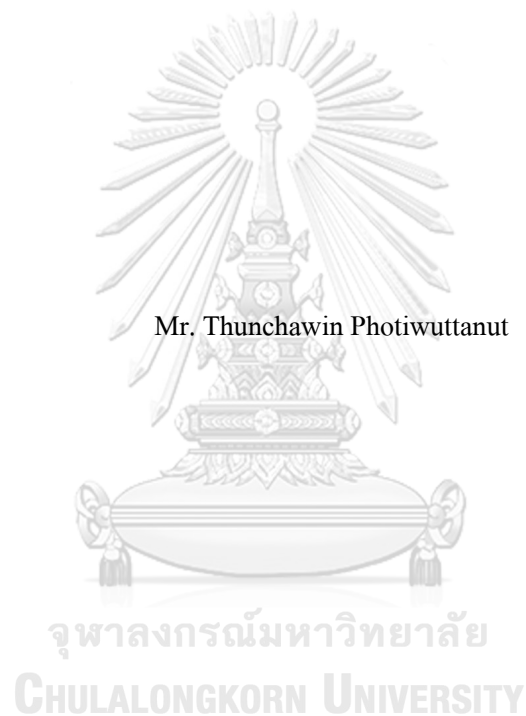


การเรียนรู้การถ่ายทอดสำหรับการจำแนกภาพด้วยโครงข่ายคอนโวลูชัน: กรณีศึกษาภาพถ่ายรังสี  
ทรวงอกของผู้ป่วยที่ติดเชื้อโควิด19



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต  
สาขาวิชาสถิติ ภาควิชาสถิติ  
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย  
ปีการศึกษา 2564  
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

IMAGE TRANSFER LEARNING FOR IMAGE CLASSIFICATION USING  
CONVOLUTIONAL NEURAL NETWORKS: A CASE STUDY OF COVID19-INFECTED  
CHEST X-RAY IMAGES.



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

Department of Statistics

FACULTY OF COMMERCE AND ACCOUNTANCY

Chulalongkorn University

Academic Year 2021

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การเรียนรู้การถ่ายทอดสำหรับการจำแนกภาพด้วยโครงข่าย คอนโวลูชัน: กรณีศึกษาภาพถ่ายรังสีทรวงอกของผู้ป่วยที่ติด เชื้อโควิด19
โดย	นายชญัญชวิน โปธิวัฒน์รัตน์
สาขาวิชา	สถิติ
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ผู้ช่วยศาสตราจารย์ ดร.อักรินทร์ ไพบุญย์พานิช

---

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้แนบวิทยานิพนธ์ฉบับนี้เป็น  
ส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

.....	คณบดีคณะพาณิชยศาสตร์และการ บัญชี
(รองศาสตราจารย์ ดร.วิเลิศ ภูริวัชร)	
คณะกรรมการสอบวิทยานิพนธ์	ประธานกรรมการ
.....	
(ผู้ช่วยศาสตราจารย์ ดร.ภูริพันธุ์ รุจิจักร)	อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
.....	
(ผู้ช่วยศาสตราจารย์ ดร.อักรินทร์ ไพบุญย์พานิช)	กรรมการ
.....	
(ผู้ช่วยศาสตราจารย์ ดร.สุรณพีร์ ภูมิวุฒิสาร)	กรรมการภายนอกมหาวิทยาลัย
.....	
(ผู้ช่วยศาสตราจารย์ ดร.บุญสิทธิ์ ยี่มวาสนา)	

ชัยณัฐวิน โปธิวัฒน์รัตน์ : การเรียนรู้การถ่ายทอดสำหรับการจำแนกภาพด้วยโครงข่ายคอน  
 โวลูชัน: กรณีศึกษาภาพถ่ายรังสีทรวงอกของผู้ป่วยที่ติดเชื้อโควิด19. (IMAGE TRANSFER  
 LEARNING FOR IMAGE CLASSIFICATION USING CONVOLUTIONAL NEURAL  
 NETWORKS: A CASE STUDY OF COVID19-INFECTED CHEST X-RAY IMAGES.)  
 อ.ที่ปรึกษาหลัก : ผศ. ดร.อัครินทร์ ไพบูลย์พานิช

เทคนิคการประมวลผลจากภาพถูกนำมาใช้กันอย่างแพร่หลายในหลากหลายอุตสาหกรรมใน  
 ปัจจุบัน โดยการนำมาประยุกต์ใช้กับทางการแพทย์ก็เป็นอีกหนึ่งอุตสาหกรรมที่ได้รับความนิยม ทั้งนี้  
 ปัญหาในการจำแนกภาพสามารถทำได้หลายวิธีด้วยกัน หนึ่งในนั้น คือการนำการเรียนรู้เชิงลึกมา  
 ประยุกต์ใช้ในการแก้ไขปัญหา โดยการจำแนกประเภทผ่านการเรียนรู้เชิงลึกสามารถแก้ไขได้อย่าง  
 รวดเร็วและแม่นยำผ่านการนำโครงข่ายการเรียนรู้เชิงลึกแบบคอนโวลูชัน หรือ ซีเอ็นเอ็น  
 (Convolutional Neural Networks หรือ CNN) มาใช้กับเทคนิคการเรียนรู้ถ่ายทอด (Transfer  
 Learning) งานวิจัยนี้จึงนำเสนอวิธีการประยุกต์ใช้เทคนิคการเรียนรู้ถ่ายทอดในการฝึกสอนแบบจำลอง  
 โครงข่ายคอนโวลูชันเชิงลึกเพื่อจำแนกภาพถ่ายรังสีทรวงอกออกเป็น 3 ประเภท คือ 1) ภาพถ่ายรังสี  
 ทรวงอกของผู้ป่วยปกติ 2) ภาพถ่ายรังสีทรวงอกของผู้ป่วยที่ติดเชื้อโควิด19 3) ภาพถ่ายรังสีทรวงอก  
 ของผู้ติดเชื้อปอดอักเสบจากไวรัส ผ่านแบบจำลองที่ถูกฝึกมาเรียบร้อยแล้ว (Pre-trained Model) แล้วสาม  
 แบบจำลอง ประกอบด้วย โมบายล์เน็ตวี2 (MobileNetV2) เรสเน็ต50 (Resnet50) และอินเซปชันวี3  
 (InceptionV3) ซึ่งได้ถูกเลือกมาใช้ในการทดสอบเพื่อสร้างแบบจำลองทั้งหมด 3 ตัว ประกอบด้วย ซี  
 เอ็นเอ็น+โมบายล์เน็ตวี2 ซีเอ็นเอ็น+เรสเน็ต50 และ ซีเอ็นเอ็น+อินเซปชันวี3 ซึ่งพบว่า สมรรถนะ  
 แบบจำลองซีเอ็นเอ็น+อินเซปชันวี3 ให้ผลลัพธ์ที่ดีที่สุด จึงถูกเลือกนำไปปรับรายละเอียด การ  
 ประเมินผลบนชุดข้อมูลทดสอบของแบบจำลองซีเอ็นเอ็น+อินเซปชันวี3 หลังจากทำการปรับ  
 รายละเอียด (Fine Tuning) ทั้งหมดด้วยกัน 8 ชั้น คือ ชั้นที่ 280, 250, 230, 200, 160, 150, 130 และ 120  
 ซึ่งแตกต่างจากบทความวิจัยส่วนใหญ่ที่ทำการละทิ้งการตรึงเพียงชั้นเดียว โดยเห็นได้ว่าการปรับ  
 รายละเอียดของแบบจำลองที่ทำการละทิ้งการตรึงตั้งแต่ชั้น 150 ให้ผลการทดสอบการจำแนกภาพถ่าย  
 รังสีทรวงอกของผู้ป่วยที่ติดเชื้อโควิด19 ได้ความแม่นยำที่ดีที่สุดที่ 95% ซึ่งเห็นได้ว่าแนวทางการ  
 จำแนกประเภทภาพที่นำเสนอมีความหวังสามารถนำไปพัฒนาต่อยอด เพื่อเป็นประโยชน์ต่อ  
 อุตสาหกรรมการแพทย์ได้

สาขาวิชา สถิติ  
 ปีการศึกษา 2564

ลายมือชื่อนิสิต .....  
 ลายมือชื่อ อ.ที่ปรึกษาหลัก .....

## 6380153326 : MAJOR STATISTICS

KEYWORD: transfer learning convolutional neural networks chest x-ray images

Thunchawin Photiwuttanut : IMAGE TRANSFER LEARNING FOR IMAGE CLASSIFICATION USING CONVOLUTIONAL NEURAL NETWORKS: A CASE STUDY OF COVID19-INFECTED CHEST X-RAY IMAGES.. Advisor: Asst. Prof. AKARIN PHAIBULPANICH, Ph.D.

Image processing techniques are widely used in a variety of industries today. In medicine, problems in image classification can be solved quickly and accurately through the application of convolutional neural networks (CNNs) with Transfer Learning. This research, therefore, presents a method for applying transfer learning techniques with an in-depth convolution network model to classify chest radiographs into 3 categories: 1) chest radiographs of Normal patients 2) chest radiographs of patients infected with COVID-19 3) chest radiographs of patients with viral pneumonia, using pre-trained models, and three models: Mobile Net V2, Resnet50, and InceptionV3 (InceptionV3). The following 3 models were selected for comparison: CNN+MobileNetV2, CNN + Resnet50, and CNN + Inception V3. The result shows that the CNN + Inception V3 model gives the best result. After adjustment by fine-tuning in 8 layers in the 280th, 250th, 230th, 200th, 160th, 150th, 130th, and 120th layer which is different from other research that normally chooses one layer to unfreeze, the accuracy of the model is improved. The model unfrozen from the 150th and above gives the best accuracy at 95%. The proposed image classification approach be further developed to benefit the medical field.

Field of Study: Statistics

Student's Signature .....

Academic Year: 2021

Advisor's Signature .....

## กิตติกรรมประกาศ

ด้วยความกรุณาอย่างยิ่งจาก รองศาสตราจารย์ ดร.อักรินทร์ ไพบูลย์พานิช อาจารย์ที่ปรึกษา วิทยานิพนธ์ผู้ที่ให้ความรู้และให้คำแนะนำทั้งคำปรึกษา คอยผลักดัน และ ตรวจแก้ไขข้อบกพร่องมาโดยตลอด วิทยานิพนธ์เล่มนี้จึงสำเร็จด้วยดี ผู้วิจัยขอกราบขอบพระคุณด้วยความเคารพอย่างสูงไว้ ณ ที่นี้

ขอขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร. ภูริพันธุ์ รุจิขจร ประธานกรรมการสอบวิทยานิพนธ์ และ ผู้ช่วยศาสตราจารย์ ดร. สุรณพีร์ ภูมิวุฒิสาร กรรมการสอบวิทยานิพนธ์ที่กรุณาเสียสละเวลาอันมีค่า ในการตรวจสอบและให้คำแนะนำที่เป็นประโยชน์ในการทำวิทยานิพนธ์ในครั้งนี้

ขอขอบพระคุณผู้ที่คอยให้กำลังใจตลอดระยะเวลาในการทำวิทยานิพนธ์เล่มนี้ ทั้งบิดามารดา วงศาคณาญาติทั้งหลาย และเพื่อนๆ พี่ๆ น้องๆ ขอบพระคุณที่คอยเป็นห่วงเป็นใย ถวายได้เสมอมาตั้งแต่เข้ารับการศึกษา จนกระทั่งทำวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วง รวมถึงขอบคุณตัวเองที่ไม่ยอมแพ้และท้อถอยและยืนหยัดและหวังว่าวิทยานิพนธ์ฉบับนี้จะเป็นประโยชน์ต่อผู้ที่สนใจต่อไปในภายภาคหน้า

ธัญญ์ชวิน โพธิ์วัฒนรัตน์

จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY

## สารบัญ

	หน้า
บทคัดย่อภาษาไทย .....	ก
บทคัดย่อภาษาอังกฤษ .....	ง
กิตติกรรมประกาศ.....	จ
สารบัญ.....	ฉ
สารบัญตาราง .....	ช
สารบัญภาพ .....	ฌ
บทที่ 1    บทนำ .....	1
1.1 ความเป็นมาและความสำคัญของปัญหา .....	1
1.2 วัตถุประสงค์การวิจัย .....	3
1.3 ขอบเขตของการวิจัย .....	3
1.4 วิธีการดำเนินการวิจัย.....	3
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	3
บทที่ 2    ทฤษฎีและงานวิจัยที่เกี่ยวข้อง .....	4
2.1 ทฤษฎีและกรอบแนวคิดที่เกี่ยวข้อง .....	4
2.2 เอกสารและงานวิจัยที่เกี่ยวข้อง .....	16
บทที่ 3    แนวคิดและวิธีการวิจัย.....	21
3.1 การเตรียมชุดข้อมูล.....	22
3.2 การสร้างแบบจำลองเพื่อเปรียบเทียบประสิทธิภาพแบบจำลองการเรียนรู้การถ่ายทอด .....	23
3.3 การวัดประสิทธิภาพของแบบจำลอง .....	23
3.4 การปรับละเอียดสมรรถนะของแบบจำลอง .....	23
3.5 สภาพแวดล้อมที่ใช้ในการทดสอบแบบจำลอง .....	23

3.6 การเตรียมชุดข้อมูลที่ใช้ในการทดสอบแบบจำลอง.....	24
3.7 การดำเนินการสร้างแบบจำลอง .....	27
3.7.1 แบบจำลองโครงข่ายประสาทคอนโวลูชันแบบอินเซปชันเวอร์ชัน 3 (InceptionV3) ..	27
3.7.2 แบบจำลองโครงข่ายประสาทเชิงลึกแบบ Residual Network (ResNet50) .....	29
3.7.3 แบบจำลองโครงข่ายประสาทคอนโวลูชันแบบ โมไบล์เน็ตเวอร์ชัน 2 (MobileNetV2)	30
บทที่ 4 ผลการวิจัย .....	31
4.1 การฝึกสอนแบบจำลอง.....	31
4.2 ผลการทดสอบแบบจำลอง.....	33
4.3 การปรับละเอียดสมรรถนะแบบจำลอง.....	37
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ .....	41
5.1 สรุปผลการวิจัย .....	41
5.2 ข้อเสนอแนะ .....	42
บรรณานุกรม .....	44
ประวัติผู้เขียน .....	47



## สารบัญตาราง

หน้า

ตารางที่ 1 ตารางเปรียบเทียบสมรรถนะของ MobileNetV2, ResNet-50 และInceptionV3 .....	12
ตารางที่ 2 ตารางสรุปจำนวนข้อมูลสอน ข้อมูลตรวจสอบ ข้อมูลทดสอบของภาพถ่ายรังสีทรวงอก ทั้งสามประเภท .....	22
ตารางที่ 3 ประสิทธิภาพในการฝึกสอนแบบจำลองการเรียนรู้การถ่ายทอดแบบ InceptionV3, MobileNetV2 และResNet50.....	33
ตารางที่ 4 ประสิทธิภาพของการทดสอบแบบจำลองการเรียนรู้การถ่ายทอดแบบ InceptionV3, MobileNetV2 และResnet50.....	35
ตารางที่ 5 ประสิทธิภาพของการทดสอบแบบจำลองการเรียนรู้การถ่ายทอดแบบ InceptionV3 ใน การละทิ้งการตรึงในชั้นที่กำหนด.....	39

## สารบัญภาพ

หน้า

รูปที่ 1 เพอร์เซปตรอน ( Perceptron).....	4
รูปที่ 2 โครงข่ายประสาทเทียม (Neural networks).....	6
รูปที่ 3 การเรียนรู้แบบคอนโวลูชัน (Convolutional Neural Networks).....	8
รูปที่ 4 การทำงานของตัวกรอง(kernel) .....	8
รูปที่ 5 การทำงานของ max pooling และ average pooling.....	9
รูปที่ 6 Residual Block.....	10
รูปที่ 7 โครงสร้างของสถาปัตยกรรมโครงข่ายประสาทแบบ Inception V3 .....	11
รูปที่ 8 โครงสร้างของสถาปัตยกรรมโครงข่ายประสาทแบบ MobileNetV2.....	12
รูปที่ 9 ตัวอย่างภาพถ่ายรังสีทรวงอกที่ผ่านการทำ Data Augmentation .....	13
รูปที่ 10 ตัวอย่างชุดข้อมูลภาพถ่ายรังสีทรวงอก ในการจำแนกผู้ป่วยที่ติดเชื้อ โควิด19 ผู้ป่วยที่เป็น ปอดอักเสบจากไวรัส และผู้ป่วยปกติ .....	18
รูปที่ 11 ตัวอย่างชุดข้อมูลภาพถ่ายรังสีทรวงอก เพื่อคัดแยกผู้ป่วยที่เป็นมะเร็งปอด และ ผู้ป่วยปกติ .....	19
รูปที่ 12 ตัวอย่างชุดข้อมูลภาพถ่ายรังสีทรวงอก ผู้ป่วยที่เป็นโรคเกี่ยวกับปอด 5 โรค และผู้ป่วยปกติ .....	20
รูปที่ 13 ตัวอย่างชุดข้อมูลภาพถ่ายรังสีทรวงอก 3 ประเภท ที่ใช้ในการฝึกฝนแบบจำลอง .....	22
รูปที่ 14 โครงสร้างชุดข้อมูลในการจำแนกภาพถ่ายรังสีทรวงอก .....	24
รูปที่ 27 ฟังก์ชันและกราฟแสดงประสิทธิภาพของการฝึกสอน โมเดลใน 30 epochs .....	32
รูปที่ 28 ฟังก์ชันในการวัดสมรรถนะของการฝึกสอนแบบจำลอง.....	32
รูปที่ 29 ตัวอย่างการใช้ฟังก์ชัน softmax เพื่อดึงค่าคำตอบในการทำนาย.....	34
รูปที่ 30 ตัวอย่างผลลัพธ์ของ confusion matrix และแสดงผลการทำนายของข้อมูลแต่ละคลาสด้วย classification report .....	34



จุฬาลงกรณ์มหาวิทยาลัย  
**CHULALONGKORN UNIVERSITY**

# บทที่ 1

## บทนำ

### 1.1 ความเป็นมาและความสำคัญของปัญหา

โรคโควิด 19 เป็นการติดเชื้อไวรัสทางระบบทางเดินหายใจ และสามารถติดต่อกันได้ง่าย ซึ่งมีที่มาจากการติดเชื้อไวรัส SARS-CoV-2 ที่มีต้นกำเนิดมาจากเมือง อู่ฮั่น ประเทศจีน ในปี 2019 [1] โดยเชื้อไวรัสโควิด 19 ได้แพร่กระจายไปทั่วภูมิภาค ทำให้มีผู้ป่วยที่ติดเชื้อ และเสียชีวิตจากไวรัสโควิด 19 จำนวนมาก โดยอาการของผู้ป่วยที่ติดเชื้อโควิด 19 มีตั้งแต่อาการปวดหัว มีไข้ ไอ จาม ท้องเสีย หรือ มีการหายใจติดขัด ซึ่งส่งผลกระทบต่อการทำงานของชีวิตประจำวันและในผู้ป่วยบางรายถึงแม้จะหายจากโรคโควิด 19 แล้ว แต่ยังคงต้องเผชิญผลข้างเคียงในระยะยาวเป็นหลายสัปดาห์ [2] ในขณะเดียวกันผู้ป่วยที่เป็นโรคปอดอักเสบจากการติดเชื้อไวรัสก็มีลักษณะของอาการที่คล้ายเคียงกันกับผู้ป่วยที่ติดเชื้อโควิด 19 ในขณะที่ภาพถ่ายรังสีทรวงอกมีความแตกต่างกันไม่มาก ซึ่งการจำแนกผู้ป่วยที่ติดเชื้อไวรัสโควิด 19 และผู้ป่วยที่ติดเชื้อปอดบวมจากไวรัสที่แม่นยำจะส่งผลต่อการรักษา เนื่องจากวิธีการรักษานั้นมีความแตกต่างกันอย่างสิ้นเชิง [3]

สำหรับการจำแนกโรคของผู้ป่วยที่ติดเชื้อระบบทางเดินหายใจนั้นมีการนำเทคนิคการประมวลผลด้วยภาพที่หลากหลายมาใช้ในการจำแนกโรคของผู้ป่วย หนึ่งในเทคนิคที่นิยมใช้กันคือการใช้โครงข่ายแบบคอนโวลูชัน (Convolutional Neural Networks) ซึ่งเป็นหนึ่งในตัวแบบการเรียนรู้ที่เป็นที่นิยมอย่างมากในการทำงานกับข้อมูลที่เป็นรูปภาพ และเสียง ซึ่งได้รับการคิดค้นครั้งแรกในปี 1980 โดย Yann LeCun และได้ถูกนำไปใช้ในหลายอุตสาหกรรม [4] การทำงานของโครงข่ายแบบคอนโวลูชัน จะแบ่งออกเป็น 4 ส่วนหลัก คือการทำคอนโวลูชัน การจัดการความเป็นเชิงเส้น การทำพูลลิ่ง และการเชื่อมต่อกันของแต่ละเลเยอร์อย่างสมบูรณ์ ซึ่งผลลัพธ์ที่ได้สามารถที่จะเป็นได้ทั้งตัวแปรตามที่เป็น 2 กลุ่ม หรือมากกว่า โดยหนึ่งในเทคนิคที่นิยมใช้ควบคู่ไปกับการเรียนรู้แบบคอนโวลูชัน คือการเรียนรู้แบบถ่ายทอด (Transfer learning) ซึ่งเป็น การนำสถาปัตยกรรมการเรียนรู้แบบคอนโวลูชันที่ถูกฝึกมาแล้วในส่วนของการทำงานคอนโวลูชัน การจัดการความเป็นเชิงเส้น และการทำพูลลิ่ง มาเชื่อมต่อกับชั้นการเชื่อมต่อกันของแต่ละเลเยอร์อย่างสมบูรณ์ที่ถูกสร้างขึ้นเองภายหลัง และใช้ข้อมูลใหม่ของผู้สร้างแบบจำลองต้องการทำนาย

ในทางการแพทย์ มีนักวิจัยหลายท่านนำเทคนิคการประมวลผลด้วยภาพที่หลากหลายมาใช้เพื่อการจำแนกโรคของผู้ป่วยที่ติดเชื้อในระบบทางเดินหายใจผ่านภาพถ่ายรังสีทรวงอก เช่น Jawad และคณะ [5] ที่ใช้ Logistic regression และ CNN มาจำแนกผู้ป่วยที่ติดเชื้อ โควิด 19 และผู้ป่วยปอดปกติ หรือ Lakhani [6] ที่ใช้ CNN 2 โมเดล คือ AlexNet และ GoogleNet เพื่อการคัดแยกระหว่างปอดของผู้ป่วยติดเชื้อ วัณโรคทางปอด (Pulmonary Tuberculosis) และผู้ป่วยปกติ ซึ่งได้ประสิทธิภาพของแบบจำลองที่สูง แต่อย่างไรก็ตามการนำเทคนิคโครงข่ายแบบคอนโวลูชัน ควบคู่ไปกับการเรียนรู้แบบถ่ายทอดที่เปรียบเทียบกับโมเดลที่ถูกฝึกมาแล้ว (Pre-trained model) หลายรูปแบบนั้นยังมีจำนวนน้อย ยิ่งไปกว่านั้นจำนวนของข้อมูลที่นำมาฝึกฝนแบบจำลองก็มีจำนวนน้อย และประสบกับปัญหาการไม่เท่ากันของจำนวนข้อมูลในแต่ละคลาส (Imbalanced data)

ในส่วนของที่มาของข้อมูล ผู้วิจัยได้ทำการเก็บรวบรวมมาจากแหล่งข้อมูลสาธารณะที่นำเชื่อถือคือ <https://www.kaggle.com/tawsifurrahman/covid19-radiography-database> เพื่อนำมาฝึกฝนแบบจำลอง โดยชุดของข้อมูลเป็นที่รู้จักกันอย่างแพร่หลายผ่านชื่อ COVID-19 radiography database ซึ่งเป็นชุดข้อมูลที่ชนะเลิศในการแข่งขัน Covid-19 dataset และได้รับการยกย่องโดย Kaggle community โดยจำนวนของภาพถ่ายรังสีทรวงอกที่นำมาใช้มีทั้งหมด 3480 ภาพ โดยแบ่งออกเป็น 1) ภาพถ่ายรังสีทรวงอกของผู้ป่วยปกติ 1160 ภาพ 2) ภาพถ่ายรังสีทรวงอกของผู้ป่วยที่ติดเชื้อ โควิด 19 ที่ 1160 ภาพ และ 3) ภาพถ่ายรังสีทรวงอกของผู้ติดเชื้อปอดอักเสบจากไวรัส ที่ 1160 ภาพ โดยแบ่งเป็นชุดข้อมูลสอน 2100 ภาพ ชุดข้อมูลทดสอบ 690 ภาพ และชุดข้อมูลตรวจสอบ 690 ภาพ

ผู้วิจัยสนใจที่จะทำการศึกษาการจำแนกประเภทภาพถ่ายรังสีทรวงอกออกเป็น 3 ประเภทคือ 1) ภาพถ่ายรังสีทรวงอกของผู้ป่วยปกติ 2) ภาพถ่ายรังสีทรวงอกของผู้ป่วยที่ติดเชื้อ โควิด 19 3) ภาพถ่ายรังสีทรวงอกของผู้ติดเชื้อปอดอักเสบจากไวรัส จากชุดข้อมูลจากแหล่งนำเชื่อถือบนอินเทอร์เน็ต ด้วยเทคนิคการเรียนรู้เชิงลึกแบบคอนโวลูชันร่วมกับเทคนิคการเรียนรู้แบบถ่ายทอด จาก Pre-trained models ทั้งหมด 3 แบบ คือ MobileNetV2 ResNet50 และ InceptionV3 และประเมินสมรรถนะของแบบจำลองที่ได้ด้วย accuracy, precision, recall และ F1 และนำโมเดลที่มีประสิทธิภาพดีที่สุดนำมาปรับพารามิเตอร์ของแบบจำลองเพื่อพัฒนาประสิทธิภาพให้ดียิ่งขึ้น

โดยทำการละทิ้งการตรึงในแต่ละชั้นของโมเดล และหาชั้นเลเยอร์ที่เหมาะสมที่สุดที่ทำให้โมเดลมีประสิทธิภาพสูงสุด

## 1.2 วัตถุประสงค์การวิจัย

1. เปรียบเทียบประสิทธิภาพในการทำนายของแบบจำลองการเรียนรู้แบบคอนโวลูชัน ทั้ง 3 แบบ คือ MobileNetV2 ResNet50 และ InceptionV3 ที่ใช้ควบคู่กับการเรียนรู้การถ่ายทอดในการทำนายข้อมูลภาพถ่ายรังสีทรวงอก
2. หาชั้นเลเยอร์จากการ fine-tuning ของ pre-trained model ที่เหมาะสมที่สุดที่ทำให้โมเดลมีประสิทธิภาพมากที่สุด

## 1.3 ขอบเขตของการวิจัย

1. แหล่งที่มาของข้อมูลภาพถ่ายรังสีทรวงอกของผู้ป่วยที่ติดเชื้อโควิด 19 ผู้ป่วยปกติ และผู้ป่วยติดเชื้อปอดอักเสบจากไวรัส จากแหล่งข้อมูลบนอินเทอร์เน็ตที่เชื่อถือได้
2. Pre-trained models ประกอบด้วย MobileNetV2 Resnet50 และ InceptionV3
3. ประเมินประสิทธิภาพแบบจำลองด้วยค่า accuracy, confusion matrix และ F1-score

## 1.4 วิธีการดำเนินการวิจัย

1. ค้นคว้าเอกสาร ทฤษฎี และกรอบแนวคิดที่เกี่ยวข้อง
2. ทำการรวบรวมข้อมูลภาพถ่ายรังสีทรวงอกของผู้ป่วยที่ติดเชื้อโควิด 19 ผู้ป่วยปกติ และผู้ป่วยปอดอักเสบจากไวรัส
3. ทำการเตรียมข้อมูลสำหรับการสร้างแบบจำลอง
4. สร้างแบบจำลองเพื่อจำแนกผู้ป่วยที่ติดเชื้อโควิด 19 ผู้ป่วยปกติ และผู้ป่วยปอดอักเสบจากไวรัส
5. พัฒนาประสิทธิภาพของแบบจำลอง
6. สรุปการดำเนินงาน

## 1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. ได้แบบจำลองที่สามารถใช้จำแนกภาพถ่ายรังสีทรวงอกของผู้ป่วยที่ติดเชื้อโควิด 19 ผู้ป่วยปกติ และผู้ป่วยปอดอักเสบจากไวรัส
2. สามารถนำไปใช้การตรวจจับการติดเชื้อทางปอดได้อย่างอัตโนมัติ

## บทที่ 2

### ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

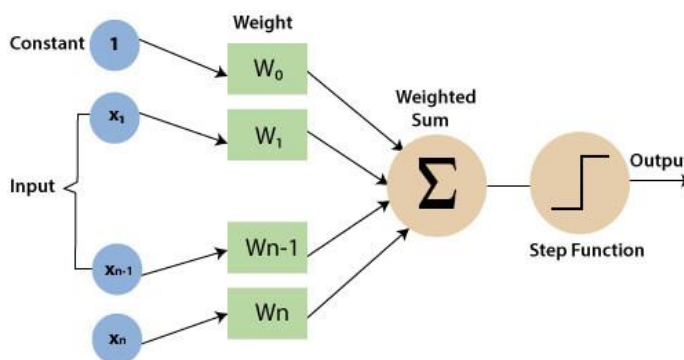
#### 2.1 ทฤษฎีและกรอบแนวคิดที่เกี่ยวข้อง

##### 2.1.1 การเรียนรู้เชิงลึก (Deep Learning) [7]

การเรียนรู้เชิงลึกสร้างขึ้นจากแนวคิดการทำงานของเซลล์ประสาทของสมองมนุษย์ที่มีการเชื่อมต่อกันจำนวนมาก ซึ่งมีมากกว่า 1,000 ล้าน เซลล์ ซึ่งทำให้มนุษย์สามารถจัดลำดับความคิด รับรู้ความรู้สึก และวิเคราะห์ปัญหาที่มีความซับซ้อนออกได้ โดยการประมวลผลข้อมูลที่เป็นรูปภาพ วิดีโอ หรือ เสียง จะถูกนำมาวิเคราะห์ผ่านการเรียนรู้เชิงลึกเป็นส่วนมาก ในการทำงานของ การเรียนรู้เชิงลึกไม่จำเป็นต้องถูกควบคุมโดยผู้ใช้งาน หลังจากทำการประมวลผลจะนำไปสู่การได้ผลลัพธ์จากการทำนายในขั้นตอนสุดท้าย

##### 2.1.2 เพอร์เซปตรอน (Perceptron) [8]

การทำงานของเพอร์เซปตรอน ดังรูปที่ 1 คล้ายคลึงกับการทำงานของเซลล์ประสาทสมองมนุษย์ ประกอบด้วย 5 ส่วนหลัก คือส่วนรับข้อมูลเข้า (input) ค่าถ่วงน้ำหนัก (weight) ส่วนรวมผลคูณของค่าข้อมูลนำเข้าและค่าถ่วงน้ำหนัก (weighted sum) ค่าเบี่ยงเบน (bias) และส่วนขจัดความเป็นเชิงเส้น (activation function)



รูปที่ 1 เพอร์เซปตรอน (Perceptron)

ที่มา: <https://www.javatpoint.com/pytorch-perceptron>

### 1. ค่าถ่วงน้ำหนัก (Weight)

ค่าถ่วงน้ำหนักเป็นค่าพารามิเตอร์ที่เป็นส่วนประกอบของเพอร์เซปตรอน ซึ่งจะมีการปรับเปลี่ยนไปตลอดจนกว่าจะได้ค่าที่ดีที่สุด โดยค่าถ่วงน้ำหนักจะถูกคูณเข้ากับค่าข้อมูลสำหรับแต่ละหน่วยของข้อมูลนำเข้า ก่อนที่จะถูกรวมและส่งผ่านส่วนจัดเชิงเส้น

### 2. ค่าเบี่ยงเบน (Bias)

ค่าเบี่ยงเบนเป็นค่าพารามิเตอร์ที่มี 1 ตัวต่อ 1 เพอร์เซปตรอน ซึ่งจะมีการปรับเปลี่ยนไปจนกว่าจะได้ค่าที่ดีที่สุดเหมือนกับค่าถ่วงน้ำหนัก โดยค่าเบี่ยงเบนจะถูกคูณเข้ากับผลรวมของผลคูณระหว่างค่าถ่วงน้ำหนักกับค่าข้อมูลนำเข้าทั้งหมดที่เข้าสู่เพอร์เซปตรอน

### 3. ส่วนจัดความเป็นเชิงเส้น (Activation Function) [7] [9]

ส่วนจัดความเป็นเชิงเส้นเป็นฟังก์ชันที่รับผลรวมการประมวลผลทั้งหมดแล้วนำมาคำนวณในสมการที่ไม่เป็นเชิงเส้น หลังจากผ่านส่วนรวมผลคูณของค่าข้อมูลนำเข้าและค่าถ่วงน้ำหนัก

#### 3.1 ฟังก์ชัน Rectified linear unit (ReLU)

ผลลัพธ์จากการดำเนินการของฟังก์ชันมีค่าเท่ากับ  $[0, \infty)$  คือเมื่อข้อมูลที่รับเข้ามีค่าตั้งแต่ 0 ขึ้นไป ผลลัพธ์ก็จะมีค่าเท่ากับ ส่วนข้อมูลที่รับเข้าที่น้อยกว่า 0 จะกำหนดให้ผลลัพธ์เป็น 0 ตลอดการใช้ ReLU ใน Hidden layer นอกจากจะทำให้การเทรนทำได้รวดเร็วกว่า Sigmoid หลายเท่าแล้ว ยังช่วยแก้ปัญหา Vanishing gradients ได้เป็นอย่างมาก เพราะความชันของฟังก์ชันจะเป็นค่าคงตัวคือ 1 เสมอ

$$relu(z) = \max(0, x)$$



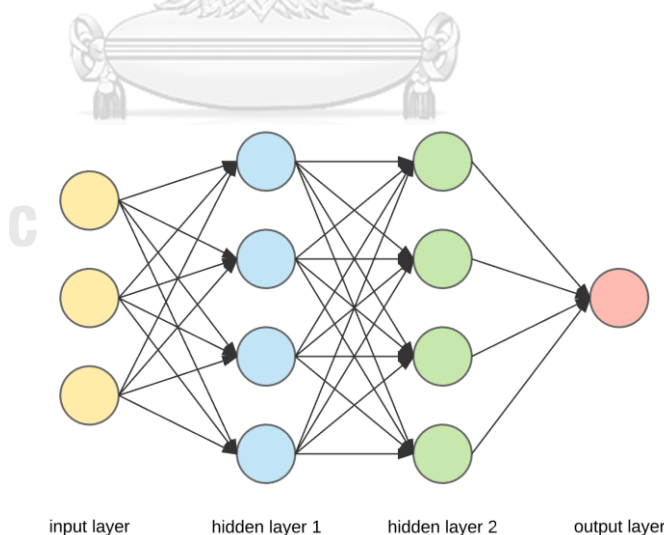
### 3.2 ฟังก์ชัน Softmax

เป็นฟังก์ชันที่มีการดำเนินการทางคณิตศาสตร์ โดยผลลัพธ์อยู่ในรูปแบบความน่าจะเป็น ฟังก์ชัน softmax จะถูกใช้ในชั้นสุดท้ายเมื่อใช้ฟังก์ชัน ReLU สำหรับการแก้ปัญหาค่าจำแนกประเภท

$$\sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^k e^{z_j}}$$

#### 2.1.3 โครงข่ายประสาทเทียม (Neural networks) [7] [9]

โครงข่ายประสาทเทียม ดังรูปที่ 2 ประกอบด้วยชั้นการดำเนินการหลัก 3 ชั้น ได้แก่ชั้นข้อมูลนำเข้า (Input Layer) ชั้นการประมวลผลซ่อน (Hidden Layer) และชั้นผลลัพธ์ (Output Layer) ในชั้นนำเข้าจะมี โหนด (Node) ที่เป็นข้อมูลที่เรานำเข้ามา หลังจากนั้นจะไหลผ่านไปในชั้นประมวลผลซ่อนซึ่งจะเป็นชั้นที่มีการคำนวณส่วนรวมผลคูณของค่าข้อมูลนำเข้าและค่าถ่วงน้ำหนัก และส่วนจัดความเป็นเชิงเส้น เป็นชั้นที่มีการเรียงตัวของเพอร์เซปตรอนหลายชั้นต่อกัน ซึ่งจะมีที่ขึ้นขึ้นอยู่กับการพิจารณาของผู้สร้างโมเดล ชั้นผลลัพธ์จะเป็นชั้นสุดท้ายของโครงข่ายประสาทเทียม ซึ่งจำนวนโหนดจะขึ้นอยู่กับจำนวนผลลัพธ์ของปัญหา โดยแต่ละชั้นจะมีการเชื่อมต่อกันอย่างสมบูรณ์



รูปที่ 2 โครงข่ายประสาทเทียม (Neural networks)

ที่มา: <https://towardsdatascience.com/applied-deep-learning-part-1-artificial-neural-networks-d7834f67a4f6>

### 2.1.4 การเรียนรู้ของโครงข่ายประสาทเทียม (Network Learning) [7] [9]

การเรียนรู้ของโครงข่ายประสาทเทียมประกอบด้วย 2 ส่วนหลัก คือการดำเนินงานไปข้างหน้า (Forward Propagation) และ การดำเนินงานย้อนกลับ (Backward Propagation) การดำเนินการไปข้างหน้าจะมีการคำนวณค่าถ่วงน้ำหนัก และค่าเบี่ยงเบน อย่างสุ่ม หลังจากนั้นจะทำการดำเนินงานย้อนกลับ โดยเรียนรู้จากการคลาดเคลื่อนหลังจากนั้นจะทำการปรับปรุงค่าถ่วงน้ำหนักและค่าเบี่ยงเบนเพื่อที่จะสะท้อน ผลความต่างระหว่างค่าพยากรณ์กับค่าจริง

### 2.1.5 อัลกอริทึมปรับปรุงประสิทธิภาพ (Optimization Algorithms) [7] [9]

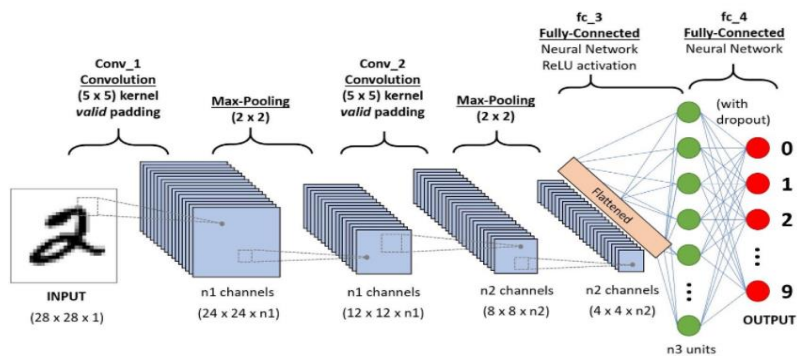
อัลกอริทึมปรับปรุงประสิทธิภาพทำหน้าที่สำหรับการปรับปรุงค่าให้โมเดลมีการเรียนรู้ที่เหมาะสม เพื่อการหาตำแหน่งที่ทำให้มีข้อผิดพลาดน้อยที่สุด รวมถึงขั้นตอนการปรับค่าพารามิเตอร์ของโครงข่าย ตัวอย่างของอัลกอริทึมปรับปรุงประสิทธิภาพพื้นฐาน คือ Gradient descents , Adagrad optimizer, Momentum optimizer, Adam optimizer, Ftrl optimizer, RMSProp optimizer เป็นต้น

### 2.1.6 ทฤษฎีการเรียนรู้แบบคอนโวลูชัน (Convolutional Neural Networks) [1,2,6]

สถาปัตยกรรมของการเรียนรู้เชิงลึกมีเป้าหมายที่จะทำให้คอมพิวเตอร์สามารถที่จะเข้าใจข้อมูลที่ได้รับ โดยการเรียนรู้แบบคอนโวลูชัน จะแบ่งการทำงานเป็น 4 ส่วน ดังรูปที่ 3

1. การทำคอนโวลูชัน
2. การขจัดความเป็นเชิงเส้น
3. การทำพูลลิ่ง
4. การเชื่อมต่อกันของแต่ละเลเยอร์อย่างสมบูรณ์

ข้อ 1 ถึง 3 เป็นส่วนที่เรียกว่า body หรือ convolutional layers โดยมีตัวกรองที่มีหน้าที่แยกลักษณะของพื้นที่ย่อย เช่น ขอบ ลายเส้น การตัดกันของสี และ รูปทรง ของรูปภาพ หลักจากนั้นจะนำไปคำนวณผ่านสมการการขจัดความเป็นเชิงเส้นเพื่อทำให้ข้อมูลง่ายต่อการคำนวณ และ ส่วนของ pooling ที่จะทำให้ข้อมูลมีขนาดเล็กลง แต่รายละเอียดของข้อมูลยังคงเดิม ในส่วนของ head คือ fully connected layer ที่จะเชื่อมข้อมูลจากส่วน body แล้วนำมาคำนวณหาคำตอบ



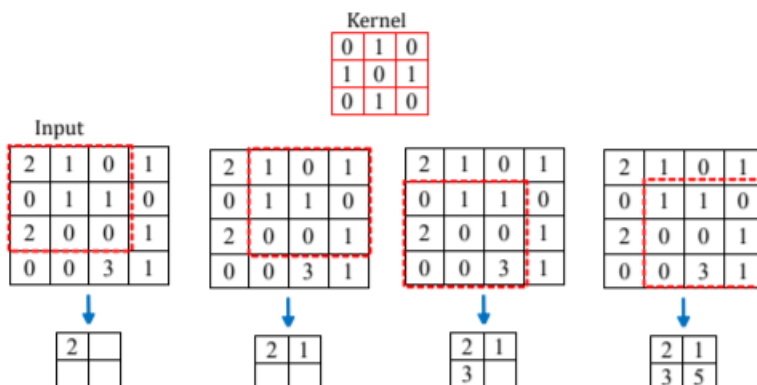
A CNN sequence to classify handwritten digits

### รูปที่ 3 การเรียนรู้แบบคอนโวลูชัน (Convolutional Neural Networks)

ที่มา: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>

#### 1. การทำคอนโวลูชัน (Convolution)

การทำคอนโวลูชันเป็นส่วนของการทำงานที่ตัวกรอง (kernel) จะทำการกรองพื้นที่บนเมทริกภาพนำเข้าตามการเลื่อนของพิกเซลที่กำหนด โดยการเลื่อนของพิกเซลจะเรียกว่า stride ถ้าเลื่อน 2 พิกเซล คือ 2 stride ค่าข้อมูลในช่องเมทริกที่ซ้อนกันจะถูกดำเนินการคูณ จากนั้นนำผลคูณแต่ละช่องรวมเข้าด้วยกัน ค่าผลรวมที่ได้จะถูกวางลงในเมทริกใหม่และสิ่งที่ได้จะเรียกว่า feature map



รูปที่ 4 การทำงานของตัวกรอง(kernel)

ที่มา: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>

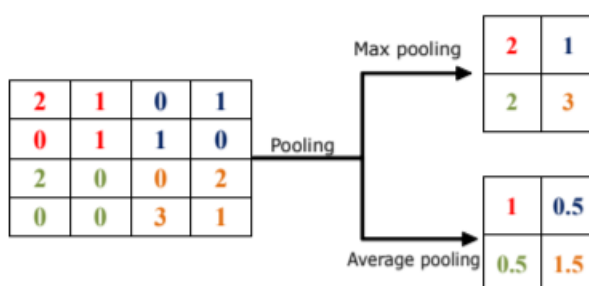
feature map ที่เป็นผลลัพธ์จะมีขนาดที่เล็กกว่าภาพที่นำเข้ามา เมื่อมีการทำงานของ convolution ไปเรื่อยต่อเนื่องภาพข้อมูลเข้าก็จะมีความถี่ลดลงไปต่อเนื่อง ซึ่งทำให้การทำงานของ convolution นั้นจำกัด แต่เมื่อมีการพิจารณาความสำคัญของข้อมูลที่ส่วนขอบของภาพนั้นมีความสำคัญน้อยกว่า ส่วนกลางของภาพ ซึ่งสามารถแก้ไขปัญหานี้ได้โดยการเพิ่มขอบของภาพ (padding) ยกตัวอย่างเช่น ภาพข้อมูลเข้ามีขนาด  $6 \times 6$  เมื่อเราเพิ่มขอบของภาพไป 1px padding ภาพของเราจะมีขนาด  $8 \times 8$  เมื่อใช้ตัวกรองที่มีขนาด  $3 \times 3$  ผลลัพธ์ของ feature map จะมีขนาด  $6 \times 6$  โดยปกติแล้วค่าในส่วนขอบที่เพิ่มเติมจะเป็น 0 โดยขนาดของ padding จะขึ้นอยู่กับขนาดของตัวกรอง

## 2. การขจัดความเป็นเชิงเส้น (activation function rectified linear)

หลังจากที่ได้ feature map ก็จะนำมาคำนวณเพื่อทำให้ข้อมูลไม่มีลักษณะเชิงเส้น โดยการใช้ฟังก์ชันคณิตศาสตร์ที่จะใช้จะขึ้นอยู่กับผู้ควบคุมแบบจำลอง เช่น ฟังก์ชัน Relu หรือ Sigmoid

## 3. การทำพูลลิ่ง (Pooling)

การพูลลิ่งเป็นการลดขนาดเชิงพื้นที่ด้วยวิธีการสุ่มตัวอย่างแต่ยังคงรักษาข้อมูลที่สำคัญไว้ และสามารถลดจำนวนพีเจอร์ที่จะนำไปคำนวณต่อซึ่งจะทำให้กระบวนการทำงานไวขึ้น การพูลลิ่งจะแบ่งเป็น 2 ประเภท คือ max pooling และ average pooling ซึ่งการทำงานของทั้ง 2 ประเภทจะคล้ายตัวกรองที่จะเลื่อนไปแต่ละพื้นที่ของข้อมูลแต่จะแตกต่างกันที่ max pooling จะสุ่มตัวอย่างที่มีค่ามากที่สุด ในขณะที่ average pooling จะทำการสุ่มตัวอย่างค่าเฉลี่ย ดังรูปที่ 5



รูปที่ 5 การทำงานของ max pooling และ average pooling

ที่มา: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks->

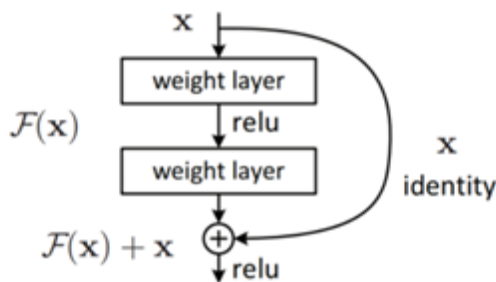
the-eli5-way-3bd2b1164a53

#### 4. การเชื่อมต่อกันของแต่ละเลเยอร์อย่างสมบูรณ์ (Fully Connected Layer)

การเชื่อมต่อกันของแต่ละเลเยอร์อย่างสมบูรณ์จะเป็นขั้นสุดท้ายของโครงข่ายคอนโวลูชัน โดยข้อมูลหลังจากผ่านการ pooling จะถูกนำมาเชื่อมต่อกับชั้นการเชื่อมต่อกันของแต่ละเลเยอร์อย่างสมบูรณ์ โดยการทำงานจะนำผลลัพธ์มาคัดกรองข้อมูลที่รับเข้ามาทำนายว่าอยู่ในประเภทไหน ซึ่งทำงานเหมือนโครงข่ายประสาททั่วไป ดังนั้นในขั้นสุดท้ายของชั้นจะทำการหาความน่าจะเป็นของข้อมูลแต่ละประเภทผ่านฟังก์ชัน Softmax

#### 2.1.7 โครงข่ายประสาทเชิงลึกแบบ Residual Network (ResNet) [10]

ชื่อเต็มของ ResNet คือ Deep Residual Network ได้รับการนำเสนอครั้งแรกในงานวิจัยชื่อ Deep Residual Learning for Image Recognition ซึ่งนำเสนอวิธีการแก้ปัญหาเรื่อง Vanishing gradient ซึ่งเกิดขึ้นกับโครงข่ายที่มีความลึกค่อนข้างมาก ด้วยการใส่ทางลัด (shortcut) ลงในโครงข่าย หรือที่เรียกว่าเทคนิคข้ามการเชื่อมต่อดังรูปที่ 6



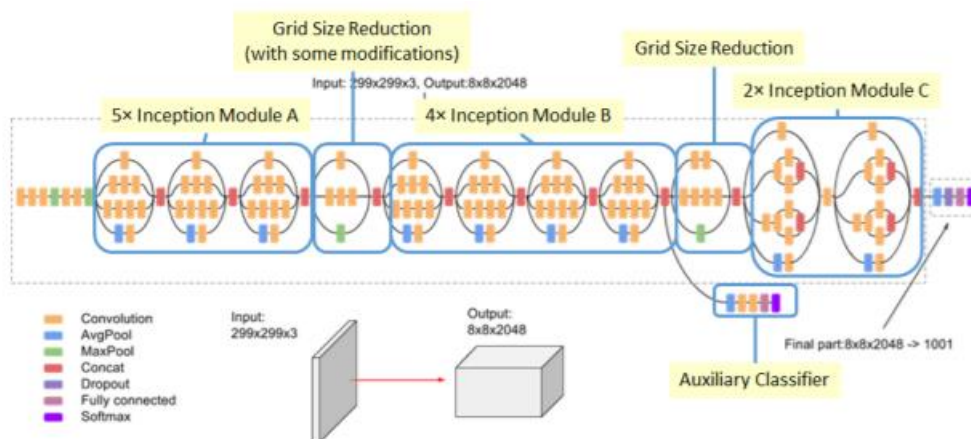
รูปที่ 6 Residual Block

ที่มา: <https://towardsdatascience.com/residual-blocks-building-blocks-of-resnet-fd90ca15d6ec>

โดยจำนวนชั้นของโครงข่ายจะถูกนำมาต่อท้ายชื่อ Resnet นั้นหมายความว่า ResNet-50 มีจำนวนโครงสร้างสร้างทั้งหมด 50 ชั้น ซึ่งการแก้ปัญหา vanishing gradient เป็นสิ่งที่สำคัญอย่างมาก เนื่องจากเมื่อเราฝึกสอนแบบ โดยการปรับพารามิเตอร์เดิมติดต่อกันในกระบวนการ gradient descent การปรับพารามิเตอร์ในครั้งถัด ๆ ไป ค่าสูญเสียจะเปลี่ยนแปลงน้อยลงมากจนไม่เกิดผลในการ Optimize หรืออาจจะต้องใช้จำนวนรอบในการฝึกสอนจำนวนหลายรอบ

### 2.1.8 โครงข่ายประสาทแบบ InceptionV3 [11]

สถาปัตยกรรม InceptionV3 ดังรูปที่ 7 เป็นรุ่นที่ 3 ในสถาปัตยกรรม Inception ซึ่งถูกคิดค้นและพัฒนาโดยทีมพัฒนาจาก google ซึ่งถูกต่อยอดจาก Inception1 และ 2 โดยการลดโครงสร้างภายในออกเป็น 5 ขั้นตอน ดังรูปที่ 7



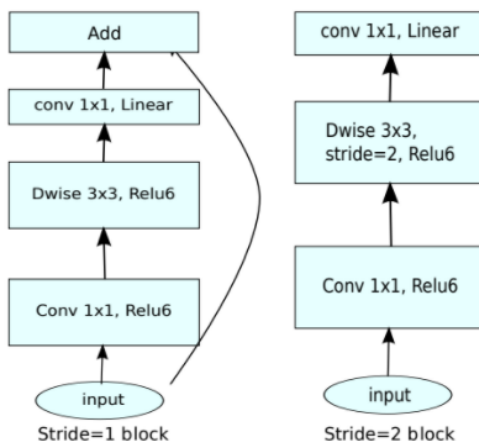
รูปที่ 7 โครงสร้างของสถาปัตยกรรมโครงข่ายประสาทแบบ Inception V3

ที่มา: <https://sh-tsang.medium.com/review-inception-v3-1st-runner-up-image-classification-in-ilsvrc-2015-17915421f77c>

โดย InceptionV3 มีจุดมุ่งหมายที่จะลดความลึกของชั้นโครงข่ายลง ในขณะที่จำนวนพารามิเตอร์ไม่มีขนาดที่ใหญ่เกินไป เพื่อที่จะคงประสิทธิภาพของแบบจำลองและสามารถดำเนินงานได้เร็วขึ้น

### 2.1.9 โครงข่ายประสาทแบบ MobileNetV2 [12]

สถาปัตยกรรม MobileNetV2 ดังรูปที่ 8 พัฒนาต่อมาจาก MobileNet โดยมีจุดประสงค์ที่จะลดขนาดของแบบจำลองการเรียนรู้เชิงลึกที่มีขนาดใหญ่ให้มีขนาดเล็กลงเพื่อที่จะสามารถนำไปประยุกต์ใช้กับอุปกรณ์ขนาดเล็กที่สามารถพกพาได้ เช่น โทรศัพท์มือถือ โดยโครงข่ายของสถาปัตยกรรม MobileNetV2 มีความคล้ายคลึงกับ MobileNet มาก แต่แตกต่างกันตรงที่ Inverted residual block กับ bottlenecking features ซึ่งมีโครงสร้างของจำนวน channels ในลักษณะ narrow-wide-narrow



รูปที่ 8 โครงสร้างของสถาปัตยกรรมโครงข่ายประสาทแบบ MobileNetV2

ที่มา: <https://paperswithcode.com/method/mobilenetv2>

การเปรียบเทียบสมรรถนะของ pre-trained model ทั้ง MobileNetV2, ResNet-50 และ InceptionV3 ที่ใช้กับงานจำแนกประเภทของรูปภาพจากชุดข้อมูล ImageNet ได้ผลดังแสดงใน

ตารางที่ 1 ตารางเปรียบเทียบสมรรถนะของ MobileNetV2, ResNet-50 และ InceptionV3

Model	Size (MB)	Top-5 Accuracy	Parameters
MobileNetV2	14	0.901	3,538,984
ResNet50	98	0.921	25,636,712
InceptionV3	92	0.937	23,851,784

### 2.1.10 ทฤษฎีการเรียนรู้แบบถ่ายทอด (Transfer learning) [13]

การฝึกโมเดลตั้งแต่ต้นจนจบของการเรียนรู้เชิงลึกต้องใช้เวลาการทำงานที่ค่อนข้างนานเนื่องจากมีตัวแปรค่าถ่วงน้ำหนักที่มีจำนวนมาก การเรียนรู้แบบถ่ายทอดสามารถเข้ามาช่วยลดระยะเวลาในการฝึกโมเดล และสามารถเรียนรู้จากโมเดลที่ไม่ปริมาณไม่มากได้ โดยการทำงานจะใช้ค่าถ่วงน้ำหนัก และ ค่าเบี่ยงเบนจากโมเดลที่มีฐานข้อมูลที่ยังขนาดใหญ่ซึ่งถูกฝึกมาแล้วในส่วน of body ของ CNN ซึ่งทำให้ประหยัดเวลาในการฝึกโมเดลให้ converge เพราะไม่ต้องปรับ ค่าถ่วงน้ำหนัก และ ค่าเบี่ยงเบนตั้งแต่เริ่มต้น หลังจากนั้นจะนำมาเชื่อมต่อกับส่วนชั้นการเชื่อมต่อกันอย่างสมบูรณ์ที่ผู้พัฒนาโมเดลสร้างขึ้นมาจากฐานข้อมูลดั้งเดิม โดยหลังจากทำการเชื่อมต่อกับโมเดลที่ผู้พัฒนาสร้างขึ้นเอง ชั้นตอนถัดไปจะเป็นการปรับรายละเอียด (Fine-tuning) โดยจะทำการละเว้น

การตรึง (Unfrozen) ในส่วนปลายของ pre-trained model ในชั้นที่ใกล้กับชั้นการเชื่อมต่อกันอย่างสมบูรณ์โดยโครงข่ายแบบคอนโวลูชันที่ถูกนำมาใช้สำหรับการเรียนรู้แบบถ่ายทอดมีหลากหลายโครงสร้าง เช่น VGG16, VGG19, MobileNetV2, MobileNetV3, Resnet50, Vgg16, Vgg19 และ Xception โดย weights และ biases ของโครงสร้างจะมาจากฐานข้อมูล imagenet ซึ่งมีรูปภาพความละเอียดสูงมากกว่า 15,000,000 และประเภทมากกว่า 22,000 ประเภท

### 2.1.11 Regularization [9]

#### 1. Data augmentation

Data augmentation เป็นกระบวนการสำคัญและใช้งานกันมากในงานด้าน Computer Vision ขั้นตอนในการทำงานคือใช้วิธีการนำรูปภาพเดิมมาทำการเลื่อน หมุน พลิก ตัดบางส่วนออก ใส่ Noise ลงไป หรือปรับแสงสี ความคมของภาพ เพื่อให้ได้ข้อมูลที่ต่างไปจากเดิมเล็กน้อย แต่ยังคงมีความหมายเหมือนเดิมเพื่อเพิ่มปริมาณของข้อมูลที่นำมาใช้ในการเทรนโมเดล ซึ่งนอกจากเป็นการขยายจำนวนข้อมูลแล้ว Augmentation ยังช่วยเพิ่มความหลากหลายของภาพที่จะนำไปฝึกอีกด้วย ดังภาพ



รูปที่ 9 ตัวอย่างภาพถ่ายรังสีทรวงอกที่ผ่านการทำ Data Augmentation เช่น การหมุนภาพ การซูมเข้าและออก



## 2. Batch Normalization

Batch Normalization เป็นเทคนิคในการปรับค่าข้อมูลให้อยู่ในขอบเขตที่กำหนด ก่อนส่งออก จาก โหนด ใน Neural Network Layer เป็นการนำเข้าของชั้นเลเยอร์ถัดไป การทำ Data Normalization จะทำให้ข้อมูลมีน้ำหนักเท่ากัน มีการกระจายตัวเหมือนกัน ไม่มีตัวหนึ่งตัวใดมีอิทธิพลมากกว่ากัน ทั้งยังเป็นการเพิ่มความเร็วในการฝึกโมเดลและทำให้ค่าเบี่ยงเบนลดลงเมื่อเทียบกันตอนที่ยังไม่ได้ทำ Normalization เพราะมีค่าข้อมูลที่เล็กกว่า

## 3. Dropout

Dropout เป็นเทคนิคในการทำ Regularization ที่เรียบง่าย แต่มีประสิทธิภาพอย่างมาก โดยเมื่อมีการใช้งาน Dropout ภายในเลเยอร์ที่กำหนดแล้ว โหนดในเลเยอร์นั้นจะถูกสุ่มเพื่อปิดการทำงานชั่วคราวในแต่ละรอบของการทำ Forward Propagation และ Back-propagation ตามอัตราที่กำหนด ในขณะที่มีการฝึกโมเดลทำให้ไม่มีการคำนวณค่าถ่วงน้ำหนักที่ถูกเชื่อมต่อกับโหนดที่กำลังถูกปิด

### 2.1.12 คอนฟิวชันเมทริกซ์ (Confusion Matrix)

คอนฟิวชันเมทริกซ์ คือ การประเมินผลลัพธ์การทำนายของโมเดล โดย จะนำผลลัพธ์ที่ทำนายได้มาเปรียบเทียบกับข้อมูลจริง สามารถวัดผลลัพธ์จากการจัดกลุ่มข้อมูล ค่าของผลลัพธ์ที่ได้คือ True Positive (TP), True Negative (TN), False Positive (FP) และ False Negative (FN)

		ค่าสังเกต	
		Negative	Positive
ค่าพยากรณ์	Negative	True negative (TN)	False positive (FP)
	Positive	False negative (FN)	True positive (TP)

- True positive (TP) คือจำนวนเหตุการณ์ที่สนใจมีผลเป็นบวก และมีผลจากการพยากรณ์เป็นบวก

- False positive (FP) คือจำนวนเหตุการณ์ที่สนใจมีผลเป็นลบ แต่มีผลจากการพยากรณ์เป็นบวก
- False negative (FN) คือจำนวนเหตุการณ์ที่สนใจมีผลเป็นบวก และมีผลจากการพยากรณ์เป็นลบ
- True negative (TN) คือจำนวนเหตุการณ์ที่สนใจมีผลเป็นลบ และมีผลจากการพยากรณ์เป็นลบ

ตัวสถิติที่ใช้วัดความถูกต้องของการพยากรณ์ซึ่งผลที่เกิดจากการพยากรณ์มี 4 ค่า คือ

Precision คือเป็นการวัดความแม่นยำของโมเดล โดยคิดแยกทีละคลาส

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

Recall คือเป็นการวัดความถูกต้องของโมเดล โดยพิจารณาแยกทีละคลาส

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

Accuracy คือเป็นการวัดความถูกต้องของโมเดล โดยพิจารณารวมกันทุกคลาส

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

F1-Score คือค่าเฉลี่ยแบบ harmonic mean ระหว่าง precision และ recall เป็น single metric ที่วัดความสามารถของโมเดล

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall}$$

## 2.2 เอกสารและงานวิจัยที่เกี่ยวข้อง

จากการศึกษาค้นคว้างานวิจัยที่เกี่ยวข้องกับการจำแนกประเภทภาพรังสีทรวงอกในหัวข้อต่าง ๆ พบว่า มีการใช้เทคนิคในการสร้างแบบจำลองที่หลากหลาย การปรับค่าพารามิเตอร์ภายในแบบจำลองที่แตกต่างกัน รวมไปถึงการใช้ค่าที่วัดประสิทธิภาพของแบบจำลอง อีกทั้งยังมีงานวิจัยที่เสนอวิธีการใช้ การเรียนรู้แบบถ่ายทอดกับชุดข้อมูลที่มีลักษณะใกล้เคียงกับชุดข้อมูลที่รวบรวมขึ้นในงานวิจัยครั้งนี้

### 2.2.1 งานวิจัยที่เกี่ยวข้องกับการคัดแยกโรคจากภาพถ่ายรังสีทรวงอก

Lakhani [6] นำเสนอการใช้ CNN 2 โมเดล คือ AlexNet และ GoogleNet เพื่อการคัดแยกระหว่างปอดของผู้ป่วยติดเชื้อวัณโรคทางปอด (Pulmonary Tuberculosis) และผู้ป่วยปกติผ่านภาพถ่ายรังสีทรวงอก โดยใช้เทคนิค Augmentation ควบคู่ไปด้วยเพื่อเพิ่มปริมาณของรูปภาพ และใช้ ค่า AUC โดยวิธี Delong มาเปรียบเทียบประสิทธิภาพของโมเดล โดยผลลัพธ์ของโมเดลที่มีค่า AUC สูงที่สุดคือการใช้โมเดล AlexNet และ GoogleNet ควบคู่กัน ซึ่งได้ค่า AUC ที่ 0.99% และการใช้ Augmentation ช่วยเพิ่มประสิทธิภาพของโมเดลให้มีการทำนายแม่นยำมากขึ้น

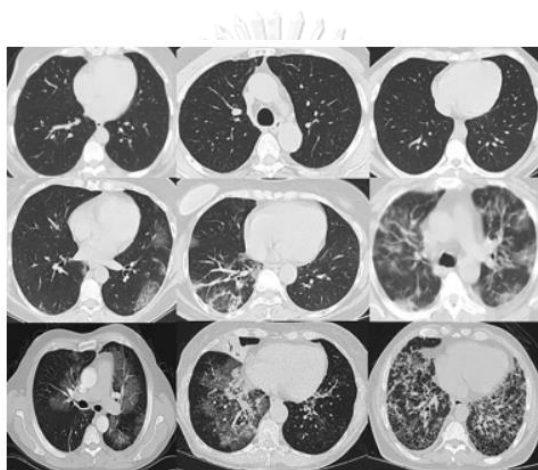
Singh และ Gupta [14] นำเสนอการคัดแยกผู้ป่วยที่เป็นมะเร็งปอด และ ผู้ป่วยปกติ ผ่านภาพถ่ายรังสีทรวงอก โดยนำโมเดล Machine learning กว่า 7 โมเดล ได้แก่ k-nearest neighbors classifier, support vector machine classifier, decision tree classifier, multinomial naive Bayes classifier, stochastic gradient descent classifier, random forest classifier และ multi-layer perceptron (MLP) classifier มาเปรียบเทียบประสิทธิภาพกัน โดยใช้ค่าความแม่นยำมาเป็นตัวชี้วัดประสิทธิภาพของโมเดล โดยใช้รูปภาพ 15,750 ภาพ โดยแบ่งเป็น ภาพถ่ายรังสีทรวงอกของผู้ป่วยที่เป็นมะเร็งปอด 8,840 ภาพ และภาพถ่ายรังสีทรวงอกของผู้ป่วยปกติ 6,910 ภาพ และแบ่งข้อมูลเป็น 2 ชุดข้อมูล คือ ชุดข้อมูลการฝึก (Train set) และชุดข้อมูลการทดสอบ (Test set) โดย multi-layer perceptron (MLP) classifier เป็นโมเดลที่มีค่าความแม่นยำสูงที่สุดเมื่อเทียบกับตัวอื่น โดยมีค่าความแม่นยำอยู่ที่ 88.55 %

Jawad และคณะ [5] นำเสนอการคัดแยกผู้ป่วยที่ติดเชื้อโควิด 19 และผู้ป่วยปกติ ผ่านภาพถ่ายรังสีทรวงอก โดยใช้รูปภาพ 500 ภาพ โดยแบ่งเป็น ภาพถ่ายรังสีทรวงอกของผู้ป่วยติดเชื้อโควิด 250 ภาพ และภาพถ่ายรังสีทรวงอกของผู้ป่วยปกติ 250 ภาพ โดยใช้โมเดล Logistic regression และ CNN ที่ใช้ Relu activation function ทั้งในคอนโวลูชันเลเยอร์และชั้นการเชื่อมต่อกันอย่างสมบูรณ์ควบคู่ไปกับการใช้ Dropout โดยมี optimizer เป็น Adam optimizer มาเปรียบเทียบประสิทธิภาพกัน นอกจากนี้ยังใช้เทคนิค GAN Augmentation มาเพิ่มปริมาณรูปภาพ และเทคนิค PCA มาช่วยลดทอนข้อมูลที่มีความสำคัญน้อย โดย Logistic Regression และ CNN ที่ไม่ใช้เทคนิค PCA มีความแม่นยำระหว่าง 95.2% ถึง 97.6% และที่ใช้เทคนิค PCA อยู่ที่ระหว่าง 97.6% ถึง 100%

Ucar และ Korkmaz [15] นำเสนอการคัดแยกผู้ป่วยที่ติดเชื้อโควิด 19 ผู้ป่วยที่เป็นปอดอักเสบจากไวรัส และผู้ป่วยปกติ ผ่านภาพถ่ายรังสีทรวงอก โดยใช้รูปภาพ 5310 ภาพ โดยแบ่งเป็นภาพถ่ายรังสีทรวงอกของผู้ป่วยติดเชื้อโควิด 66 ภาพ และภาพถ่ายรังสีทรวงอกของผู้ป่วยปกติ 1349 ภาพ และภาพถ่ายรังสีทรวงอกของผู้ป่วยที่เป็นปอดอักเสบจากไวรัส 3895 ภาพ และใช้เทคนิค Augmentation มาเพิ่มจำนวนรูปภาพ โดยใช้โมเดล Deep Bayes-SqueezeNet และ confusion matrix มาเป็นตัววัดประสิทธิภาพ โดยมีค่าความแม่นยำในทั้ง 3 คลาส ที่ 98.26 %

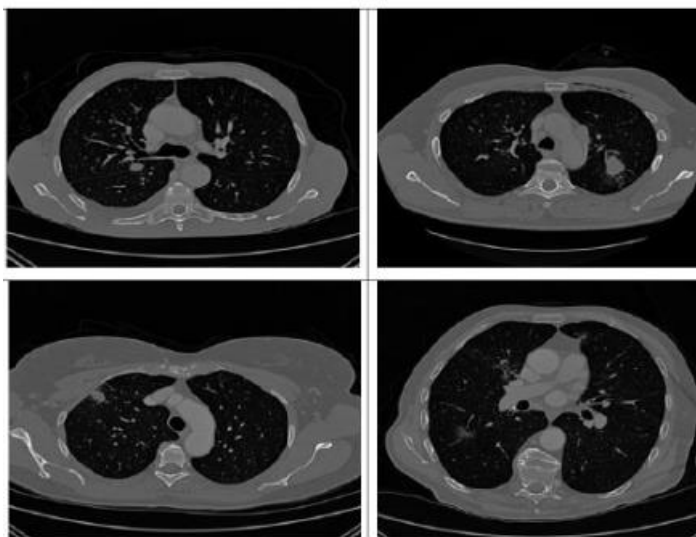
Dey และคณะ [16] นำเสนอการคัดแยกผู้ป่วยที่เป็นมะเร็งปอด และผู้ป่วยปกติผ่านภาพถ่ายรังสีทรวงอกจากการทำ CT Scan ซึ่งเป็นภาพ 3 มิติ จาก ข้อมูลภาพ LIDC-IDRI โดยใช้ 4 แบบจำลองคือ 3D CNN, a novel multi-output network, a 3D DenseNet, และ an augmented 3D DenseNet with multi-outputs และนำเทคนิคการถ่ายทอดการเรียนรู้มาใช้กับแบบจำลองทั้ง 4 ด้วย และใช้ค่าความแม่นยำ และ AUC score มาวัดประสิทธิภาพของแบบจำลอง ซึ่งโมเดลที่ใช้เทคนิคการเรียนรู้การถ่ายทอดมีค่าความแม่นยำ และ AUC score ที่สูงกว่าโมเดลที่ไม่ใช้ และ a novel multi-output network หรือ MoDenseNet มีประสิทธิภาพของแบบจำลองที่ดีที่สุด โดยมีความแม่นยำบนข้อมูลภาพ LIDC-IDRI ที่ 90.40% และ AUC score ที่ 0.95%

Joshi และคณะ [17] นำเสนอการใช้โครงข่ายคอนโวลูชันควบคู่กับการใช้เทคนิคการเรียนรู้ผ่านการถ่ายทอดเพื่อคัดแยกระหว่างผู้ป่วยที่ติดเชื้อโควิด19 ผู้ป่วยที่เป็นปอดอักเสบจากไวรัส และผู้ป่วยปกติ ผ่านภาพถ่ายรังสีทรวงอกจากการทำ CT Scan โดยใช้รูปภาพทั้งหมด 1200 ภาพ แบ่งเป็นชุดละ 400 ภาพ โดยได้นำแบบจำลอง VGG16 และ ResNet50 มาใช้เป็นแบบจำลองพื้นฐาน และนำเทคนิค data augmentation เช่น Horizontal flip Vertical flip Rotation range และเทคนิค fine-tuning มาใช้ด้วย ซึ่งผลทดสอบสรุปได้ว่า ResNet50 มีความแม่นยำมากกว่าอยู่ที่ 88.52% ในขณะที่ VGG-16 มีความแม่นยำอยู่ที่ 86.74%



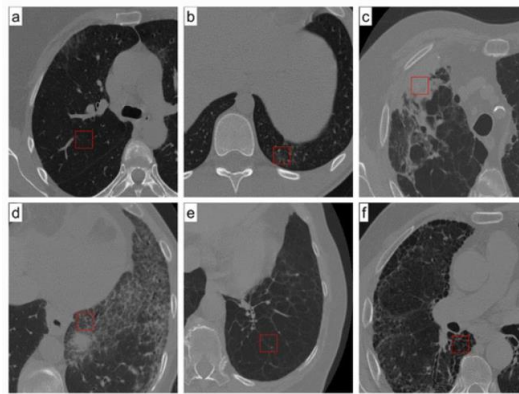
รูปที่ 10 ตัวอย่างชุดข้อมูลภาพถ่ายรังสีทรวงอก ในการจำแนกผู้ป่วยที่ติดเชื้อโควิด19 ผู้ป่วยที่เป็นปอดอักเสบจากไวรัส และผู้ป่วยปกติ

Cinar และ คณะ [18] นำแบบจำลอง AlexNet, ResNet18, Googlenet, และ ResNet50 ควบคู่กับเทคนิคการเรียนรู้การถ่ายทอด และ data augmentation มาใช้ เพื่อคัดแยกผู้ป่วยที่เป็นมะเร็งปอด และ ผู้ป่วยปกติผ่านภาพถ่ายรังสีทรวงอกจากการทำ CT Scan โดยใช้ภาพ 1018 ภาพ จากชุดข้อมูล LIDC-IDRI และแบ่งชุดข้อมูลฝึกที่ 70 % และข้อมูลทดสอบที่ 30 % โดย AlexNet มีค่าความแม่นยำมากที่สุดที่ 98.52 % อีกทั้งยังใช้เวลาในการฝึกฝนน้อยที่สุดที่ 64 นาที ในขณะที่ ResNet18, Googlenet, และ ResNet50 มีค่าความแม่นยำอยู่ที่ 97.97%, 94.10%, และ 97.05%



รูปที่ 11 ตัวอย่างชุดข้อมูลภาพถ่ายรังสีทรวงอก เพื่อคัดแยก  
ผู้ป่วยที่เป็นมะเร็งปอด และ ผู้ป่วยปกติ

Kim และ คณะ [19] นำเสนอการคัดแยกผู้ป่วยปกติ กับ ผู้ป่วยที่เป็นโรคเกี่ยวกับปอด 5 โรค คือ ground-glass opacity, consolidation, reticular opacity, emphysema, และ honeycombing ผ่านภาพถ่ายรังสีทรวงอก จากการทำ CT Scan โดยใช้แบบจำลองแบบคอนโวลูชัน โดยมีชั้นคอนโวลูชัน 4 ชั้น และชั้นการเชื่อมต่อกันอย่างสมบูรณ์ 2 ชั้น และแบบจำลอง a shallow learning of a support vector machine (SVM) โดย แบบจำลอง CNN มีความแม่นยำมากกว่า SVM 6-9% ในขณะที่เมื่อเพิ่มจำนวนชั้นคอนโวลูชันขึ้น ค่าความแม่นยำของ CNN เพิ่มขึ้นจาก 81.27 ถึง 95.12% โดยเฉพาะแบบจำลองที่เปรียบเทียบระหว่างกลุ่มผู้ป่วยปกติ และผู้ป่วยที่เป็น emphysema ผู้ป่วยที่เป็น honeycombing และผู้ป่วยที่เป็น reticular opacity



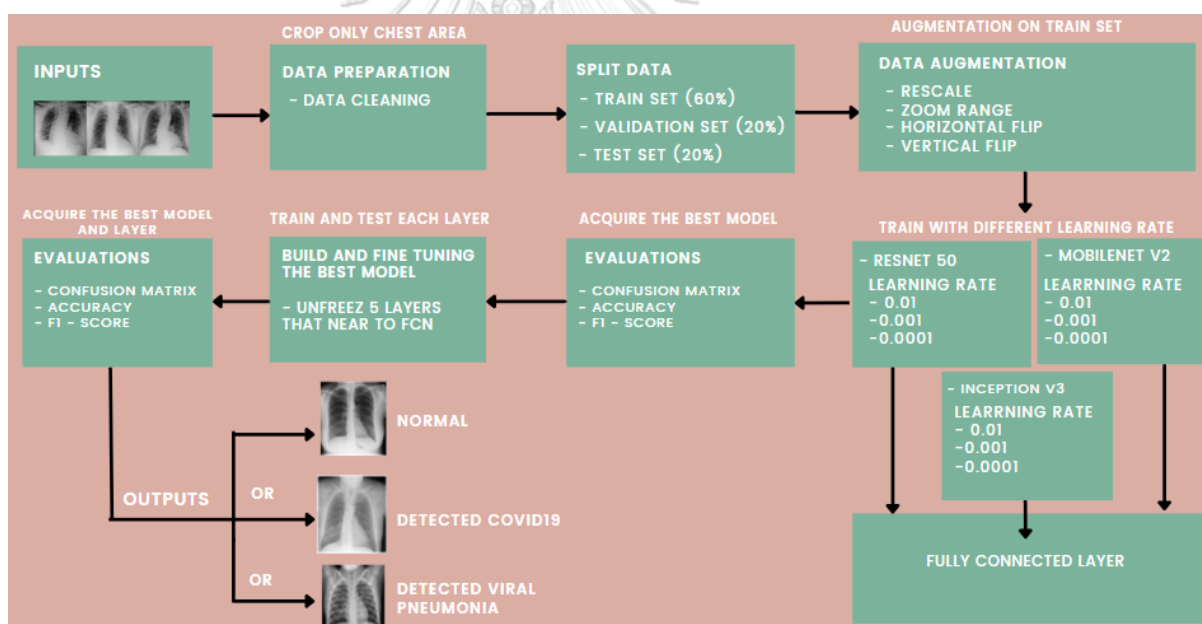
รูปที่ 12 ตัวอย่างชุดข้อมูลภาพถ่ายรังสีทรวงอก ผู้ป่วยที่เป็นโรคเกี่ยวกับปอด 5 โรค และผู้ป่วยปกติ



### บทที่ 3

## แนวคิดและวิธีการวิจัย

งานวิจัยนี้นำเสนอการสร้างแบบจำลองด้วยโครงข่ายแบบคอนโวลูชันร่วมกับเทคนิคการเรียนรู้การถ่ายทอด เพื่อการจำแนกภาพถ่ายรังสีทรวงอกออกเป็น 3 ประเภท คือ 1) ภาพถ่ายรังสีทรวงอกของผู้ป่วยปกติ 2) ภาพถ่ายรังสีทรวงอกของผู้ป่วยที่ติดเชื้อโควิด 3) ภาพถ่ายรังสีทรวงอกของผู้ติดเชื้อปอดอักเสบจากไวรัส ลำดับของการวิจัยจะแบ่งเป็น 4 ขั้นตอนหลัก คือ 1) การเตรียมชุดข้อมูลรูปภาพ 2) การเปรียบเทียบแบบจำลองการเรียนรู้การถ่ายทอดจาก pre-trained โมเดล 3) การปรับแต่งประสิทธิภาพของแบบจำลองที่มีประสิทธิภาพมากที่สุด 4) การประเมินประสิทธิภาพของแบบจำลองการจำแนกภาพถ่ายรังสีทรวงอก



รูปที่ 32 แผนผัง โครงสร้างวิธีการดำเนินงานวิจัย



### 3.1 การเตรียมชุดข้อมูล

เก็บรูปภาพถ่ายรังสีทรวงอกจากแหล่งข้อมูลออนไลน์ โดยคัดเลือกภาพถ่ายรังสีทรวงอกออกเป็น 3 ประเภท คือ 1) ภาพถ่ายรังสีทรวงอกของผู้ป่วยปกติ 2) ภาพถ่ายรังสีทรวงอกของผู้ป่วยที่ติดเชื้อโควิด 19 3) ภาพถ่ายรังสีทรวงอกของผู้ติดเชื้อปอดอักเสบจากไวรัส จาก <https://www.kaggle.com/tawsifurrahman/covid19-radiography-database> ดังแสดงในรูปที่ 13



1) ปอดของผู้ป่วยปกติ

2) ปอดของผู้ป่วยที่ติดเชื้อโควิด19

3) ปอดของผู้ป่วยที่ติดเชื้อปอดบวมจากไวรัส

รูปที่ 13 ตัวอย่างชุดข้อมูลภาพถ่ายรังสีทรวงอก 3 ประเภท ที่ใช้ในการฝึกฝนแบบจำลอง

ในการฝึกสอนและทดสอบแบบจำลอง จะทำการแบ่งข้อมูลออกเป็น ข้อมูลสอน ข้อมูลตรวจสอบ และข้อมูลทดสอบ ดังตารางที่ 2

ตารางที่ 2 ตารางสรุปจำนวนข้อมูลสอน ข้อมูลตรวจสอบ ข้อมูลทดสอบของภาพถ่ายรังสีทรวงอก ทั้งสามประเภท

ประเภทรูปภาพ	ข้อมูลสอน	ข้อมูลตรวจสอบ	ข้อมูลทดสอบ	จำนวนทั้งหมด
ปอดของผู้ป่วยปกติ	700	230	230	1160
ปอดของผู้ป่วยที่ติดเชื้อโควิด19	700	230	230	1160
ปอดของผู้ป่วยที่ติดเชื้อปอดอักเสบจากไวรัส	700	230	230	1160
<b>รวม</b>	<b>2100</b>	<b>690</b>	<b>690</b>	<b>3480</b>

### 3.2 การสร้างแบบจำลองเพื่อเปรียบเทียบประสิทธิภาพแบบจำลองการเรียนรู้การถ่ายทอด

ทำการทดสอบแบบจำลองการเรียนรู้ถ่ายโอนทั้ง 3 แบบ คือ MobileNetV2, ResNet-50 และ InceptionV3 โดยใช้ pre-trained weight จาก ImageNet หลังจากนั้นทำการละเว้นการตรึงชั้นการเชื่อมต่อกันอย่างสมบูรณ์เพื่อที่จะให้มีการอัปเดตค่าถ่วงน้ำหนักรวมไปถึงค่าเบี่ยงเบนและทำการตรึงชั้นชั้นคอนโวลูชันในทุกชั้นของ pre-trained model และทำการปรับชั้นผลลัพธ์ให้ตรงกับคลาสที่เราต้องการ และทำการบันทึกผลแบบจำลองทั้ง 3 แบบ

### 3.3 การวัดประสิทธิภาพของแบบจำลอง

การวัดประสิทธิภาพของแบบจำลองในงานวิจัยนี้ จะทำการเปรียบเทียบค่าความแม่นยำของชุดข้อมูลทดสอบ และเลือกแบบจำลองที่มีค่าความแม่นยำมากที่สุดจาก learning rate ที่ต่างกัน และ epoch ที่ดีที่สุด โดยการวัดประสิทธิภาพของแบบจำลองบนชุดข้อมูลทดสอบ จะใช้ค่าสถิติทดสอบ confusion matrix จากนั้นนำไปคำนวณค่าของผลลัพธ์ที่ได้คือ True Positive (TP), True Negative (TN), False Positive (FP) และ False Negative (FN)

### 3.4 การปรับละเอียดสมรรถนะของแบบจำลอง

ทำการสร้างแบบจำลองโครงข่ายประสาทคอนโวลูชัน ร่วมกับ pre-trained โมเดลที่มีประสิทธิภาพที่ดีที่สุดจากภาพถ่ายรังสีทรวงอก 3 คลาส และทำการปรับแบบจำลองโดยการ fine-tuning เป็นการปรับชั้นคอนโวลูชัน โดยทำการละเว้นการตรึงในชั้นที่ใกล้กับชั้นการเชื่อมต่อกันอย่างสมบูรณ์เพื่อให้มีการอัปเดตค่าถ่วงน้ำหนักรวมไปถึงค่าเบี่ยงเบนและจะทดลองการตรึงในชั้นที่แตกต่างกันเพื่อหาชั้นการตรึงที่ให้ค่าความแม่นยำดีที่สุด

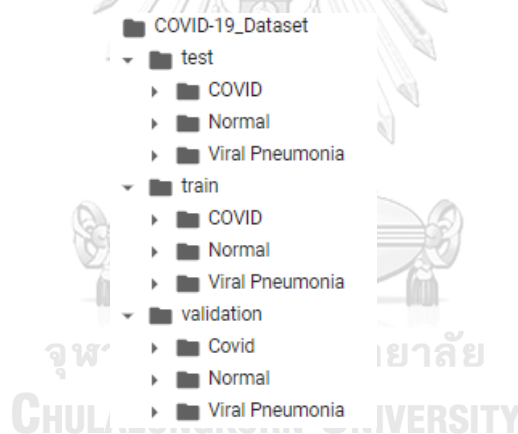
### 3.5 สภาพแวดล้อมที่ใช้ในการทดสอบแบบจำลอง

ใช้ Google Colaboratory หรือ Google Colab โดยมีโฮสต์โปรแกรมเป็น Jupyter Notebook ทำงานบน Cloud โดยชื่อเต็ม ซึ่งใช้ภาษา Python เป็นภาษาหลักที่ใช้ในการเขียนและรันโมเดล โดย Google Colab เป็นบริการที่อำนวยความสะดวกการพัฒนาซอฟต์แวร์เป็นอย่างมาก เนื่องจากการสร้างโมเดลจะต้องใช้คอมพิวเตอร์ที่มีความเร็ว หรือ มีประสิทธิภาพสูงเพื่อลดระยะเวลาในการประมวลผลของเครื่อง โดยผู้ใช้สามารถที่จะเลือกใช้บริการแบบมีค่าใช้จ่าย (Google Colab Pro)

หรือไม่มีค่าใช้จ่ายก็ได้ ซึ่งแบบมีค่าใช้จ่ายจะมีการประมวลผลที่เร็วกว่าหลายเท่า และมีพื้นที่ในการจัดเก็บข้อมูลมากกว่า รวมไปถึงฟังก์ชันพิเศษที่สามารถรันโมเดลทิ้งไว้ในการพัฒนาแบบจำลอง โดยร่วมกับการใช้ไลบรารีจาก Tensorflow, Scikit-learn และ Pandas

### 3.6 การเตรียมชุดข้อมูลที่ใช้ในการทดสอบแบบจำลอง

ทำการอัปโหลดชุดข้อมูลฝึกสอน ชุดข้อมูลตรวจสอบ และชุดข้อมูลทดสอบ ที่อยู่ภายใต้โฟลเดอร์ COVID19-Dataset ที่ Google Drive ซึ่งเป็นบริการเก็บข้อมูลแบบ cloud สามารถเข้าถึงชุดข้อมูลได้จากทุกที่ โดยมีโครงสร้างโฟลเดอร์เป็น test, train, และ validation โดยในแต่ละโฟลเดอร์มี โฟลเดอร์ Covid, Normal และ Viral Pneumonia ซึ่งเป็นที่เก็บข้อมูลของแต่ละคลาส ดังรูปที่ 14



รูปที่ 14 โครงสร้างชุดข้อมูลในการจำแนกภาพถ่ายรังสีทรวงอก

ในขั้นตอนการเรียกใช้รูปภาพจาก Google drive จำเป็นต้องมีการ Mount drive เพื่อดึงข้อมูลมาใช้ใน google colab ใช้คำสั่ง `drive.mount('/content/drive')` ภายใต้ไลบรารี `google.colab` และทำการ unzip ไฟล์ด้วยคำสั่ง `unzip`

สำหรับการนำเข้าข้อมูลที่อยู่ในไดเรกทอรีของ google drive ที่ถูกแบ่งเป็นชุดข้อมูลฝึกสอนชุดข้อมูลตรวจสอบ และชุดข้อมูลทดสอบนั้น จะอยู่ใน `train`, `validation` และ `test` ตามลำดับ โดยมีการกำหนดไดเรกทอรีเพื่อนำเข้าข้อมูลฝึกสอน ข้อมูลตรวจสอบและข้อมูลทดสอบ เพื่อให้ข้อมูลภาพถ่ายรังสีทรวงอกเป็นข้อมูลชุดเดียวกัน

เนื่องจากชุดข้อมูลมีขนาดเล็ก จึงใช้เทคนิคการเพิ่มจำนวนข้อมูล (data augmentation) ดังรูปที่ 15 สำหรับชุดข้อมูลที่ใช้ในการฝึกสอนแบบจำลอง (training dataset) ด้วยการ `rescale`, `zoom` และ `flip` ผ่านคำสั่ง `ImageDataGenerator` โดยมีการกำหนดค่าให้พารามิเตอร์ ดังนี้

1. ทำการ Normalized ข้อมูลแต่ละพิกเซลของรูปภาพให้อยู่ช่วง 0 ถึง 1 ด้วยพารามิเตอร์ (`Rescale = 1/255`)
2. ชุมอินพุต (`zoom_range = 0.2`)
3. พลิกอินพุตในแนวนอน (`horizontal_flip = True`)
4. พลิกอินพุตในแนวตั้ง (`vertical_flip = True`)

สำหรับชุดข้อมูลตรวจสอบ (validation dataset) และชุดข้อมูลทดสอบ (test dataset) จะไม่ใช้เทคนิคการเพิ่มจำนวนข้อมูล (data augmentation) แต่จะทำการทำให้ข้อมูลแต่ละพิกเซลของรูปภาพให้อยู่ช่วง 0 ถึง 1 ด้วยพารามิเตอร์ (Normalization) เท่านั้น หลังจากนั้นทำการกำหนด flow ของแต่ละชุดโดยมีการกำหนด path ของชุดข้อมูลต่าง ๆ กำหนดขนาดของรูปภาพเป็น `256x256` และกำหนด `batch_size` เท่ากับ 32 ซึ่งคือจำนวนข้อมูลที่น่าไปใช้ในการฝึกสอนแต่ละรอบ

```

image_gen_train = ImageDataGenerator(
    rescale=1./255,
    zoom_range=0.2,
    horizontal_flip=True,
    vertical_flip=True)

train_data_gen = image_gen_train.flow_from_directory(batch_size=BATCH_SIZE,
    directory=PATH + "/train/",
    shuffle=True,
    target_size=(IMG_SHAPE,IMG_SHAPE),
    class_mode='categorical')

=====
image_gen_val = ImageDataGenerator(rescale=1./255)

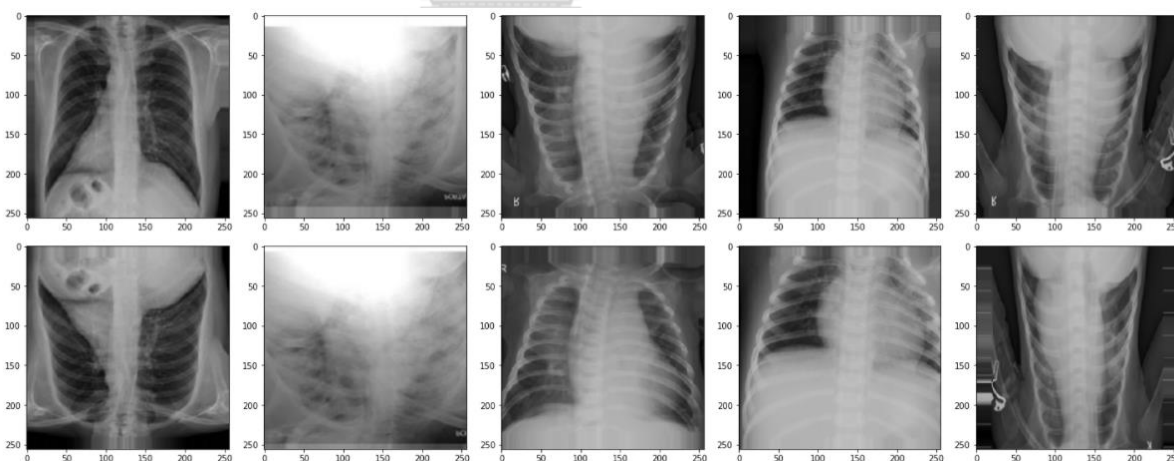
val_data_gen = image_gen_val.flow_from_directory(batch_size=BATCH_SIZE,
    directory=PATH + "/validation/",
    target_size=(IMG_SHAPE, IMG_SHAPE),
    class_mode='categorical')

=====
image_gen_test = ImageDataGenerator(rescale=1./255)

test_data_gen = image_gen_val.flow_from_directory(batch_size=BATCH_SIZE,
    directory=PATH + "/test/",
    target_size=(IMG_SHAPE, IMG_SHAPE),
    class_mode='categorical')

```

รูปที่ 15 การเพิ่มจำนวนข้อมูล (data augmentation) สำหรับชุดข้อมูลที่ใช้ในการทดสอบ



รูปที่ 16 ตัวอย่างข้อมูลที่ได้จากการทำ data augmentation

### 3.7 การดำเนินการสร้างแบบจำลอง

ทำการนำเข้าไลบรารีที่จำเป็นต่อการสร้างแบบจำลอง จาก tensorflow โดยจะนำมาใช้ในการสร้างแบบจำลองแบบจำลองโครงข่ายประสาทคอนโวลูชัน โดยจะนำเข้าโมเดล pre-trained ทั้งหมด และนำมาใช้สร้างชั้นการเชื่อมต่อกันอย่างสมบูรณ์สำหรับจำแนกประเภทภาพถ่ายรังสีทรวงอกของผู้ป่วยปกติ ผู้ป่วยที่ติดเชื้อโควิด19 และผู้ป่วยปอดอักเสบจากการติดเชื้อของไวรัสให้กับแบบจำลองเรียนรู้ถ่ายทอด

#### 3.7.1 แบบจำลองโครงข่ายประสาทคอนโวลูชันแบบอินเซปชันเวอร์ชัน3 (InceptionV3)

ทำการนำเข้าไลบรารีแบบจำลองการเรียนรู้แบบถ่ายทอดแบบอินเซปชันเวอร์ชัน3 จากนั้นทำการรับข้อมูลนำเข้าขนาด 250x250x3 โดยนำเข้าตามความกว้าง ความสูง และจากชั้นสีแบบ RGB ของรูปภาพ โดยไม่ต้องการชั้นการเชื่อมต่อกันอย่างสมบูรณ์เพราะจะนำแบบจำลองการเรียนรู้ถ่ายทอดมาใช้ เนื่องจากชุดข้อมูลที่ต้องการจำแนกมีลักษณะการจำแนกต่างกับ โครงสร้างแบบจำลองดั้งเดิม โดยจะทำการตั้งค่าพารามิเตอร์ include\_top ให้เป็น False และนำค่าถ่วงน้ำหนักจาก imagenet ที่ถูกฝึกสอนมาใช้กับแบบจำลอง ดังรูปที่ 17 โดยในการสร้างแบบจำลองการเรียนรู้ถ่ายโอนครั้งนี้จะต้องทำการตรึงค่าถ่วงน้ำหนักโดยการตั้งค่าให้แต่ละเลเยอร์ถูกตรึงค่าให้ trainable เป็น False ของชั้นคอนโวลูชันทั้งหมด เพื่อที่จะให้ไม่มีการฝึกสอนซ้ำ ดังรูปที่ 18

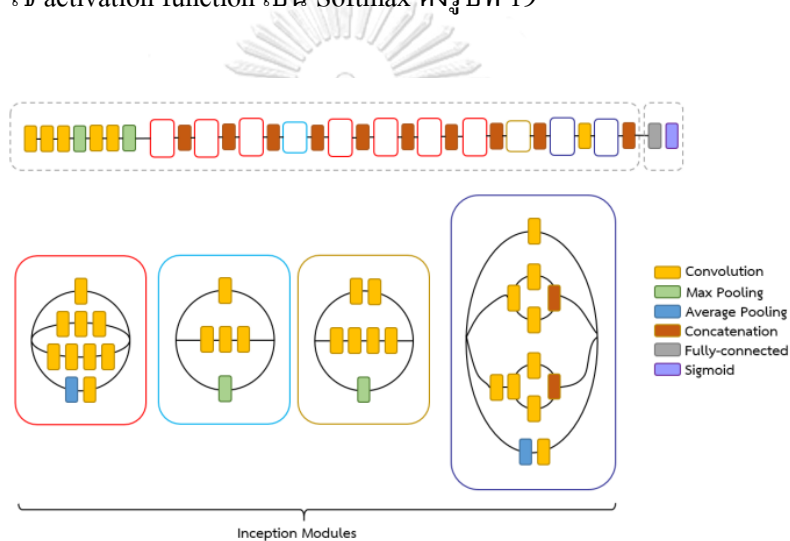
```
base_model = tf.keras.applications.InceptionV3(input_shape=IMG_SHAPE,
                                              include_top=False,
                                              weights='imagenet')
```

รูปที่ 17 การเรียกใช้ pre-trianed weights จากแบบจำลอง inceptionV3

```
for layer in base_model.layers:
    layer.trainable = False
```

รูปที่ 18 การตรึงค่าของ *pre-trained model* ไม่ให้มีการฝึกสอนซ้ำ

จากนั้นใช้ชั้น Fully-connected layer (FCN) เหมือนแบบจำลองโครงข่ายประสาทเทียมแบบคอนโวลูชันมาเชื่อมต่อในชั้นถัดไป ซึ่งจะรับข้อมูลจากชั้นก่อนหน้า โดยชั้นสุดท้ายคือชั้น output ใช้ชั้น dense ขนาด 3 ใช้ activation function เป็น Softmax ดังรูปที่ 19



รูปที่ 19 โครงสร้างของโครงข่ายประสาทคอนโวลูชันแบบอินเซปชันเวอร์ชัน 3 (*InceptionV3*)

สำหรับขนาดของโครงสร้างของแบบจำลองโครงข่ายประสาทแบบอินเซปชันเวอร์ชัน 3 หลังจากทำการปรับปรุงโครงสร้างให้เหมาะกับแบบจำลองการจำแนกภาพถ่ายรังสีทรวงอกจากหัวข้อ 3.3 ในงานวิจัยนี้แสดงดังรูปที่ 20

Model: "sequential"

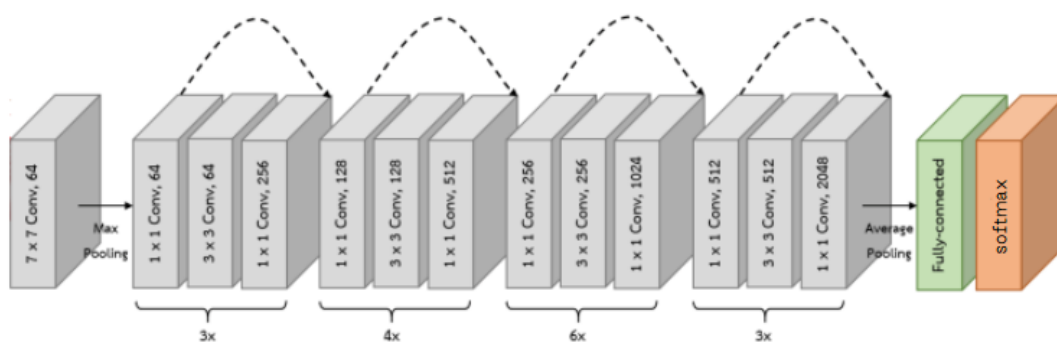
Layer (type)	Output Shape	Param #
inception_v3 (Functional)	(None, 6, 6, 2048)	21802784
global_average_pooling2d (GlobalAveragePooling2D)	(None, 2048)	0
dense (Dense)	(None, 3)	6147

=====  
 Total params: 21,808,931  
 Trainable params: 6,147  
 Non-trainable params: 21,802,784

รูปที่ 20 โครงสร้างแบบจำลองและจำนวนพารามิเตอร์ทั้งหมดของโครงข่ายคอนโวลูชันแบบ  
*inceptionV3*

### 3.7.2 แบบจำลองโครงข่ายประสาทเชิงลึกแบบ Residual Network (ResNet50)

สำหรับการนำเข้าไลบรารีแบบจำลองการเรียนรู้แบบถ่ายโอนแบบโครงข่ายประสาทเชิงลึกแบบ Residual Network 50 ชั้น ไปจนถึงการการสร้างแบบจำลองหลังจากทำการปรับปรุงโครงสร้างให้เหมาะสมกับข้อมูลจากหัวข้อ 3.3 มีขั้นตอนการสร้างเหมือนกับแบบจำลองโครงข่ายประสาทเทียมแบบอินเซปชันเวอร์ชัน 3 สำหรับขนาดของโครงสร้างของแบบจำลองเป็นไปดัง รูปที่ 21

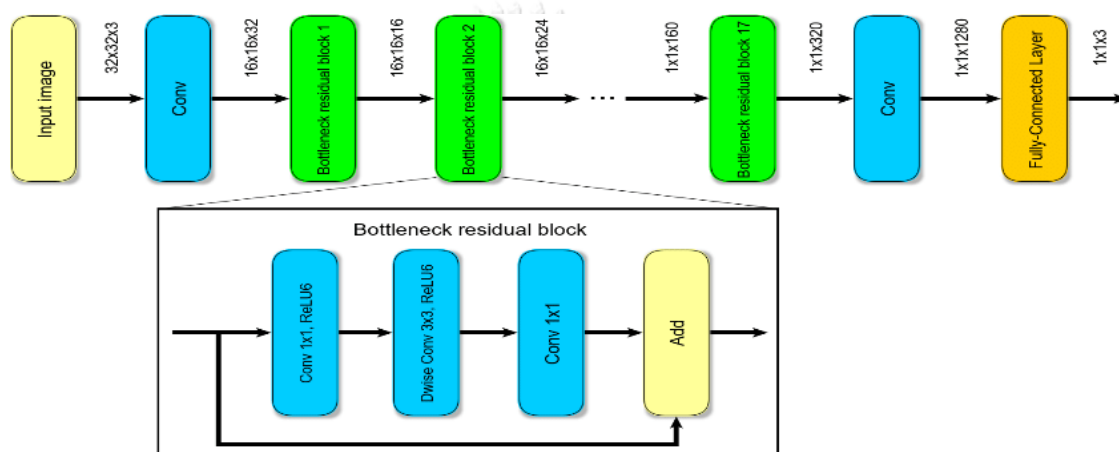


รูปที่ 21 โครงสร้างของโครงข่ายประสาทเชิงลึกแบบ Residual Network (ResNet50)



### 3.7.3 แบบจำลองโครงข่ายประสาทคอนโวลูชันแบบโมบายล์เน็ตเวอร์ชัน 2 (MobileNetV2)

สำหรับการนำเข้าไลบรารีแบบจำลองการเรียนรู้แบบถ่ายโอนแบบโครงข่ายประสาทแบบ MobileNetV2 ไปจนถึงการการสร้างแบบจำลองหลังจากทำการปรับปรุงโครงสร้างให้เหมาะสมให้เหมาะกับข้อมูลจากหัวข้อ 3.3 มีขั้นตอนการสร้างเหมือนกับแบบจำลองโครงข่ายประสาทเทียมแบบอินเซปชันเวอร์ชัน 3 สำหรับขนาดของโครงสร้างของแบบจำลองเป็นไปดังรูปที่ 22



รูปที่ 22 โครงสร้างของแบบจำลองโครงข่ายประสาทคอนโวลูชันแบบโมบายล์เน็ตเวอร์ชัน 2 (MobileNetV2)

## บทที่ 4

### ผลการวิจัย

#### 4.1 การฝึกสอนแบบจำลอง

แบบจำลองทั้ง 3 โครงสร้าง มีการคำนวณ loss แบบ categorical crossentropy และมีการปรับเปลี่ยนค่า learning rate เป็น 0.01, 0.001 และ 0.0001 ตามลำดับ ในการวัดสมรรถนะในการฝึกสอนและตรวจสอบแบบจำลองจะใช้ค่าความแม่นยำ (accuracy) เป็นตัววัด โดยใช้ Adaptive moment estimation optimizer (Adam) เป็น optimizer แล้วทำการ compile แบบจำลอง หลังจากการ compile แบบจำลองสำเร็จ จะทำการ fit แบบจำลอง โดยกำหนด batch\_size = 32 และทำการฝึกสอนจำนวน 30 epoches ดังรูปที่ 23 ในการฝึกสอนแต่ละรอบใช้ค่า step\_per\_epoch จากการนำจำนวนข้อมูลฝึกสอนทั้งหมดหารด้วย batch\_size โดยนำ ModelCheckpoint มาทำการเก็บ weight ของการฝึกสอนแบบจำลอง ที่ให้ค่าความแม่นยำ (accuracy) ที่ดีที่สุด โดยตั้งค่า mode = max และทำการบันทึกใน google drive ที่ได้ทำการ mount ไว้สมรรถนะของการฝึกสอนแบบจำลองทั้ง 3 โดยใช้ model.evaluate ในการวัดสมรรถนะของแบบจำลองขณะฝึกสอน ดังรูปที่ 25

```

model.compile(optimizer=tf.keras.optimizers.Adam(learning_rate=base_learning_rate),
              loss=tf.keras.losses.CategoricalCrossentropy(from_logits=False),
              metrics=['accuracy'])
=====
%%time
epochs = 30
history = model.fit_generator(
    train_data_gen,
    callbacks=[model_checkpoint_callback],
    steps_per_epoch=int(np.ceil(2100 / float(BATCH_SIZE))),
    epochs=epochs,
    validation_data=val_data_gen,
    validation_steps=int(np.ceil(690 / float(BATCH_SIZE))))

```

รูปที่ 23 ฟังก์ชันในการฝึกสอนแบบจำลอง

```

acc = history.history['accuracy']
val_acc = history.history['val_accuracy']

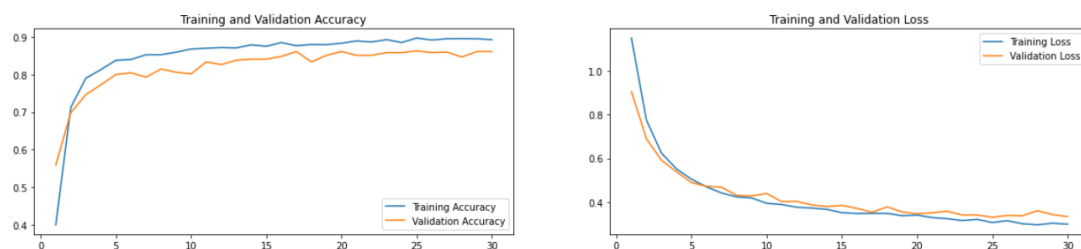
loss = history.history['loss']
val_loss = history.history['val_loss']

epochs_range = range(1, epochs+1, 1)

plt.figure(figsize=(20, 4))
plt.subplot(1, 2, 1)
plt.plot(epochs_range, acc, label='Training Accuracy')
plt.plot(epochs_range, val_acc, label='Validation Accuracy')
plt.legend(loc='lower right')
plt.title('Training and Validation Accuracy')

plt.subplot(1, 2, 2)
plt.plot(epochs_range, loss, label='Training Loss')
plt.plot(epochs_range, val_loss, label='Validation Loss')
plt.legend(loc='upper right')
plt.title('Training and Validation Loss')
plt.show()

```



รูปที่ 15 ฟังก์ชันและกราฟแสดงประสิทธิภาพของการฝึกสอน โมเดลใน 30 epochs

```

loss, accuracy = model.evaluate(test_data_gen)
print('Test accuracy :', accuracy)

```

```

22/22 [=====] - 17s 283ms/step - loss: 0.2204 - accuracy: 0.9348
Test accuracy : 0.9347826242446899

```

รูปที่ 16 ฟังก์ชันในการวัดสมรรถนะของการฝึกสอนแบบจำลอง

จุฬาลงกรณ์มหาวิทยาลัย

โดยจะทำการเปรียบเทียบค่าความแม่นยำ ค่าความสูญเสีย เวลาในการฝึกสอนแบบจำลองรวมไปจนถึงจำนวนพารามิเตอร์ของแต่ละแบบจำลอง ดังตารางที่ 3

ตารางที่ 3 ประสิทธิภาพในการฝึกสอนแบบจำลองการเรียนรู้การถ่ายทอดแบบ InceptionV3, MobileNetV2 และ ResNet50

Models	Learning rate	Total parameters	Training time (s)	Final Loss	Final Accuracy (%)
InceptionV3	0.01	21,808,931	1424	0.2204	93.47%
MobileNetV2		2,261,827	4021	0.3115	92.17%
ResNet50		23,593,859	1463	0.3922	88.41%
InceptionV3	0.001	21,808,931	<b>1044</b>	<b>0.1687</b>	93.47%
MobileNetV2		2,261,827	3879	0.2103	91.45%
ResNet50		23,593,859	1417	0.6298	74.78%
InceptionV3	0.0001	21,808,931	1308	0.2167	93.33%
MobileNetV2		2,261,827	4018	0.1944	<b>93.62%</b>
ResNet50		23,593,859	1355	0.9387	66.23%

#### 4.2 ผลการทดสอบแบบจำลอง

จากการสร้างแบบจำลองทั้ง 3 โครงสร้าง และทำการทดสอบแบบจำลองด้วยชุดข้อมูลทดสอบใช้ prediction ค่าทำนายเท่ากับ 0 หมายถึง ทำนายว่าเป็นปอดของผู้ติดเชื้อ โควิด 19 เมื่อค่าทำนายเท่ากับ 1 หมายถึง ทำนายว่าเป็นปอดของผู้ป่วยปกติ และเมื่อค่าทำนายเท่ากับ 2 หมายถึง ทำนายว่าเป็นปอดของผู้ติดเชื้อปอดอักเสบจากไวรัส โดยใช้ฟังก์ชัน  $\text{np.argmax}$  มาดึงค่าคำตอบจากการทำนาย ดังรูปที่ 29 จากนั้นทำการสรุปผลเป็น confusion matrix และแสดงผลการทำนายแต่ละคลาสด้วย classification\_report และจะแสดงค่า average macro ของ precision, recall, f1 score และ accuracy ดังรูปที่ 30

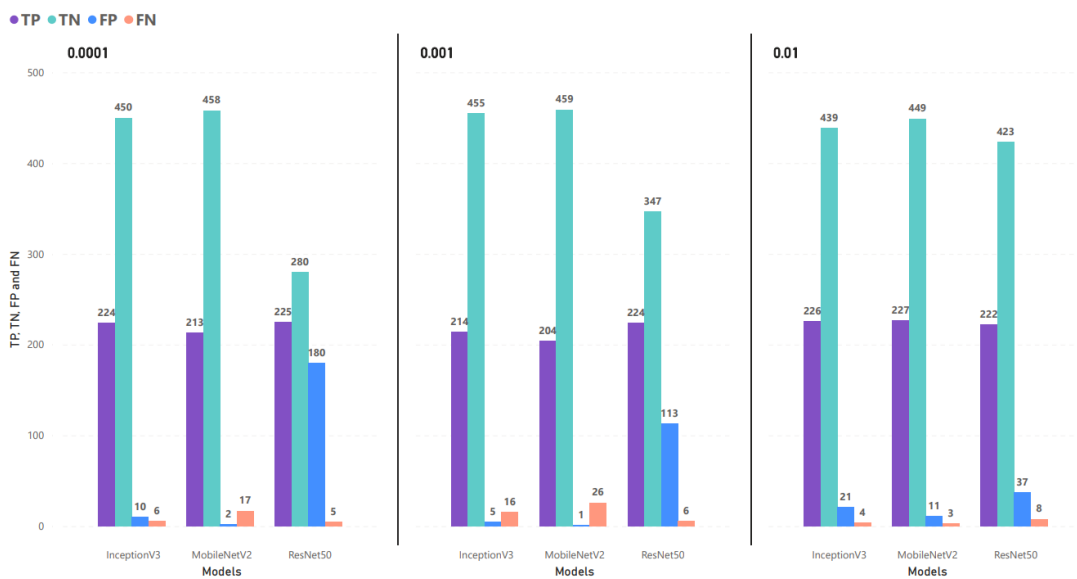


โดยสมรรถนะของแบบจำลองทั้ง 3 โครงสร้าง คือ InceptionV3, Resnet50 และ MobileNetV2 สามารถดูผลการทดสอบแบบจำลองได้ที่ ตารางที่ 4

ตารางที่ 4 ประสิทธิภาพของการทดสอบแบบจำลองการเรียนรู้การถ่ายทอดแบบ InceptionV3, MobileNetV2 และ Resnet50

Models	Learning rate	Confusion matrix				Macro average			Accuracy
		TP	FN	FP	TN	Precision	Recall	F1 score	
InceptionV3	0.01	226	4	21	439	<b>0.94</b>	<b>0.93</b>	<b>0.93</b>	<b>0.93</b>
MobileNetV2		<b>227</b>	<b>3</b>	11	449	0.93	0.92	0.92	0.92
ResNet50		222	8	37	423	0.89	0.88	0.88	0.88
InceptionV3	0.001	214	16	5	455	<b>0.94</b>	<b>0.93</b>	<b>0.93</b>	<b>0.93</b>
MobileNetV2		204	26	<b>1</b>	<b>459</b>	0.92	0.92	0.92	0.92
ResNet50		224	6	113	347	0.79	0.75	0.72	0.75
InceptionV3	0.0001	224	6	10	450	0.93	<b>0.93</b>	<b>0.93</b>	<b>0.93</b>
MobileNetV2		213	17	2	458	0.92	0.92	0.92	0.92
ResNet50		225	5	180	280	0.76	0.66	0.59	0.66

### TP, TN, FP AND FN BY MODELS AND LEARNING RATE



รูปที่ 31 กราฟแท่งแสดงการเปรียบเทียบ TP, TN, FP, และ FN ในแต่ละ learning rate

### PRECISION, RECALL, F1 SCORE AND ACCURACY BY MODELS AND LEARNING RATE



รูปที่ 32 กราฟแท่งแสดงการเปรียบเทียบ Precision, Recall, F1-score, และ Accuracy ในแต่ละ learning rate

จากตารางที่ 4 พบว่า แบบจำลองที่ปรับค่า learning rate เท่ากับ 0.001 ทุกตัวแบบ ให้ค่าความสูญเสียที่ต่ำกว่า เมื่อเทียบกับตัวแบบเดียวกัน ในขณะเดียวกัน โมเดลที่ใช้การเรียนรู้การถ่ายทอดจาก InceptionV3 ให้ค่าค่าความสูญเสียที่ต่ำที่สุด และให้ค่าความแม่นยำสูงที่สุดเมื่อเทียบกับโมเดลอื่นที่ใช้ learning rate ที่ 0.01 และ 0.001 โดยมีเพียงแค่ learning rate ที่ 0.0001 ที่มีค่าค่าความสูญเสียที่สูงกว่า และค่าความแม่นยำที่ต่ำกว่า ตัวแบบที่ใช้ MobileNetV2 เพียงเล็กน้อยเท่านั้น โดยโมเดลที่ใช้การเรียนรู้การถ่ายทอดจาก Resnet50 ให้ค่าค่าความสูญเสียและค่าความแม่นยำต่ำที่สุดเมื่อเทียบกับตัวแบบอื่นที่มี learning rate ที่เท่ากันและมีค่าความแม่นยำลดลงและค่าความสูญเสียเพิ่มขึ้นอย่างต่อเนื่อง เมื่อ learning rate เพิ่มขึ้น ซึ่งจะเห็นได้ว่าเมื่อตัวแบบทำนายข้อมูลจากชุดข้อมูลทดสอบ โมเดลไม่สามารถดึงคุณลักษณะที่สำคัญจากชุดข้อมูลทดสอบได้มากเท่าที่ควร โดยมีค่า false positive ที่สูงกว่าตัวแบบอื่นอย่างเห็นได้ชัดเนื่องจาก คุณลักษณะเด่นของแบบจำลองที่สามารถข้ามแต่ละชั้น และไปประมวลผลในชั้นถัดไป อาจทำให้ไม่สามารถแยกคุณลักษณะของข้อมูลชุดใหม่ได้คืนักเมื่อใช้สำหรับทดสอบสมรรถนะแบบจำลองในการดึงคุณลักษณะรูปภาพชุดใหม่

#### 4.3 การปรับละเอียดสมรรถนะแบบจำลอง

จากผลการทดสอบความสามารถในการทำนายของแบบจำลอง MobileNetV2 และ InceptionV3 มีสมรรถนะที่ใกล้เคียงกันทั้งด้านของค่าความสูญเสียและค่าความแม่นยำในทุก learning rate ที่กำหนด ในขั้นตอนการฝึกแบบจำลอง ไม่ว่าจะอย่างไรก็ตามเมื่อนำแบบจำลองทั้งสองมาใช้กับข้อมูลทดสอบ ที่เป็นรูปภาพที่ไม่เคยเห็นมาก่อน ซึ่งค่าทำนายเท่ากับ 0 หมายถึง ทำนายว่าเป็นปอดของผู้ติดเชื้อโควิด19 เมื่อค่าทำนายเท่ากับ 1 หมายถึง ทำนายว่าเป็นปอดของผู้ป่วยปกติ และเมื่อค่าทำนายเท่ากับ 2 หมายถึง ทำนายว่าเป็นปอดของผู้ติดเชื้อปอดอักเสบจากไวรัส แบบจำลองการเรียนรู้การถ่ายทอดจาก InceptionV3 ให้ค่า macro average ที่ดีกว่าทุกด้าน เช่น precision, recall, f1-score, และ accuracy โดยในแง่ของ learning rate จะเลือกใช้ที่ 0.0001 เนื่องจาก ตัวแบบให้ค่า TP และ TN ที่สูง ในขณะที่ FN และ FP ที่ต่ำ ดังนั้น สำหรับงานวิจัยนี้จึงเลือกการเรียนรู้ถ่ายโอนแบบ InceptionV3 มาเพื่อการพัฒนาศักยภาพของแบบจำลอง โดยนำ pre-trained weight จาก imagenet



มาใช้ และ ทำการพัฒนาประสิทธิภาพของแบบจำลองโดยการตรึงชั้นที่สามารถทำให้การดึงคุณลักษณะที่สำคัญของรูปภาพดีขึ้น

จากการตรวจสอบแบบจำลอง inceptionV3 และแสดงจำนวนชั้นทั้งหมดในแบบจำลองการเรียนรู้แบบถ่ายโอน InceptionV3 ทำให้ทราบว่ามีชั้นจำนวนชั้นทั้งหมด 311 ชั้น ดังรูปที่ 32 ดังนั้นในการปรับสมรรถนะแบบจำลองจะทำการละทิ้งการตรึงจาก top model ไปทีละชั้นโดยเลือกพิจารณาจากการละทิ้ง ตามชั้น ทั้ง 8 แบบ คือ ชั้นที่ 120, 130, 150, 160, 200, 230, 250, และ 280 โดยผลการทดสอบในขั้นตอนการละทิ้งการตรึงตามชั้นที่ถูกระบุทั้งหมดของแบบจำลองการเรียนรู้แบบถ่ายโอนแบบ InceptionV3 สามารถดูสมรรถนะได้จาก ตารางที่ 5



```
# Fine-tune from this layer onwards
fine_tune_at = 150
```

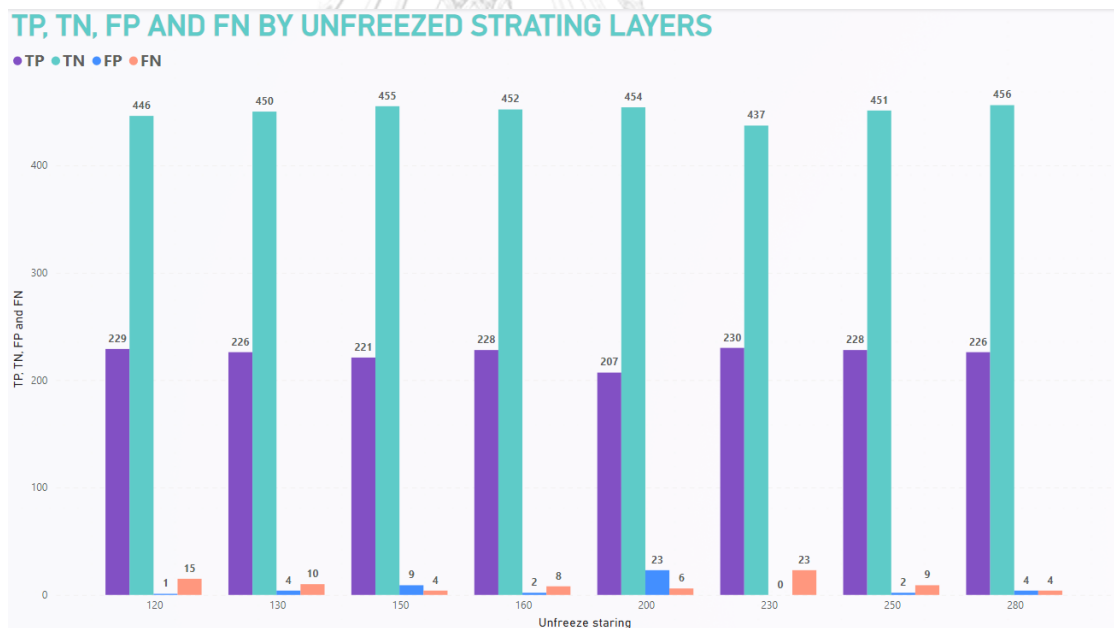
```
# Freeze all the layers before the `fine_tune_at` layer
for layer in base_model.layers[fine_tune_at:]:
    layer.trainable = False
```

รูปที่ 32 ฟังก์ชันในการ unfreeze เลเยอร์ และแสดงจำนวนเลขเอร์  
ของ pre-trained โมเดลทั้งหมด

จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY

ตารางที่ 5 ประสิทธิภาพของการทดสอบแบบจำลองการเรียนรู้การถ่ายทอดแบบ InceptionV3 ในการละทิ้งการตรึงในชั้นที่กำหนด

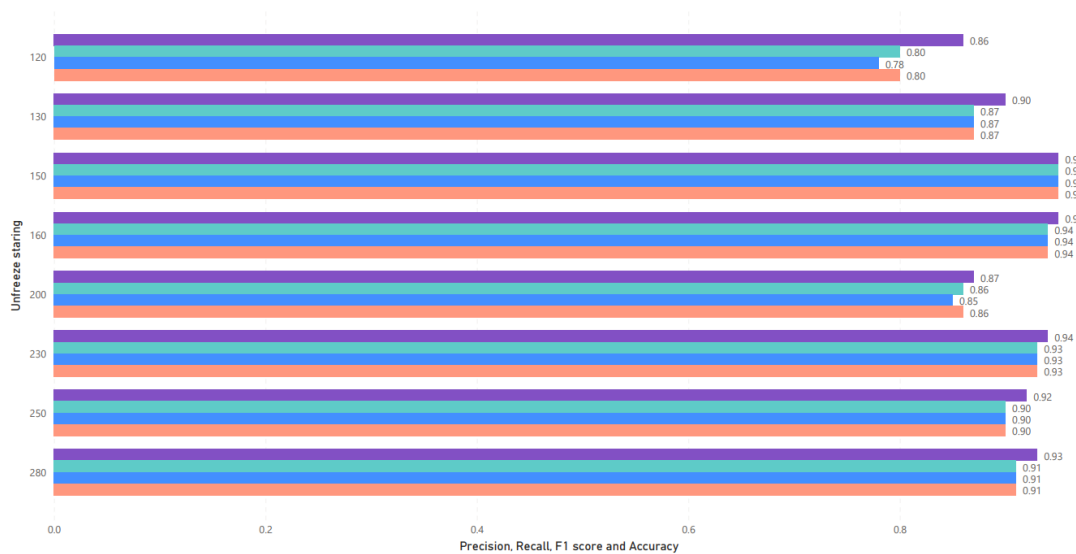
Model	Unfreeze starting	Confusion matrix				Macro average			Accuracy
		TP	FN	FP	TN	Precision	Recall	F1 score	
Inception V3	120	229	15	1	446	0.86	0.80	0.78	0.80
	130	226	10	4	450	0.90	0.87	0.87	0.87
	150	221	4	9	455	<b>0.95</b>	<b>0.95</b>	<b>0.95</b>	<b>0.95</b>
	160	228	8	2	452	0.95	0.94	0.94	0.94
	200	207	6	23	454	0.87	0.86	0.85	0.86
	230	<b>230</b>	23	<b>0</b>	437	0.94	0.93	0.93	0.93
	250	228	9	2	451	0.92	0.90	0.90	0.90
	280	226	4	4	<b>456</b>	0.93	0.91	0.91	0.91



รูปที่ 33 กราฟแท่งแสดงการเปรียบเทียบ TP, TN, FP, และ FN ในแต่ละเลเยอร์ที่เริ่มต้นการละทิ้งการตรึง

### PRECISION, RECALL, F1-SCORE AND ACCURACY BY UNFREEZED STARTING LAYERS

● Precision ● Recall ● F1 score ● Accuracy



รูปที่ 34 กราฟแท่งแสดงการเปรียบเทียบ Precision, Recall, F1-score, และ Accuracy ในแต่ละเลเยอร์ที่เริ่มต้นการละทิ้งการตรึง

## บทที่ 5

### สรุปผลการวิจัยและข้อเสนอแนะ

#### 5.1 สรุปผลการวิจัย

วิทยานิพนธ์นี้ได้นำเสนอประยุกต์ใช้การเรียนรู้ถ่ายทอดร่วมกับโครงข่ายประสาทเทียมแบบคอนโวลูชัน เพื่อจำแนกประเภทภาพถ่ายรังสีทรวงอกของผู้ป่วยที่ติดเชื้อโควิด 19 ผู้ป่วยปอดปกติ และผู้ป่วยปอดอักเสบจากไวรัส โดยชุดข้อมูลที่ใช้เป็นชุดข้อมูลที่ถูกรวบรวมจากออนไลน์ที่นำเชื่อถือ และทำการเตรียมข้อมูลขึ้นมา ในขณะที่จำนวนของข้อมูลภาพเป็นข้อมูลขนาดเล็ก ขนาด  $256 \times 256 \times 3$  จำนวนเพียง 3480 ภาพเท่านั้น โดยแบ่งเป็นชุดข้อมูลสอน 2100 ภาพ ชุดข้อมูลทดสอบ 690 ภาพ และชุดข้อมูลตรวจสอบ 690 ภาพ ดังนั้นการเปรียบเทียบการเรียนรู้ถ่ายทอดของโครงข่ายประสาทเทียมแบบคอนโวลูชันที่เหมาะสมกับรูปภาพที่จะนำมาใช้ให้แบบจำลองสามารถดึงคุณลักษณะที่สำคัญจากชุดข้อมูลชุดใหม่ที่มีขนาดเล็กจึงเป็นขั้นตอนที่จำเป็นในการคัดเลือกแบบจำลอง

สำหรับการเปรียบเทียบสมรรถนะของแบบจำลอง ผู้วิจัยได้ทำการสร้างแบบจำลองเรียนรู้การถ่ายทอดประกอบไปด้วยแบบจำลอง 3 แบบด้วยกัน คือโครงข่ายประสาทคอนโวลูชันแบบโครงข่ายประสาทเชิงลึกแบบ ResNet50, โครงข่ายประสาทแบบ InceptionV3 และโครงข่ายประสาทแบบ MobileNetV2 โดยแบบจำลองทั้ง 3 โครงสร้าง มีการคำนวณ loss แบบ categorical cross entropy และมีการปรับเปลี่ยนค่า learning rate เป็น 0.01, 0.001 และ 0.0001 ตามลำดับ โดยในการวัดสมรรถนะในการฝึกสอนและตรวจสอบแบบจำลองจะใช้ค่าความแม่นยำเป็นตัววัด โดยใช้ Adaptive moment estimation optimizer เป็น optimizer แล้วทำการ compile แบบจำลอง หลังจากการ compile แบบจำลองสำเร็จจะทำการ fit แบบจำลอง โดยกำหนด batch\_size = 32 และทำการฝึกสอนจำนวน 30 epoches ในการฝึกสอนแต่ละรอบใช้ค่า step\_per\_epoch จากการนำจำนวนข้อมูลฝึกสอนทั้งหมดหารด้วย batch\_size จากผลการทดสอบแบบจำลอง ทั้ง 3 แบบ พบว่าที่ learning rate เท่ากับ 0.0001 แบบจำลอง MobileNetV2 และ InceptionV3 ให้ค่าความแม่นยำในการจำแนกประเภทภาพถ่ายรังสีทรวงอกที่ไม่เคยเห็นมาก่อน ที่ใกล้เคียงกันที่ 93% และ 92% ตามลำดับ ในขณะที่ learning rate เท่ากับ 0.01 และ 0.0001 แบบจำลอง InceptionV3 และ MobileNetV2 ให้ค่าความแม่นยำที่เท่าเดิม สำหรับแบบจำลอง ResNet50 มีค่าความแม่นยำน้อยที่สุดในทุกการปรับ learning rate ซึ่ง มีค่าความแม่นยำที่มากที่สุดที่

88% เมื่อปรับ learning rate เป็น 0.01 และต่ำที่สุดที่ 66% เมื่อปรับ learning rate เป็น 0.0001 ทำให้สรุปได้ว่าแบบจำลองการเรียนรู้ถ่ายทอดแบบ InceptionV3 สามารถใช้จำแนกประเภทรูปภาพถ่ายรังสีทรวงอกได้ดีกว่าแบบจำลองอื่นที่กำหนด

ในขั้นตอนต่อมาได้นำแบบจำลองการเรียนรู้ถ่ายทอดแบบ InceptionV3 มาทำการปรับปรุงสมรรถนะของแบบจำลอง โดยได้ทำการละทิ้งการตรึงใน pre-trained model ทำให้แบบจำลองสามารถเรียนรู้ข้อมูลรูปภาพชุดใหม่ได้ดียิ่งขึ้น โดยใช้ค่าถ่วงน้ำหนักจากชุดข้อมูล imagenet จากนั้นจึงทำการละทิ้งการตรึงในแต่เริ่มต้นทั้งหมด 8 ชั้น คือ 280, 250, 230, 200, 160, 150, 130 และ 120 ตามลำดับ ผลปรากฏว่า แบบจำลองการเรียนรู้ถ่ายโอน InceptionV3 ที่ learning rate เท่ากับ 0.0001 และทำการละทิ้งการตรึง ตั้งแต่ชั้น 120 ให้ผลการทดสอบการจำแนกภาพถ่ายรังสีทรวงอกของผู้ป่วยที่ติดเชื้อโควิด19 ได้ความแม่นยำที่น้อยที่สุด ในขณะที่แบบจำลองที่ทำการละทิ้งการตรึงตั้งแต่ชั้น 150 ให้ผลการทดสอบการจำแนกภาพถ่ายรังสีทรวงอกของผู้ป่วยที่ติดเชื้อโควิด19 ได้ความแม่นยำที่ดีที่สุด เท่ากับ 95% ด้วยค่า precision, recall, f1 score และ accuracy มีค่าเป็น 95%, 95%, 95% และ 95% ตามลำดับ

จากผลการทดสอบทั้งสิ้น 2 การทดสอบในงานวิจัยการเรียนรู้ถ่ายโอนสำหรับการจำแนกประเภทภาพถ่ายรังสีทรวงอกออกเป็น 3 ประเภท คือ 1) ภาพถ่ายรังสีทรวงอกของผู้ป่วยปกติ 2) ภาพถ่ายรังสีทรวงอกของผู้ป่วยที่ติดเชื้อโควิด19 3) ภาพถ่ายรังสีทรวงอกของผู้ติดเชื้อปอดอักเสบจากไวรัส ซึ่งจะเห็นได้ว่า 1) แบบจำลอง InceptionV3 มีประสิทธิภาพและสามารถจำแนกภาพถ่ายรังสีทรวงอกได้ดีกว่าแบบจำลองอื่นที่ทำการทดสอบ 2) แบบจำลองการเรียนรู้การถ่ายทอดสามารถทำนายข้อมูลภาพที่ไม่เคยเห็นมาก่อนได้อย่างมีประสิทธิภาพแม้ข้อมูลจะมีขนาดที่จำกัด 3) การปรับละเอียดสมรรถนะแบบจำลองทั้ง แบบการละทิ้งการตรึงในชั้นต่าง ๆ ส่งผลต่อสมรรถนะของแบบจำลอง

## 5.2 ข้อเสนอแนะ

จากการใช้การใช้โครงข่ายคอนโวลูชันในการสร้างโมเดลจะเห็นได้ว่าข้อมูลที่ใช้สำหรับการเรียนรู้ของโมเดลนั้นมีอยู่ค่อนข้างน้อย ดังนั้นควรมีการเพิ่มจำนวนข้อมูลของภาพถ่ายรังสีทรวงอกให้มากขึ้นในอนาคต นอกจากนั้นจากการทบทวนวรรณกรรมทำให้ทราบว่าสามารถนำเทคนิคต่าง ๆ เช่น PCA ที่สามารถที่จะลดทอนข้อมูลที่มีความสำคัญน้อยลงออก หรือ stratified k-fold cross

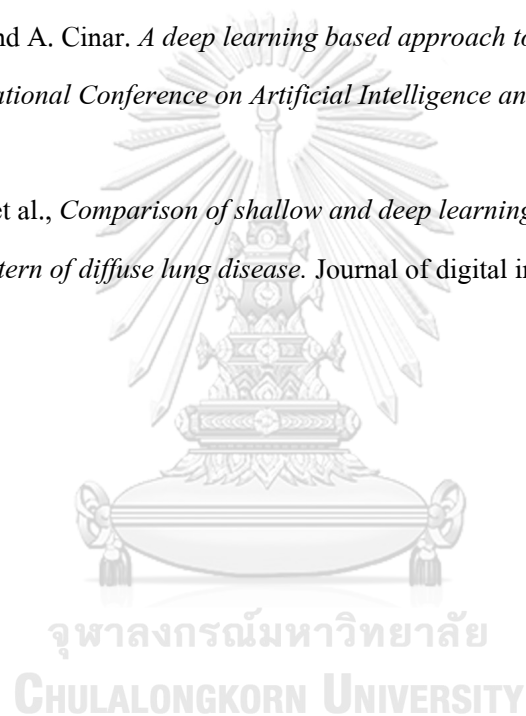
validation ที่ช่วยในการแบ่งคลาสของข้อมูลให้เท่ากันสำหรับการฝึกฝนโมเดล ยิ่งไปกว่านั้นการนำโครงข่ายชนิดอื่น ๆ มาใช้ร่วมกันอาจจะทำให้โมเดลสามารถเพิ่มขีดจำกัดในการเรียนรู้ได้มากขึ้น ตัวอย่างเช่น การนำเอา LSTM มาใช้ร่วมกัน เพราะแต่ละโครงข่ายมีความสามารถในการสกัดคุณลักษณะที่สำคัญของข้อมูลต่างกันไป



## บรรณานุกรม

1. Shereen, M.A., et al., *COVID-19 infection: Emergence, transmission, and characteristics of human coronaviruses*. Journal of advanced research, 2020. **24**: p. 91-98.
2. Diseases, N.C.f.I.a.R. *Post-COVID Conditions*. [cited 2021 29 October].
3. แก้วอมตวงศ์, ร.น.ช. ปอดอักเสบ กับ โควิด-19 เกี่ยวข้องกันอย่างไร. [cited 2021 29 October].
4. Dickson, B. *What are convolutional neural networks (CNN)*. [cited 2021 29 October].
5. Rasheed, J., et al., *A machine learning-based framework for diagnosis of COVID-19 from chest X-ray images*. Interdisciplinary Sciences: Computational Life Sciences, 2021. **13**(1): p. 103-117.
6. Lakhani, P. and B. Sundaram, *Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks*. Radiology, 2017. **284**(2): p. 574-582.
7. Anurag Bhardwaj, J.W., and Wei Di, *Deep Learning Essentials: Your Hands-on Guide to the Fundamentals of Deep Learning and Neural Network Modeling*. 2018: Packt Publishing, 271.
8. R, A., *Deep learning for computer vision with python*. 1 ed. 2017: PYIMAGESEARCH.
9. Geron, A., *Hands-On Machine Learning with Scikit-Learn and TensorFlow*. 2 ed. 2019: O'Reilly Media, Inc.
10. He, K., et al. *Deep residual learning for image recognition*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
11. Szegedy, C., et al. *Rethinking the inception architecture for computer vision*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
12. Sandler, M., et al. *Mobilenetv2: Inverted residuals and linear bottlenecks*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.
13. Fchollet. *Transfer learning & fine-tuning* 2020.
14. Singh, G.A.P. and P. Gupta, *Performance analysis of various machine learning-based approaches for detection and classification of lung cancer in humans*. Neural Computing and Applications, 2019. **31**(10): p. 6863-6877.
15. Ucar, F. and D. Korkmaz, *COVIDiagnosis-Net: Deep Bayes-SqueezeNet based diagnosis of*

- the coronavirus disease 2019 (COVID-19) from X-ray images*. Medical hypotheses, 2020. **140**: p. 109761.
16. Dey, R., Z. Lu, and Y. Hong. *Diagnostic classification of lung nodules using 3D neural networks*. in *2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)*. 2018. IEEE.
  17. Mishra, N.K., P. Singh, and S.D. Joshi, *Automated detection of COVID-19 from CT scan using convolutional neural network*. Biocybernetics and Biomedical Engineering, 2021. **41**(2): p. 572-588.
  18. Cengil, E. and A. Cinar. *A deep learning based approach to lung cancer identification*. in *2018 International Conference on Artificial Intelligence and Data Processing (IDAP)*. 2018. IEEE.
  19. Kim, G.B., et al., *Comparison of shallow and deep learning methods on classifying the regional pattern of diffuse lung disease*. Journal of digital imaging, 2018. **31**(4): p. 415-424.







จุฬาลงกรณ์มหาวิทยาลัย  
**CHULALONGKORN UNIVERSITY**

## ประวัติผู้เขียน

ชื่อ-สกุล

ชญ์ชวิน โพธิวัฒน์ธนต์

วัน เดือน ปี เกิด

6 สิงหาคม 2539

วุฒิการศึกษา

สถิติศาสตรบัณฑิต

ผลงานตีพิมพ์

IAMBEST2022



จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY