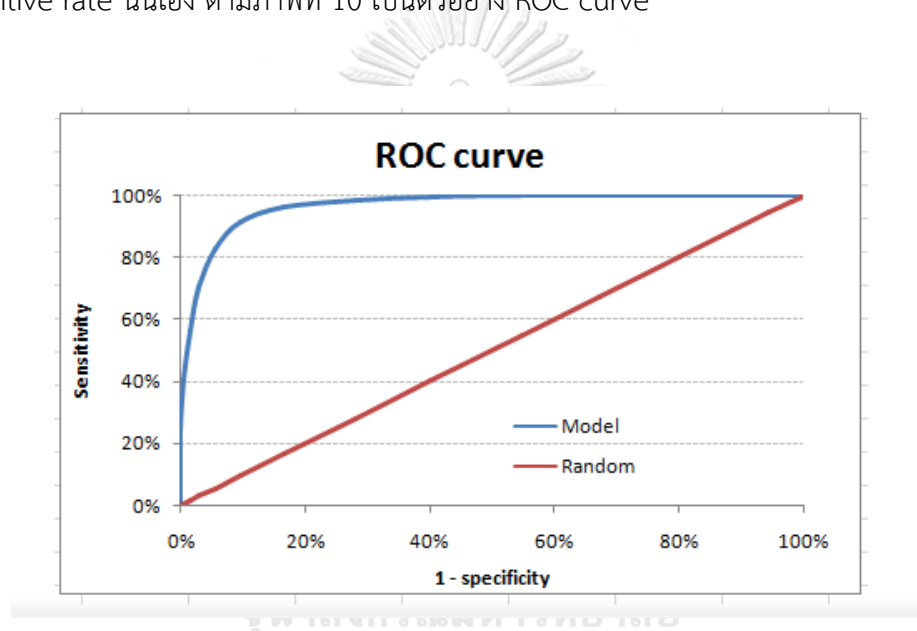
Precision คือ สัดส่วนของกรณีบวกที่เกิดขึ้นจริงต่อกรณีที่เป็นบวกทั้งหมด

Sensitivity or Recall คือ สัดส่วนของกรณีบวกที่เกิดขึ้นจริงซึ่งระบุได้อย่างถูกต้อง

Specificity คือ สัดส่วนของกรณีเชิงลบที่เกิดขึ้นจริงซึ่งระบุได้อย่างถูกต้อง

พื้นที่ใต้เส้นโค้ง ROC (AUC - ROC)

ROC (Receiver Operation Characteristic) เป็นการพล็อตระหว่าง sensitivity และ 1-specificity โดยที่ 1-specificity นั้นก็คือ false positive rate และ sensitivity นั้นก็คือ True Positive rate นั่นเอง ตามภาพที่ 10 เป็นตัวอย่าง ROC curve



ภาพที่ 10 ROC Curve

บทที่ 4

การทดลองและผลการทดลอง

4.1 ระบบปฏิบัติการและภาษาที่ใช้ในการทดลอง

การสร้างแบบจำลองโครงข่ายประสาทเทียมสังวัตนาการ (CNN) พัฒนบน keras open-source deep learning framework และ TensorFlow 2.6.0 as the backend ใช้ภาษา python 3.7.11 และฮาร์ดแวร์ GPU: Tesla T4 16 GB GDDR6

4.2 ชุดข้อมูลที่ใช้ในการทดลอง

ข้อมูลที่ใช้ในการทดลองเป็นข้อมูลภาพเอกซเรย์ที่รวบรวมจากฐานข้อมูลเปิด 2 ชุด คือ Montgomery 138 ภาพ Shenzen 662 ภาพ และภาพของกวงโจวโรค กรมควบคุมโรคอีก 1 ชุด จำนวน 943 ภาพ (โดยได้ทำการนิรนามข้อมูลแล้ว) รายละเอียดข้อมูลแต่ละฐานข้อมูลเป็นดังนี้

ตารางที่ 3 รายละเอียดของภาพเอกซเรย์ที่รวมแยกตามฐานข้อมูล

ฐานข้อมูล	เอกซเรย์ปกติ	เอกซเรย์ที่เป็นวัณโรค	รวม	ความละเอียดภาพ
Montgomery	80	58	138	4020 x 4020, 4892 x 4892
Shenzen	326	336	662	948 x 1130, 3001 x 3007
กวงโจวโรค	438	505	943	3072 x 3072
รวม	844	899	1743	

4.3 การเตรียมชุดข้อมูล

เพื่อทำการทดลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) ด้วยสถาปัตยกรรมแบบ VGG16, DenseNet121, ResNet50, EfficientNetB0 นี้มีการเตรียมชุดข้อมูลด้วยวิธี Data augmentation ดังต่อไปนี้

4.3.1 การปรับภาพให้เป็นสีขาว-ดำ (Image adjustment to be black and white)

ก่อนการเรียนรู้ของแบบจำลองในการทำนายวัณโรคปอดจากภาพเอกซเรย์ทรวงอก จะทำการนำชุดข้อมูลที่นำมาใช้มาปรับสีที่ได้ให้อยู่ในช่วงสีที่เป็นขาวและดำ

4.3.2 การปรับความสว่างของภาพ (Random brightness adjustment)

แบบจำลองในการทำนายวัณโรคปอดจากภาพเอกซเรย์ทรวงอกที่สว่างมากหรือน้อยกว่าภาพปกติ จะทำการนำชุดข้อมูลที่นำมาใช้มาปรับความสว่างของภาพแต่ละภาพด้วยการสุ่มความสว่างระหว่าง มีดลง 20% ถึงสว่างขึ้น 20 %

4.3.3 การหมุนภาพ (Random rotation range)

แบบจำลองในการทำนายวัณโรคปอดจากภาพเอกซเรย์ทรวงอกที่บิดเบี้ยวจากภาพปกติ จะทำการนำชุดข้อมูลที่นำมาใช้มาหมุนภาพแต่ละภาพด้วยการสุ่มมุมระหว่าง -10 องศา ถึง 10 องศา

4.3.4 การหดหรือขยายภาพ (Random scaling zoom range)

แบบจำลองในการทำนายวัณโรคปอดจากภาพเอกซเรย์ทรวงอกที่เป็นภาพขยายกว่าภาพปกติ จะทำการนำชุดข้อมูลที่นำมาใช้มาปรับระยะของแต่ละภาพด้วยการสุ่มกำลังขยายจาก -10% ถึง 10 %

4.3.5 การกลับภาพซ้าย-ขวา (Horizontal image flipping)

แบบจำลองในการทำนายวัณโรคปอดจากภาพเอกซเรย์ทรวงอกที่เป็นภาพกลับข้าง จะทำการนำชุดข้อมูลที่นำมาใช้มาหมุนภาพในแนวนอน 1 ครั้งเพื่อกลับซ้าย-ขวา

4.5 การแบ่งข้อมูลเพื่อให้โมเดลเรียนรู้

ในการทดลองข้อมูลภาพเอกซเรย์ทรวงอกแต่ละชุด จะแบ่งข้อมูลออกเป็น 3 ชุดคือ Training dataset, Validation dataset และ Testing dataset โดยแบ่งออกเป็นชุดข้อมูลสำหรับการสอน (Training dataset) ของแบบจำลอง 70% ของจำนวนข้อมูลทั้งหมด ชุดข้อมูลสำหรับการตรวจสอบ (Validation dataset) ของแบบจำลอง 10% ของจำนวนข้อมูลทั้งหมด และข้อมูลสำหรับการทดสอบ (Test dataset) ของแบบจำลอง 20% ของจำนวนข้อมูลทั้งหมด

4.6 การดำเนินการทดลองและผลที่ได้ในแต่ละแบบจำลอง

การดำเนินการทดลองในงานวิจัยนี้จะแบ่งเป็น 2 ช่วง ในช่วงแรกนี้จะเป็นการทดลองเพื่อหาโมเดลที่ดีที่สุดในการทำนายผลที่เหมาะสมกับการภาพเอกซเรย์ปอด เพราะเนื่องจากในแต่ละสถาปัตยกรรมจะให้ผลที่ดีกับภาพในแต่ละแบบที่ไม่เหมือนกัน ดังนั้นเราจึงต้องทดสอบภาพเอกซเรย์

ปอดซึ่งเป็นสีขาว ดำ เทาเป็นส่วนใหญ่ และมีการใช้การถ่ายทอดการเรียนรู้มาเพื่อช่วยลดระยะเวลาในการเรียนรู้ของแต่ละแบบจำลอง โดยคณะวิจัยได้เลือกใช้แบบจำลองซึ่งเป็นที่นิยมได้แก่ VGG16, ResNet50, DenseNet121, EfficientNetB0 เพื่อเปรียบเทียบดูว่าโมเดลใดที่จะเหมาะสมที่สุดใน การนำมาแยกภาพเอกซเรย์ โดยชุดข้อมูลที่คณะวิจัยได้นำมาใช้ในการเรียนรู้ของแบบจำลองจะนำมา จาก Opensource 2 dataset คือภาพจาก Montgomery และ Shenzen ที่เป็นชุดข้อมูลที่นิยมใช้ สำหรับภาพเอกซเรย์วัณโรคปอด โดยในการทดลองต่อไปนั้นจะเป็นการนำเอาแบบจำลองที่ประเมิน แล้วว่าดีที่สุด มาทำการเรียนรู้จากชุดข้อมูลจาก 3 แหล่งคือ Montgomery Shenzen และกองวัณ โรค ให้แบบจำลองได้ทำการเรียนรู้และทดสอบดูประสิทธิภาพว่า เมื่อเพิ่มชุดข้อมูลเข้าไปแล้วจะทำให้ แบบจำลองสามารถทำงานได้ดีขึ้นหรือไม่ แล้วหลังจากนั้นจะนำแบบจำลองที่เทรนจากทั้ง 3 แหล่ง ลงมาทดสอบแยกแต่ละฐานข้อมูลเพื่อดูว่ามีความแตกต่างกันอย่างไรบ้าง โดยรายละเอียดแต่ละการ ทดลองในแต่ละสถาปัตยกรรมแบบจำลองต่าง ๆ เป็นดังนี้

การทดลองนี้ได้ใช้ภาพจาก Montgomery และ Shenzen จำนวน 800 ภาพ ทำ data augmentation แล้วแบ่งข้อมูลของการเรียนรู้ออกเป็นโดยแบ่งออกเป็นชุดข้อมูลสำหรับการสอน 70% ของจำนวนข้อมูลทั้งหมด ชุดข้อมูลสำหรับการตรวจสอบ 10% ของจำนวนข้อมูลทั้งหมด และ ข้อมูลสำหรับการทดสอบ 20% ของจำนวนข้อมูลทั้งหมด โดยตั้งค่าพารามิเตอร์ adam optimizer ด้วย learning rate 0.0001 batch size 32 และสอนโมเดลจำนวน 15 epoch แล้วนำแบบจำลองที่ สร้างจากชุดข้อมูลสอนแล้วนำไปทดสอบค่าความถูกต้อง (Accuracy) รวมถึงค่าความแม่นยำ (Precision) ความไว (Sensitivity) และค่าพื้นที่ใต้กราฟ ROC (AUC-ROC) ของแบบจำลอง VGG16, ResNet50, DenseNet121, EfficientNetB0 ที่เรียนรู้ด้วยชุดข้อมูลจากทั้ง 2 dataset จะทำนายกับ ชุดข้อมูลทดสอบได้ผลลัพธ์ดังตาราง

ตารางที่ 4 แสดง Confusion matrix ของแบบจำลอง VGG16

Confusion matrix: VGG16		Actual class	
		Normal	TB
Predicted class	Normal	34	47
	TB	7	72

การประเมินประสิทธิภาพของแบบจำลอง VGG16

จากการนำแบบจำลอง VGG16มาใช้กับชุดข้อมูลสำหรับการตรวจสอบ เมื่อนำมาคำนวณหาค่าความถูกต้องจะได้ 66% ความแม่นยำ 72% ความไว 67% พื้นที่ใต้กราฟ ROC 90.27%

ตารางที่ 5 แสดง Confusion matrix ของแบบจำลอง ResNet50

Confusion matrix: ResNet50		Actual class	
		Normal	TB
Predicted class	Normal	55	26
	TB	10	69

การประเมินประสิทธิภาพของแบบจำลอง ResNet50

จากการนำแบบจำลอง ResNet50 มาใช้กับชุดข้อมูลสำหรับการตรวจสอบ เมื่อนำมาคำนวณหาค่าความถูกต้องจะได้ 78% ความแม่นยำ 79% ความไว 78% พื้นที่ใต้กราฟ ROC 92.14%

ตารางที่ 6 แสดง Confusion matrix ของแบบจำลอง DenseNet121

Confusion matrix: DenseNet121		Actual class	
		Normal	TB
Predicted class	Normal	66	15
	TB	12	67

การประเมินประสิทธิภาพของแบบจำลอง DenseNet121

จากการนำแบบจำลอง DenseNet121 มาใช้กับชุดข้อมูลสำหรับการตรวจสอบ เมื่อนำมาคำนวณหาค่าความถูกต้องจะได้ 83% ความแม่นยำ 83% ความไว 83% พื้นที่ใต้กราฟ ROC 92.85%

ตารางที่ 7 แสดง Confusion matrix ของแบบจำลอง EfficientNetB0

Confusion matrix: EfficientNetB0		Actual class	
		Normal	TB
Predicted class	Normal	58	23
	TB	13	66

การประเมินประสิทธิภาพของแบบจำลอง EfficientNetB0

จากการนำแบบจำลอง EfficientNetB0 มาใช้กับชุดข้อมูลสำหรับการตรวจสอบ เมื่อนำมาคำนวณหาค่าความถูกต้องจะได้ 78% ความแม่นยำ 78% ความไว 78% พื้นที่ใต้กราฟ ROC 91.35%

ตารางที่ 8 แสดงประสิทธิภาพของสถาปัตยกรรมแบบจำลองต่าง ๆ

MODEL	PRECISION	RECALL	ACCURACY	AUC-ROC
VGG16	72	67	66	90.27
RESNET50	79	78	78	92.14
DENSENET121	83	83	83	92.85
EFFICIENTNETB0	78	78	78	91.35

4.7 ผลการทดลองของแบบจำลองต่าง ๆ

จากตารางที่ 8 จะเห็นได้ว่าสถาปัตยกรรมแบบจำลองที่มีประสิทธิภาพสูงที่สุดในการทำนาย วัณโรคปอดคือ แบบจำลอง DenseNet121 ซึ่งจะเห็นได้ว่าทำคะแนนได้ดีกว่า EfficientNetB0 ที่ขณะนี้ เป็นสถาปัตยกรรมที่ใช้แยกรูปภาพที่มีสีใน Imagenet dataset ได้สูงที่สุด ซึ่งโมเดลรุ่นเก่า ResNet50 ก็สามารถทำได้ดีกว่าเล็กน้อย ส่วน VGG16 ที่มีจำนวนชั้นน้อยที่สุดก็มีประสิทธิภาพต่ำที่สุด

เราจะเลือกโมเดลที่มีประสิทธิภาพมากที่สุดในกลุ่มนี้คือ DenseNet121 มาใช้ทดสอบกับฐานข้อมูลทั้ง 3 แหล่งที่มีและแยกแต่ละแหล่งเพื่อดูว่ามีความเหมือนหรือต่างกันอย่างไร โดยโครงสร้างหลักของแบบจำลอง DenseNet121ที่เรานำมาใช้ในการทดลองต่อไป เป็นดังภาพที่ 11

การพิจารณาเลือกจำนวนชั้นของแบบจำลอง DenseNet

ผู้วิจัยได้ทำการทดลองกับแบบจำลอง DenseNet ตามจำนวนชั้นที่ต่างกัน แต่ใช้ค่าพารามิเตอร์ที่เหมือนกันในชุดข้อมูลเดิม เพื่อทดสอบว่าจำนวนชั้นที่เพิ่มขึ้นนั้นจะช่วยให้ประสิทธิภาพของโมเดลเพิ่มขึ้นหรือไม่ โดยได้ผลดังตารางที่ 9

ตารางที่ 9 แสดงประสิทธิภาพของ DenseNet ที่จำนวนชั้นต่าง ๆ กัน

MODEL	PRECISION	RECALL	ACCURACY	AUC-ROC
DENSENET121	83	83	83	92.85
DENSENET169	82	82	82	91.71
DENSENET260	82	81	82	90.12

จะเห็นได้ว่าการเพิ่มจำนวนชั้นของแบบจำลอง DenseNet ที่เพิ่มขึ้นกับภาพเอกซเรย์ปอด ไม่ได้ช่วยให้ประสิทธิภาพของแบบจำลองดีขึ้น และมีแนวโน้มจะลดลงเล็กน้อย รวมทั้งเวลาในการทำงานของโมเดลใช้เวลานานกว่า ดังนั้นในการทดลองขั้นต่อไปที่มีการเพิ่มชุดข้อมูลเราจะใช้แบบจำลอง DenseNet121 ในการทดสอบกับชุดข้อมูล

การไม่ทำ transfer learning เพื่อทดสอบกับภาพเอกซเรย์ปอด

การทดลองนี้จะไม่ใช่ pretrained weight เพื่อดูว่าเมื่อทำการทดลองกับภาพเอกซเรย์ที่เป็นภาพขาวดำนั้น จะสามารถทำการเทรนโดยใช้ค่า random weight โดยทั่วไปได้หรือไม่ ซึ่งจากการทดลองจะได้ผลดังตารางที่ 10

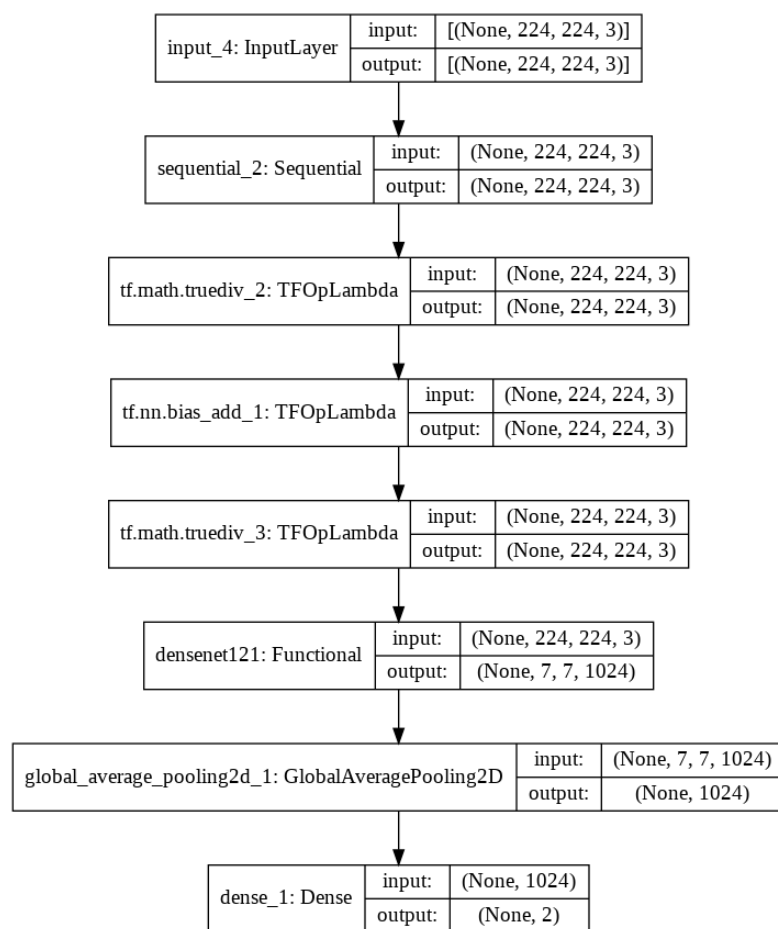
ตารางที่ 10 แสดงประสิทธิภาพของแบบจำลองที่ไม่ได้ทำ transfer learning

MODEL	PRECISION	RECALL	ACCURACY	AUC-ROC
DENSENET121	25	48	49	49.31

จะเห็นได้ว่าการไม่ใช้ pretrained weight จะทำให้ประสิทธิภาพของแบบจำลองลดลงได้ ดังนั้นการใช้ transfer learning กับภาพเอกซเรย์ปอดก็ช่วยให้ประสิทธิภาพของแบบจำลองดีขึ้นได้

4.8 การนำแบบจำลอง DenseNet121 มาทดสอบกับฐานข้อมูล

จากการทดลองที่ผ่านมาเราได้สถาปัตยกรรมแบบจำลองที่ทำงานบนชุดข้อมูลภาพเอกซเรย์ทรวงอกได้ดีคือ DenseNet121 ในขั้นตอนต่อมาเราจะนำแบบจำลองนี้มาเพิ่มชุดข้อมูลในการเรียนรู้ และนำแบบจำลองใหม่ที่ได้ไปใช้ทดสอบดูในแต่ละชุดฐานข้อมูลเพื่อใช้ในการดูประสิทธิภาพว่ามีความเหมือนหรือแตกต่างกันอย่างไร



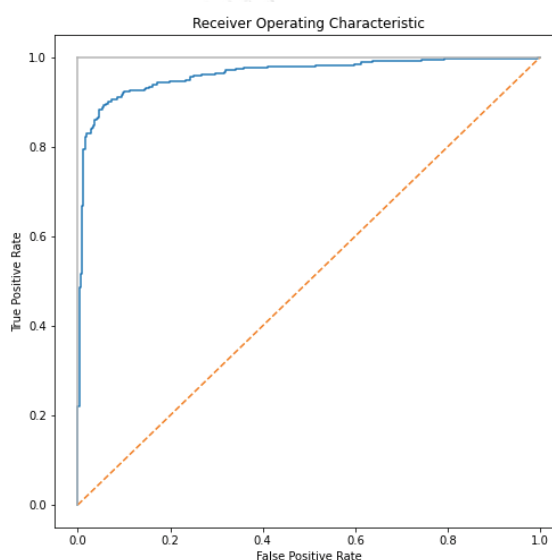
ภาพที่ 11 แสดงโครงสร้างของแบบจำลอง DenseNet121

ฐานข้อมูล 3 แหล่ง

การทดลองนี้ได้ใช้ภาพจาก Montgomery Shenzen และกองวิมโรค กรมควบคุมโรค จำนวน 1743 ภาพ ทำ data augmentation แล้วแบ่งข้อมูลของการเรียนรู้ออกเป็นโดยแบ่งออกเป็นชุดข้อมูลสำหรับการสอน 70% ของจำนวนข้อมูลทั้งหมด ชุดข้อมูลสำหรับการตรวจสอบ 10% ของจำนวนข้อมูลทั้งหมด และข้อมูลสำหรับการทดสอบ 20% ของจำนวนข้อมูลทั้งหมด โดยตั้งค่าพารามิเตอร์ adam optimizer ด้วย learning rate 0.0001 batch size 32 และสอนโมเดลจำนวน 15 epoch แล้วนำแบบจำลองที่สร้างจากชุดข้อมูลสอนแล้วนำไปทดสอบค่าความถูกต้อง (Accuracy) รวมถึงค่าความแม่นยำ (Precision) และความไว (Sensitivity) โดยใช้แบบจำลอง DenseNet121 ที่เลือกมาแล้ว มาใช้ทำนายกับชุดข้อมูลทดสอบได้ผลลัพธ์ดังตารางที่ 11

ตารางที่ 11 แสดง Confusion matrix ของแบบจำลอง DenseNet121

Confusion matrix: 3 sources		Actual class	
		Normal	TB
Predicted class	Normal	159	10
	TB	20	160



ภาพที่ 12 ROC Curve ของ DenseNet121 บน 3 dataset

การประเมินประสิทธิภาพบนฐานข้อมูล 3 แหล่ง

จากการนำแบบจำลอง DenseNet121มาใช้กับชุดข้อมูลสำหรับการตรวจสอบ เมื่อนำมาคำนวณหาค่าความถูกต้องจะได้ 91% ความแม่นยำ 91% ความไว 91% พื้นที่ใต้กราฟ ROC 95.63% ซึ่งจะเห็นได้ว่าเมื่อเพิ่มขนาดของชุดข้อมูลในการเรียนรู้ไปอีกเท่าตัว ก็จะช่วยเพิ่มประสิทธิภาพให้กับแบบจำลองได้ จากเดิมที่ค่าความถูกต้อง 83% ความแม่นยำ 83% ความไว 83% พื้นที่ใต้กราฟ ROC 92.85%

ฐานข้อมูลแต่ละแหล่ง

การทดลองนี้ได้แบบจำลอง DenseNet121 ที่ได้รับการสอนโดยชุดสอนจากทั้ง 3 แหล่งข้อมูล แล้วนำแบบจำลองไปทดสอบในชุดข้อมูลทดสอบแยกแต่ละฐานข้อมูล แล้วนำไปทดสอบ

ค่าความถูกต้อง (Accuracy) รวมถึงค่าความแม่นยำ (Precision) และความไว (Sensitivity) มาใช้ทำนายกับชุดข้อมูลทดสอบได้ผลลัพธ์ดังตารางที่ 12

ตารางที่ 12 แสดงประสิทธิภาพของ DenseNet121 บนฐานข้อมูลแต่ละฐาน

แหล่งข้อมูล	PRECISION	RECALL	ACCURACY	AUC-ROC
MONTGOMERY	88	86	86	93.29
SHENZEN	94	90	92	94.66
กองวิณโรค	99	94	96	96.85
3 SOURCES	91	91	91	95.63

จะเห็นได้ว่า การประเมินประสิทธิภาพบนฐานข้อมูลแต่ละแหล่งนั้น ฐานข้อมูลกองวิณโรคได้ประสิทธิภาพสูงที่สุดทั้งความถูกต้อง ความแม่นยำ และความไว โดยฐานข้อมูลที่ได้ประสิทธิภาพรองลงมาคือ Shenzen และประสิทธิภาพของแบบจำลองที่ทำได้น้อยที่สุดคือบนฐานข้อมูล Montgomery

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

4.9 การกำหนดค่าความไวให้เป็น 100%

เป็นการทำการหาค่า threshold ที่ทำให้ได้ค่า sensitivity หรือ recall ที่ 100% วัตถุประสงค์นี้ทำไปเพื่อให้โมเดลสามารถทำคัดกรองคนที่ป่วยเป็นวิณโรคได้ทุกคน โดยผู้วิจัยได้ให้โมเดลคำนวณหาค่า threshold ที่เหมาะสมและ fix ให้ค่า sensitivity ที่ 100% จะพบว่าประสิทธิภาพของแบบจำลองได้ตามตารางที่ 13

```
def find_100_recall_threshold(y_score, y_true):
    y_score = y_score.flatten()
    y_true = y_true.flatten()
```

```

y_score_pos = y_score[y_true == 1]
thre = min(y_score_pos)
return thre

```

ตารางที่ 13 แสดงประสิทธิภาพของแบบจำลองที่กำหนดค่า Recall 100%

	CLASS	PRECISION	RECALL	ACCURACY	AUC-ROC
3 SOURCES	0	100	52		
	1	68	100		
	Macro avg	84	76	76	80.14

จะเห็นได้ว่าเมื่อกำหนดให้ค่า sensitivity ของคนที่เป็นวัณโรคเป็น 100% จะทำให้ความแม่นยำของคนปกติเพิ่มขึ้นเป็น 100% ด้วยเช่นกัน แต่ค่าความแม่นยำรวมจะลดลงจาก 91% เหลือ 84% ค่าความไว ลดลงจาก 91% เหลือ 76% ค่าความถูกต้องลดลงจาก 91% เหลือ 76% ค่าพื้นที่ใต้กราฟ ROC ลดลงจาก 95.63% เหลือ 80.14%

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

จากการทดลองเพื่อเปรียบเทียบประสิทธิภาพและผลลัพธ์ที่ได้จากโครงข่ายประสาทเทียมแบบสังวัตนาการด้วยวิธีการถ่ายทอดความรู้โมเดลแบบต่าง ๆ เพื่อการทำนายวัณโรคปอดจากภาพเอกซเรย์ทรวงอก โดยใช้ชุดข้อมูลเปิดจากฐานข้อมูลของ Montgomery และ Shenzen ที่นิยมใช้สำหรับวัณโรคปอด และชุดข้อมูลจากกรมวัณโรค กองควบคุมโรค โดยนำฐานข้อมูลเปิดมาใช้เป็นข้อมูลในการเรียนรู้ของสถาปัตยกรรมแบบจำลองต่าง ๆ โดยคณะวิจัยเลือกได้เลือกใช้แบบจำลองซึ่งเป็นที่ยอมรับได้แก่ VGG16, ResNet50, DenseNet121, EfficientNetB0 เพราะเนื่องจากในแต่ละสถาปัตยกรรมของแต่ละแบบจำลองจะให้ผลที่ดีกับภาพที่ไม่เหมือนกัน โดยงานวิจัยนี้จะเป็นภาพจากการเอกซเรย์ทรวงอก และมีรอยโรควัณโรคอยู่ในปอด โดยแบบจำลองต้องการรูปแบบที่เป็นรอยสีขาวในภาพเอกซเรย์เพื่อการทำนายผล ซึ่งจากผลการทดลองนั้นพบว่า แบบจำลองที่มีประสิทธิภาพในการทำนายวัณโรคปอดนั้นเรียงจากมากไปน้อยคือ DenseNet121 ResNet50 EfficientNetB0 และ VGG16 ซึ่งเป็นแบบจำลองที่ค่อนข้างเก่าและจำนวนชั้นน้อยที่สุดในกลุ่ม แต่กลับพบว่าแบบจำลองที่ใหม่และทำประสิทธิภาพได้ดีในการจำแนกสัตว์ สิ่งของ ซึ่งเป็นภาพถ่ายทั่วไปอย่าง EfficientNetB0 นั้นกลับทำคะแนนได้ไม่ต่างกับ ResNet50 แสดงให้เห็นว่าแม้ว่าสถาปัตยกรรมใหม่ที่ออกแบบมาอย่างดีและใช้การทำงานที่ซับซ้อน ก็อาจจะเหมาะที่จะใช้กับงานบางชนิดจึงจำเป็นที่ผู้ใช้ต้องเลือกแบบจำลองให้เหมาะสมกับการใช้งาน

จะเห็นว่าโครงข่ายประสาทเทียมแบบจำลอง Densenet121 มีประสิทธิภาพที่ดีในการใช้งานกับภาพเอกซเรย์ทรวงอก โดยเมื่อนำมาเรียนรู้บนฐานข้อมูลทั้ง 3 แหล่งจะพบว่าประสิทธิภาพในการทำนายวัณโรคเพิ่มขึ้น และเมื่อนำแบบจำลองมาทดสอบบนแต่ละฐานข้อมูลนั้น จะพบว่าการทดสอบบนชุดข้อมูลจากกรมวัณโรค กองควบคุมโรคได้ประสิทธิภาพทั้งความถูกต้อง ความแม่นยำ และความไว ในการทำนายโรคดีที่สุด รองลงมาคือชุดข้อมูล Shenzen และ Montgomery ความแม่นยำคือ 96%, 92%, 86% ตามลำดับ ที่เป็นเช่นนี้นั้นทางคณะวิจัยได้วิเคราะห์ว่าอาจมาจากเหตุผลดังต่อไปนี้ คือ

- 1) อาจมาจากการแบบจำลองมีการเรียนรู้จากฐานข้อมูลของกองวัณโรค กรมควบคุมโรคในปริมาณมากที่สุด จึงทำให้แบบจำลองมีความคุ้นเคยกับชุดข้อมูลของกองวัณโรคได้

2) อาจมาจากการที่รูปแบบของการเป็นวัณโรคในประเทศไทยนั้นค่อนข้างชัดเจน หรือการที่ตรวจพบโรคได้ช้า หรือเชื้อค่อนข้างเจริญได้ดีในแถบนี้จึงทำให้รูปแบบของรอยโรคมีความชัดเจนมากกว่าในประเทศจีน และอเมริกาที่มีอากาศหนาวกว่า

3) เนื่องจากแพทย์ของประเทศไทยมีความชำนาญในการอ่านผลเอกซเรย์ของผู้ป่วยเป็นวัณโรคได้ดีกว่า เนื่องจากมีประสบการณ์มากกว่าจากความชุกของการเกิดวัณโรคที่มากในแถบประเทศอาเซียน ทำให้แพทย์ได้พบเจอโรคบ่อยและมีความคุ้นเคยในการอ่านฟิล์มได้เก่งกว่าแถบอเมริกา

4) การที่ฐานข้อมูล Montgomery ได้ประสิทธิภาพจากโมเดลที่ค่อนข้างต่ำ อาจเนื่องมาจากภาพที่มีค่อนข้างน้อยทำให้โมเดลเรียนรู้รูปแบบของภาพได้น้อย และรูปที่ค่อนข้างเก่าจึงมีคุณภาพของภาพที่น้อยกว่า

5) ทั้งนี้ยังมาจากเครื่องเอกซเรย์ที่ใช้ด้วย เนื่องจากเครื่องฉายเอกซเรย์แต่ละบริษัทก็จะผลิตและใช้แหล่งกำเนิดรังสีเอกซ์ที่แตกต่างกันไป และเรื่องของตัวรับสัญญาณที่แตกต่างกัน ทำให้มีการกระจายของแสงที่ไม่เหมือนกัน

ประโยชน์ที่ได้รับจากงานวิจัยในการใช้แบบจำลองเพื่อทำนายวัณโรคจากภาพเอกซเรย์ทรวงอก เป็นการที่หน่วยงานนำไปใช้ในการตรวจคัดกรองในชุมชน หรือในบริเวณที่มีความเสี่ยงจะติดเชื้อวัณโรค ซึ่งตรวจเอกซเรย์เป็นจำนวนมาก ทำให้ลดเวลาในการอ่านฟิล์มเพื่อคัดกรองและเพิ่มขีดความสามารถของบุคลากรทางการแพทย์หน้างาน เพื่อแบ่งเบาภาระและมีเวลาไปดูแลผู้ป่วยได้เพิ่มมากขึ้น การที่เราคัดกรองผู้ที่มีความเสี่ยงสูงออกมาได้ก่อนยังช่วยลดโอกาสการแพร่กระจายโรคไปยังคนอื่น ๆ ในชุมชนได้ด้วย อีกทั้งการคัดกรองผู้ที่มีความเสี่ยงสูงออกมาก่อนจะช่วยในการตัดสินใจของบุคลากรทางการแพทย์เพื่อการจัดการบริการหรือการส่งตรวจเพิ่มเติม

จุฬาลงกรณ์มหาวิทยาลัย

CHULALONGKORN UNIVERSITY

5.2 ข้อเสนอแนะ

จากการศึกษาจะพบได้ว่าการหาข้อมูลภาพมาเพิ่มเติมนั้นสามารถช่วยเพิ่มประสิทธิภาพการทำงานให้กับแบบจำลองได้ ดังนั้นการได้ความร่วมมือกันเพื่อหาภาพเอกซเรย์ทรวงอกที่เพิ่มมากขึ้นในแต่ละคลาสโดยมาจากหลายๆแหล่ง และเครื่องเอกซเรย์หลาย ๆ แบบ จะทำให้โมเดลมีความ generalize เพื่อการใช้งานที่สามารถใช้ได้กับหลาย ๆ โรงพยาบาลหรือหลาย ๆ สถานที่ที่ต้องใช้การเอกซเรย์เพื่อทำนายวัณโรคปอด

เนื่องจากแต่ละบริษัทที่ทำให้ภาพของเอกซเรย์ไม่เหมือนกัน จึงยังอาจต้องศึกษาเพิ่มเติมถึงการให้โมเดลมีการคัดกรองภาพที่มีคุณภาพไม่เข้าเกณฑ์เพื่อคัดกรองภาพเหล่านี้ออกไปก่อนการนำเข้าแบบจำลองในการทำนายโรค เพื่อเพิ่มประสิทธิภาพให้กับแบบจำลอง แล้วนำภาพบางส่วนที่ไม่

เข้าเกณฑ์อาจเนื่องจากคุณภาพของภาพไม่ดี หรือมีความสว่างของภาพไม่ดี หรือถ่ายผิดประเภท
นำส่งไปให้กับทางรังสีแพทย์พิจารณาเพื่ออ่านผลโดยแพทย์หรือพิจารณาให้ทำเอกซเรย์ใหม่อีกครั้ง

นอกจากนี้อาจมีการใช้เทคนิคอื่นๆเพื่อมาเพิ่มประสิทธิภาพโดยนำมาประมวลผลกับภาพ
ที่ใช้ เช่น การตัดภาพเฉพาะส่วนของปอด หรือการนำอาการของผู้ป่วยที่ได้มาจากการซักประวัติ การ
นำผลทางห้องปฏิบัติการเข้ามาร่วมช่วยในการตัดสินใจ เพื่อช่วยเพิ่มความแม่นยำให้กับโมเดลในการ
ทำนายได้



บรรณานุกรม

1. Chakaya, J., et al., *Global Tuberculosis Report 2020 - Reflections on the Global TB burden, treatment and prevention efforts*. Int J Infect Dis, 2021. **113 Suppl 1**: p. S7-S12.
2. Zumla, A., et al., *COVID-19 and tuberculosis—threats and opportunities*. The International Journal of Tuberculosis and Lung Disease, 2020. **24**(8): p. 757-760.
3. WHO, *The WHO Global Task Force on TB Impact Measurement*. World Health Organization, 2019.
4. Bhalla, A.S., et al., *Chest tuberculosis: Radiological review and imaging recommendations*. Indian Journal of Radiology and Imaging, 2015. **25**(03): p. 213-225.
5. LeCun, Y., Y. Bengio, and G. Hinton, *Deep Learning*. Nature, 2015. **521**: p. 436-44.
6. Sultana, F., A. Sufian, and P. Dutta, *Advancements in Image Classification using Convolutional Neural Network*. 2018, 2018 Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN). 122-129.
7. R., C., G.K. K., and J.R. C., *Convolutional Neural Network (CNN) for Image Detection and Recognition*. 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC), 2018: p. 278-82.
8. Yamashita, R., et al., *Convolutional neural networks: an overview and application in radiology*. Insights into Imaging, 2018. **9**(4): p. 611-629.
9. Cody, T. and P. Beling, *A Systems Theory of Transfer Learning*. 2021.
10. Pesapane, F., M. Codari, and F. Sardanelli, *Artificial intelligence in medical imaging: threat or opportunity? Radiologists again at the forefront of innovation in medicine*. European Radiology Experimental, 2018. **2**(1).
11. Abiyev, R.H. and M.K.S. Ma'Aitah, *Deep Convolutional Neural Networks for Chest Diseases Detection*. Journal of Healthcare Engineering, 2018. **2018**: p. 1-11.
12. Krizhevsky, A., I. Sutskever, and G. Hinton, *ImageNet Classification with Deep Convolutional Neural Networks*. Neural Information Processing Systems, 2012.

25.

13. K., S. and Z. A., *Very deep convolutional networks for large-scale image recognition*. 2014.
14. He, K., et al. *Deep Residual Learning for Image Recognition*. 2016. IEEE.
15. Huang, G., et al., *Densely Connected Convolutional Networks*. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: p. 2261-2269.
16. Mingxing, T. and Q.V. Le., *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks*. 2019.
17. Jaeger, S., et al., *Automatic Tuberculosis Screening Using Chest Radiographs*. IEEE Transactions on Medical Imaging, 2014. **33**(2): p. 233-245.
18. Hwang, S., et al., *A novel approach for tuberculosis screening based on deep convolutional neural networks*. SPIE Medical Imaging. Vol. 9785. 2016: SPIE.
19. Hashmi, M.F., et al., *Efficient Pneumonia Detection in Chest Xray Images Using Deep Transfer Learning*. Diagnostics, 2020. **10**(6): p. 417.



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ประวัติผู้เขียน

ชื่อ-สกุล	ชัยสิทธิ์ พัฒนาสุวรรณ
วัน เดือน ปี เกิด	18 กรกฎาคม 2528
สถานที่เกิด	กรุงเทพมหานคร
วุฒิการศึกษา	จุฬาลงกรณ์มหาวิทยาลัย
ที่อยู่ปัจจุบัน	163 ถนนสามเสน แขวงวัดสามพระยา เขตพระนคร กรุงเทพมหานคร



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY