

การสร้างภาพสามมิติของหลอดเลือดแดงใหญ่ส่วนท้องด้วยโครงข่าย  
การเรียนรู้เชิงลึกของสถาปัตยกรรมโครงข่ายยูเน็ตชนิดสามมิติ  
จากภาพเอกซเรย์คอมพิวเตอร์หลอดเลือด

RECONSTRUCTION OF 3D ABDOMINAL AORTA ANEURYSM FROM  
COMPUTED TOMOGRAPHIC ANGIOGRAPHY USING 3D U-NET  
DEEP LEARNING NETWORK



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต  
สาขาวิชาวิศวกรรมชีวการแพทย์  
คณะวิศวกรรมศาสตร์  
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง  
พ.ศ.2566

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

RECONSTRUCTION OF 3D ABDOMINAL AORTA ANEURYSM FROM  
COMPUTED TOMOGRAPHIC ANGIOGRAPHY USING 3D U-NET  
DEEP LEARNING NETWORK



A THESIS SUBMITTED IN PARTIAL FULFILLMENT  
OF THE REQUIREMENT FOR THE DEGREE OF  
MASTER OF ENGINEERING IN BIOMEDICAL ENGINEERING  
SCHOOL OF ENGINEERING  
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG  
2023  
KMITL-2023-EN-M-317-035

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



COPYRIGHT 2023

SCHOOL OF ENGINEERING

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์	การสร้างภาพสามมิติของหลอดเลือดแดงใหญ่ส่วนท้องด้วยโครงข่ายการเรียนรู้เชิงลึกของสถาปัตยกรรมโครงข่ายยูเน็ตชนิดสามมิติ จากภาพเอกซเรย์คอมพิวเตอร์หลอดเลือด
นักศึกษา	นางสาวศิริพร คงรัตน์
รหัสประจำตัว	63601060
ปริญญา	วิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชา	วิศวกรรมชีวการแพทย์
ปีการศึกษา	2566
อาจารย์ที่ปรึกษาวิทยานิพนธ์	รศ.ดร.สุพันธุ์ ตั้งจิตกุศลมั่น
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม	รศ.ดร.ชูชาติ ปิณฑวิรุจน์

### บทคัดย่อ

หลอดเลือดแดงใหญ่ในช่องท้องโป่งพอง (AAA) คือการโป่งของหลอดเลือดแดงใหญ่ที่เกิดขึ้นเมื่อผนังของหลอดเลือดแดงใหญ่อ่อนแอลง AAA เป็นภาวะที่อาจเป็นอันตรายถึงชีวิต โดยเฉพาะอย่างยิ่งหากมีการแตกเป็นผลทำให้เกิดเลือดออกรุนแรง วิธีดำเนินการวิจัยเราพัฒนาวิธีการแบ่งส่วนภาพแบบอัตโนมัติสำหรับการสร้างรูปทรงสามมิติของหลอดเลือดใหญ่ส่วนท้องที่โป่งพอง จากการตรวจทางรังสีด้วยเอกซเรย์คอมพิวเตอร์หลอดเลือดแดง (CTA) ตามแนวทางโครงข่ายการเรียนรู้เชิงลึกด้วยสถาปัตยกรรมยูเน็ตสามมิติสำหรับ AAA และ AAA ที่มีลิ้มเลือดในผนังที่หนาบนชุดข้อมูลการฝึกการเรียนรู้ที่ โดยจำแนกประเภทของหลอดเลือดออก 8 ชุดข้อมูลในหลอดเลือดปกติ 14 ชุดข้อมูลในหลอดเลือดที่มีการโป่งพอง และ 38 ชุดข้อมูลในหลอดเลือดที่มีโป่งพองและมีลิ้มเลือดในผนังที่หนา และขยายข้อมูลด้วยฟังก์ชันเสริมข้อมูล เช่น การลดขนาด การตัดภาพแบบสุ่ม การเปลี่ยนระดับความเทา การกลับแกนภาพ และการเฉือนภาพบนล่าง ถูกเพิ่มลงในโมเดลการฝึกการเรียนรู้เชิงลึกนี้ เพื่อให้ได้ประสิทธิภาพที่ดีที่สุดและมีความถูกต้อง ผลการทดลองที่ได้ช่วยยืนยันว่าวิธีการที่นำเสนอนี้ สามารถให้ความถูกต้องในแง่ของค่าสัมประสิทธิ์ความคล้าย Dice Coefficient Score (DSC) ที่ 0.9669 จากการประเมินประสิทธิภาพการฝึกการเรียนรู้ของโมเดล และ 0.9868 จากการประเมินการทดสอบโมเดลด้วยสถาปัตยกรรมยูเน็ตสามมิติ

<b>Thesis</b>	Reconstruction of 3D Abdominal Aorta Aneurysm from Computed Tomographic Angiography Using 3D U-Net Deep Learning Network
<b>Student</b>	Miss.Siriporn Kongrat
<b>Student ID.</b>	63601060
<b>Degree</b>	Master of Engineering
<b>Program</b>	Biomedical Engineering
<b>Year</b>	2023
<b>Thesis Advisor</b>	Assoc.Prof.Dr. Supan Tungjitkusolmun
<b>Thesis Co-Advisor</b>	Assoc.Prof.Dr. Chuchart Pintavirooj

### ABSTRACT

(1) Background: An abdominal aortic aneurysm (AAA) is a swelling (aneurysm) of the aorta that occurs when the wall of the aorta weakens. An AAA is a potentially life-threatening condition, especially if it eventually ruptures, causing severe bleeding. (2) Methods: We developed an automated segmentation method for 3D AAA reconstruction from computed tomography angiography (CTA) based on the 3D U-NET deep learning network approaches for AAA and AAA with thrombus on training dataset classified as 8 normal, 14 aneurysm volume, and 38 thrombus aneurysm volume with the data augmentations app, i.e., scaling, random crop, gray-scale variation, axial y flip, and shear, were added to the training model, achieving better performance. (3) Results: The results confirm that the proposed method can provide accuracy in terms of the Dice Similar Coefficient (DSC) scores of 0.9669 for training performance and 0.9868 for testing evaluation with the 3D U-Net model.

## กิตติกรรมประกาศ

วิทยานิพนธ์เล่มนี้สำเร็จได้ด้วยความกรุณาจากอาจารย์ที่ปรึกษา รศ.ดร.สุพันธุ์ ตั้งจิตกุศลมั่น และอาจารย์ที่ปรึกษาร่วม รศ.ดร.ชูชาติ ปิณฑวิรุจน์ ที่มีความกรุณาให้คำปรึกษา คำแนะนำ รวมทั้งความรู้ต่างๆ เพื่อใช้ในการดำเนินงานวิจัยให้สำเร็จ และสมบูรณ์

ขอขอบคุณหลักสูตรที่ทำให้ข้าพเจ้าได้เพิ่มพูนความรู้เพื่อใช้ในการดำเนินการวิจัยในครั้งนี้ ได้เรียนรู้ในสิ่งใหม่ๆ และเพิ่มความสามารถในการประกอบอาชีพได้อย่างมีประสิทธิภาพ

ขอขอบคุณแหล่งความรู้ช่องทางออนไลน์ ที่ทำให้ข้าพเจ้าหาข้อมูลเพื่อนำมาใช้เป็นความรู้ในการทำวิทยานิพนธ์เล่มนี้ได้สำเร็จ และงานวิจัยต่าง ๆ ที่ได้รวบรวมความรู้มากมายและให้ทักษะมาประยุกต์ใช้ในงานนี้

ขอขอบคุณ โรงพยาบาลรามาริบัติ มหาวิทยาลัยมหิดล ที่ให้การสนับสนุนการวิจัยด้านข้อมูล ซึ่งเป็นแหล่งการเรียนรู้ และมีฐานข้อมูลขนาดใหญ่

ขอขอบคุณอาจารย์นายแพทย์ศิโรช วงศ์ไวยวรรณ และนางสาวเมธาวี เกษขุนทด ที่ช่วยเหลือเรื่องการขอจริยธรรมในมนุษย์ของโรงพยาบาลรามาริบัติ เพื่อใช้ในการดำเนินงานวิจัยให้สำเร็จ และสมบูรณ์

สุดท้ายต้องขอขอบคุณครอบครัว และเพื่อนของข้าพเจ้า ที่เป็นกำลังใจ และสนับสนุนการศึกษาในครั้งนี้ตลอดมา

สิ่งที่ตั้งงามและคุณค่าของวิทยานิพนธ์ฉบับนี้ข้าพเจ้าขอมอบและอุทิศแด่ผู้มีพระคุณทุกท่าน

ศิริพร คงรัตน์

## สารบัญ

	หน้า
บทคัดย่อภาษาไทย .....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VII
สารบัญรูป.....	VIII
<b>บทที่ 1 บทนำ.....</b>	<b>1</b>
1.1 ความเป็นมาของปัญหา.....	1
1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา.....	2
1.3 สมมุติฐานของการศึกษา.....	2
1.3 ขอบเขตของงานวิจัย.....	2
1.4 แผนการดำเนินงานวิจัย.....	3
<b>บทที่ 2 ทฤษฎีที่เกี่ยวข้อง.....</b>	<b>4</b>
2.1 โรคหลอดเลือดแดงใหญ่ในช่องท้องโป่งพอง.....	4
2.2 โครงข่ายประสาทเทียมเบื้องต้น.....	6
2.2.1 องค์ประกอบของโครงข่ายประสาทเทียม.....	6
2.3 การเรียนรู้ของโครงข่ายประสาทเทียม (learning of neural network).....	10
2.2.1 องค์ประกอบของโครงข่ายประสาทเทียม.....	11
2.3.2 การเรียนรู้แบบไม่มีผู้สอน (unsupervised learning) .....	11
2.3.3 การเรียนรู้เชิงบังคับ (reinforcement learning) .....	12
2.4 การเรียนรู้เชิงลึก (Deep learning) .....	13
2.5 โครงข่ายประสาทคอนโวลูชัน (Convolutional Neural Networks : CNN).....	13
2.5.1 Convolutional Layer.....	14
2.5.2 Pooling layer.....	16
2.5.3 Padding.....	17
2.5.4 Stride.....	17
2.5.5 Fully Connected Layer.....	18
2.6 โครงข่ายประสาทคอนโวลูชัน (Convolutional Neural Networks : CNN).....	19
2.7 โครงข่ายประสาทคอนโวลูชันแบบ 3 มิติ (3D CNN) .....	20
2.8 การแบ่งส่วนภาพ (Image Segmentation).....	20

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.9	โครงข่ายประสาทเทียมสำหรับการแบ่งส่วน.....	21
2.10	ตัวเข้ารหัส - ตัวถอดรหัส (encoder – decoder).....	21
2.11	การเชื่อมต่อตรงข้าม (Skip connections) .....	22
2.12	การถ่ายทอดการเรียนรู้ (Transfer learning) .....	22
2.13	สถาปัตยกรรมการเรียนรู้เชิงลึกสำหรับการแบ่งส่วนภาพ.....	23
2.13.1	Fully Convolutional networks .....	23
2.13.2	คอนโวลูชันโมเดลกับกราฟฟิกโมเดล .....	27
2.13.3	โมเดลที่ใช้ตัวเข้ารหัส - ตัวถอดรหัส (encoder – decoder models)....	28
2.13.4	โมเดลตามโครงข่ายแบบหลายสเกลและพีระมิด .....	28
2.13.5	โมเดลอาร์ซีเอ็นเอ็น (R-CNN) .....	29
2.13.6	โมเดลคอนโวลูชันแบบขยายและตระกูล DeepLab .....	31
2.13.7	โมเดลตามโครงข่ายประสาทเทียมที่เก็ดย้ำ (RNN).....	31
2.13.9	โมเดลต้นกำเนิดและการฝึกอบรมด้านตรงข้าม.....	32
2.13.10	โมเดลโครงข่ายประสาทเทียมพร้อมโมเดลแอคทีฟคอนทัว.....	32
2.13.11	โมเดลอื่นๆ.....	32
2.14	วิธีการทางการแพทย์สำหรับการแบ่งส่วนภาพทางการแพทย์.....	33
2.14.1	Computed Tomography Scan (CT scan) .....	33
2.14.2	Magnetic resonance imaging (MRI) .....	35
2.14.3	เอกซเรย์ระบบดิจิตอล (DR) .....	36
2.15	รูปภาพและส่วนประกอบของภาพทางการแพทย์.....	37
2.15.1	ความลึกของพิกเซลหรือความลึกบิตหรือความลึกของสี.....	38
2.15.2	การระบุความเข้มแสง.....	39
2.15.3	ข้อมูลเมตา.....	39
2.15.4	ข้อมูลพิกเซล.....	40
2.16	รูปแบบไฟล์ภาพทางการแพทย์.....	40
2.16.1	DICOM (การถ่ายภาพดิจิทัลและการสื่อสารในการแพทย์).....	40
2.16.2	NIFTI (Neuroimaging Informatics Technology Initiative).....	41
2.16.3	NRRD.....	41
2.16.4	MINC.....	41
2.17	เมตริกและการประเมินชุดข้อมูลการแบ่งกลุ่ม.....	41
2.17.1	ค่าสัมประสิทธิ์ความเหมือน (Dice Coefficient Score : DSC).....	41
2.17.2	ค่าสัมประสิทธิ์ความคล้าย (Intersection Over Union : IoU).....	42

2.17.3 ค่าความแม่นยำของฟิสิกเซล (Accuracy) .....	43
2.17.4 ฟังก์ชันการสูญเสีย (Loss functions) .....	43
2.17.5 Binary Cross-Entropy.....	44
2.17.6 Dice Loss.....	44
2.18 การขยายข้อมูล (Data Augmentation) .....	44
2.19 Google Colab GPU.....	45
2.20 3D slicer.....	46
2.21 MATLAB.....	47
2.23.1 Volume Viewer.....	47
<b>บทที่ 3 วิธีดำเนินงานวิจัย.....</b>	<b>49</b>
3.1 คำอธิบายชุดข้อมูล.....	49
3.2 การเตรียมชุดข้อมูล.....	49
3.3 การเตรียมชุดข้อมูล CT Images.....	51
3.4 การเตรียมชุดข้อมูลแบบตัวอย่างการฝึกรวม Ground truth annotations.....	51
3.5 วิธีดำเนินงานวิจัย.....	53
3.5.1 ความจำเป็นของคอมพิวเตอร์.....	53
3.5.2 Importing files and data preprocessing.....	53
3.6 การสร้างแพตช์สำหรับเก็บชุดข้อมูลจากการทำฟังก์ชันเสริมข้อมูล (Patchwise Augmentation) .....	56
3.7 การโหลดไฟล์จากแพตช์ data augmentation.....	60
3.8 การแบ่งข้อมูลสำหรับใช้ในการฝึกการเรียนรู้ และทดสอบโมเดล.....	61
3.9 การนำข้อมูลเข้าโครงข่ายประสาทเทียมเพื่อทำการฝึกการเรียนรู้.....	63
3.10 สถาปัตยกรรมโครงข่ายประสาทเทียมยูเน็ตสามมิติ (3D U-Net architecture)..	63
3.11 เริ่มฝึกชุดข้อมูลที่เตรียมไว้สำหรับฝึกการเรียนรู้.....	66
3.12 การทำนายด้วยสถาปัตยกรรมยูเน็ต U-Net.....	68
<b>บทที่ 4 ผลการทดลอง.....</b>	<b>71</b>
4.1 ผลการทดลองการใช้สถาปัตยกรรม 3D U-Net.....	71
<b>บทที่ 5 สรุปและอภิปรายผลการทดลอง.....</b>	<b>76</b>
5.1 สรุปผลการทดลอง.....	76
5.2 ข้อจำกัดในการศึกษา.....	76
5.3 บทสรุปข้อเสนอแนะ.....	77
บรรณานุกรม.....	78

## สารบัญตาราง

	หน้า
ตารางที่ 1.1 แผนการดำเนินงานวิจัย.....	3
ตารางที่ 4.1 ผลการทดลองจากสถาปัตยกรรมยูเน็ตสามมิติ.....	71



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญรูป

	หน้า
รูปที่ 2.1 ตำแหน่งหลอดเลือดแดงใหญ่ส่วนท้อง ประเภทปกติและโป่งพอง.....	4
รูปที่ 2.2 การรักษาหลอดเลือดแดงใหญ่โป่งพอง.....	5
รูปที่ 2.3 โครงสร้างระบบประสาท.....	7
รูปที่ 2.4 ส่วนประกอบของโครงข่ายประสาทเทียม.....	7
รูปที่ 2.5 โครงสร้างการทำงานของโครงข่ายประสาทเทียม.....	8
รูปที่ 2.6 ฟังก์ชันลิคเก้เรคตีไฟด์ลินเนียยูนิต.....	10
รูปที่ 2.7 โครงสร้างการเรียนรู้แบบมีผู้สอน.....	11
รูปที่ 2.8 กระบวนการการเรียนรู้แบบมีผู้สอน.....	11
รูปที่ 2.9 โครงสร้างการเรียนรู้แบบไม่มีผู้สอน.....	12
รูปที่ 2.10 กระบวนการเรียนรู้แบบไม่มีผู้สอน.....	12
รูปที่ 2.11 กระบวนการเรียนรู้เชิงบังคับ.....	13
รูปที่ 2.12 วิธีการคำนวณคอนโวลูชัน.....	15
รูปที่ 2.13 ขั้นตอนการคำนวณคอนโวลูชัน.....	15
รูปที่ 2.14 การเขียนรูปแบบ Neural network.....	16
รูปที่ 2.15 วิธีการคำนวณ Pooling Layer.....	17
รูปที่ 2.16 พารามิเตอร์ที่ต้องกำหนดใน Convolutional layer.....	18
รูปที่ 2.17 Fully connected layer.....	18
รูปที่ 2.18 โครงข่ายประสาทคอนโวลูชันแบบ 2 มิติ (2D CNN) .....	19
รูปที่ 2.19 โครงข่ายประสาทคอนโวลูชันแบบ 3 มิติ (3D CNN) .....	20
รูปที่ 2.20 สถาปัตยกรรมตัวเข้ารหัส-ตัวถอดรหัส.....	22
รูปที่ 2.21 การเชื่อมต่อตรงข้าม (Skip connections) .....	22
รูปที่ 2.22 การถ่ายทอดการเรียนรู้ (Transfer learning) .....	23
รูปที่ 2.23 Fully Convolutional Network (FCN) .....	24
รูปที่ 2.24 คอนโวลูชันโมเดลกับกราฟฟิกโมเดล.....	25
รูปที่ 2.25 โมเดลที่ใช้ตัวเข้ารหัส - ตัวถอดรหัส (encoder – decoder models) .....	25
รูปที่ 2.26 โครงสร้างสถาปัตยกรรมยูเน็ต (U-Net) .....	26
รูปที่ 2.27 โครงสร้างสถาปัตยกรรมวีเน็ต (V-Net) .....	27
รูปที่ 2.28 โมเดลตามโครงข่ายแบบหลายสเกลและพีระมิด.....	27
รูปที่ 2.29 โมเดลอาร์ซีเอ็นเอ็น (R-CNN).....	28

รูปที่ 2.30 โมเดลคอนโวลูชันแบบขยายและตระกูล DeepLab.....	29
รูปที่ 2.31 โมเดลตามโครงข่ายประสาทเทียมที่เกิดซ้ำ (RNN) .....	29
รูปที่ 2.32 การเปรียบเทียบระหว่างโมเดลกราฟ-LSTM กับ โมเดล RNN แบบใช้ พิกเซลแบบดั้งเดิม.....	30
รูปที่ 2.33 แบบจำลองกราฟ LSTM สำหรับการแบ่งส่วน.....	30
รูปที่ 2.34 โมเดลที่เน้นความสนใจ (Attention-Based Models) .....	31
รูปที่ 2.35 โมเดลต้นกำเนิดและการฝึกอบรมด้านตรงข้าม.....	32
รูปที่ 2.36 ช่วงเวลาของอัลกอริทึมการแบ่งส่วนภาพ 2 มิติ.....	33
รูปที่ 2.37 เครื่องเอกซเรย์คอมพิวเตอร์ (CT Scan) .....	33
รูปที่ 2.38 เครื่องสร้างภาพด้วยสนามแม่เหล็กไฟฟ้า .....	35
รูปที่ 2.39 แผนผังแสดงส่วนประกอบภาพทางการแพทย์.....	38
รูปที่ 2.40 ตารางแสดงค่าจำนวนบิตที่ใช้เข้ารหัสข้อมูลของแต่ละพิกเซล.....	39
รูปที่ 2.41 แผนผังแสดงรูปแบบไฟล์ภาพทางการแพทย์.....	40
รูปที่ 2.42 แสดงถึงค่าสัมประสิทธิ์ความเหมือน (Dice Coefficient Score : DSC).....	42
รูปที่ 2.43 แสดงถึงค่าสัมประสิทธิ์ความคล้าย (Intersection Over Union : IoU).....	43
รูปที่ 2.44 แสดงถึงชนิดการขยายข้อมูล (Data Augmentation).....	45
รูปที่ 2.45 แสดงการแบ่งส่วนภาพด้วยโปรแกรม 3D slicer.....	47
รูปที่ 2.46 แอปพลิเคชัน Volume Viewer จากโปรแกรม Matlab.....	48
รูปที่ 3.1 แสดงภาพเอกซเรย์คอมพิวเตอร์แบบโวลูม.....	50
รูปที่ 3.2 แสดงการใช้ 3D Slicer.....	51
รูปที่ 3.3 แสดงภาพเอกซเรย์คอมพิวเตอร์หลอดเลือด CT image และ ชุดข้อมูลแบบตัวอย่างการฝึกอบรม (Ground truth) .....	53
รูปที่ 3.4 อัลกอริทึมการนำเข้าไลบรารี (library).....	54
รูปที่ 3.5 อัลกอริทึมการนำเข้าชุดข้อมูลและการแปลงภาพเป็นอาร์เรย์.....	55
รูปที่ 3.6 แสดงการนำเข้าชุดข้อมูลบางส่วน.....	56
รูปที่ 3.7 อัลกอริทึมการสร้างแพตช์เพื่อเก็บชุดข้อมูล.....	57
รูปที่ 3.8 ชุดคำสั่งการสร้างข้อมูลรูปแบบต่างๆ.....	58
รูปที่ 3.9 ชุดคำสั่งการลดขนาดภาพ (resize) .....	58
รูปที่ 3.10 ชุดคำสั่งการเปลี่ยนการเทา (Grayscale variation) .....	59
รูปที่ 3.11 ชุดคำสั่งการกลับด้านภาพ แนวนว x , y (Flip) .....	59
รูปที่ 3.12 ชุดคำสั่งการเลื่อนภาพบนล่าง (Shear) .....	60
รูปที่ 3.13 ไฟล์แพตช์ที่ใช้เก็บข้อมูลภาพการจากใช้ฟังก์ชันการเสริมข้อมูล.....	61
รูปที่ 3.14 อัลกอริทึมเรียกภาพจากไฟล์แพตช์.....	61

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 3.15 แสดงการแบ่งสัดส่วนข้อมูลสำหรับฝึกการเรียนรู้ และทดสอบโมเดล.....	62
รูปที่ 3.16 อัลกอริทึมใช้สำหรับการแบ่งข้อมูลสำหรับใช้ในการฝึกการเรียนรู้ และทดสอบโมเดล.....	62
รูปที่ 3.17 แผนผังสถาปัตยกรรมยูเน็ตสามมิติ.....	64
รูปที่ 3.18 โลกรารีสำหรับสร้างชุดคำสั่งโครงข่ายประสาทเทียม.....	65
รูปที่ 3.19 คำสั่งการอินพุตขนาดภาพ 3 ระนาบ.....	65
รูปที่ 3.20 ชุดคำสั่งการสร้างสถาปัตยกรรมยูเน็ตสามมิติ.....	66
รูปที่ 3.21 อัลกอริทึมค่าสัมประสิทธิ์ Dice coefficient และ ฟังก์ชันการสูญเสีย Dice coefficient loss.....	67
รูปที่ 3.22 อัลกอริทึมการเตรียมฝึกการเรียนรู้.....	68
รูปที่ 3.23 อัลกอริทึมชุดทดสอบโมเดลเพื่อประเมินประสิทธิภาพ.....	69
รูปที่ 3.24 อัลกอริทึมชุดทดสอบโมเดลที่ฝึกการเรียนรู้สำเร็จ.....	69
รูปที่ 3.25 อัลกอริทึมแปลงไฟล์อาร์เรย์เป็นไฟล์ภาพ NIFTI .....	69
รูปที่ 3.26 การสร้างรูปสามมิติจากไฟล์ภาพที่ได้จากการทำนายของโมเดล.....	70
รูปที่ 4.1 กราฟแสดงผลจากการเรียนรู้ของโมเดล.....	73
รูปที่ 4.2 แสดงภาพตัดขวางสองมิติบางส่วน.....	74
รูปที่ 4.3 แสดงภาพรูปทรงสามมิติ.....	75

# บทที่ 1

## บทนำ

### 1.1 ความเป็นมาและความสำคัญของปัญหา

หลอดเลือดแดงใหญ่ส่วนท้องโป่งพอง (AAA) เป็นภาวะที่อาจเป็นอันตรายถึงชีวิต การแตกของหลอดเลือดแดงที่โป่งพองนั้น มีความเป็นไปได้ที่จะมีความเกี่ยวข้องกับอัตราการตายของโรคสูงมากกว่า 50% ความถี่ของหลอดเลือดแดงใหญ่ในส่วนท้องโป่งพองอัตราเพิ่มขึ้นพร้อมกับอัตราการตายที่เพิ่มขึ้น ซึ่งมีค่าประมาณ 65% ถึง 85% สำหรับกรณีของการแตก เส้นผ่านศูนย์กลางสูงสุดของหลอดเลือดแดงใหญ่ส่วนท้องโป่งพอง เป็นพารามิเตอร์หลักที่ใช้ในการประเมินรอยโรคในผู้ป่วยเพื่อใช้ในการดูแลรักษา เนื่องจากมีความสัมพันธ์กับความเสียหายของการแตกของหลอดเลือด เมื่อตรวจพบจะต้องได้รับการการรักษาด้วยการผ่าตัด โดยที่จำเป็นต้องมีเส้นผ่านศูนย์กลางสูงสุดของหลอดเลือดแดงใหญ่ในส่วนท้องโป่งพองที่วัดจากบริเวณด้านหน้าไปยังด้านหลังในบริเวณที่มีการโป่งพองมากที่สุดประมาณ 50-55 มิลลิเมตร อย่างไรก็ตาม ถึงแม้ว่ามีความสัมพันธ์ระหว่างเส้นผ่านศูนย์กลางและความเสียหายของการแตกสูง แต่ยังคงมีการโป่งพองที่ทำให้หลอดเลือดแตกได้ที่เส้นผ่านศูนย์กลางที่มีขนาดเล็กกว่าโดยเฉพาะในผู้หญิง และการโป่งพองที่มีขนาดขยายมากขึ้นและไม่มีการแตกของหลอดเลือด ในความเป็นจริงสำหรับผู้ที่เป็โรคหลอดเลือดโป่งพองที่มีขนาดมากกว่า 55 มิลลิเมตร ความเสี่ยงต่อการเสียชีวิตจากสาเหตุอื่นที่ไม่ใช่ AAA นั้นสูงกว่าความเสี่ยงต่อการเสียชีวิตจากหลอดเลือดโป่งพองแล้วแตก แต่ปัจจัยเพิ่มเติมที่มีผลต่อการโป่งพองของหลอดเลือด ได้แก่ ลิ่มเลือดที่ผนังหลอดเลือดส่วนท้องที่หนา (ILT) การอักเสบของผนังหลอดเลือดที่โป่งพอง และความเค้นของผนังหลอดเลือดทางชีวกลศาสตร์ ได้รับการศึกษาแล้วว่าเป็นตัวบ่งชี้ที่มีผลต่อการเพิ่มขนาดและการแตกของ AAA อย่างรวดเร็ว และ ILT มีส่วนในการส่งผลต่อการเพิ่มขนาดของ AAA และเกิดความเสี่ยงของการแตกยังไม่เป็นที่แน่ชัด ซึ่งอาจเนื่องมาจากผลกระทบเชิงกลและทางชีวเคมีที่เสริมกัน

การเอกซเรย์คอมพิวเตอร์สำหรับดูเส้นเลือด (CTA) เป็นรูปแบบการสร้างภาพสำหรับการวินิจฉัย AAA เพื่อใช้ในการวินิจฉัยไปจนถึงการวางแผนการรักษาก่อนการผ่าตัด และการติดตามผลหลังการผ่าตัด โดยเฉพาะอย่างยิ่งการวินิจฉัยนั้น ประกอบด้วย การประเมินความเสี่ยงการแตก โดยการวัดขนาดของ AAA จากการสแกนเอกซเรย์คอมพิวเตอร์ (CT) การเอกซเรย์คอมพิวเตอร์สำหรับดูเส้นเลือดเป็นเทคนิคการสร้างภาพเบื้องต้นที่ใช้ในการประเมิน จัดการ และตรวจสอบการโป่งพองของหลอดเลือดส่วนท้อง การตัดแบ่งภาพที่แม่นยำของลูเมนของหลอดเลือด ในการสแกนเอกซเรย์คอมพิวเตอร์สำหรับดูเส้นเลือดเป็นขั้นตอนสำคัญในการวัดความยาวและเส้นผ่านศูนย์กลางของหลอดเลือด เพื่อความง่ายต่อการปรับขนาดของการทำรักษาด้วยวิธีการสวนหลอดเลือด

โครงข่ายประสาทแบบคอนโวลูชัน (CNNs) มีความสามารถในการตัดแบ่งภาพ AAA บริเวณใต้ต่อเส้นเลือดแดง renal แบบอัตโนมัติและมีความถูกต้องแม่นยำ จากการสแกนเพื่อสร้างภาพก่อนการผ่าตัดและหลังการผ่าตัด โครงข่ายประสาทแบบคอนโวลูชันเป็นการแสดงวัตถุที่ได้จากการเรียนรู้ที่มีความแม่นยำใน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ระดับต่างๆ ได้อย่างมีประสิทธิภาพจากการนำเข้าข้อมูลตัวอย่างที่ต้องการให้เรียนรู้ การตัดแบ่งภาพของหลอดเลือดแดงใหญ่ส่วนท้องที่โป่งพอง ทั้งมีและไม่มีลิ้มเลือดที่ผนังแบบอัตโนมัติเป็นเครื่องมือที่สามารถทำได้ด้วยความรวดเร็วและเป็นมาตรฐาน ในการวิเคราะห์กายวิภาคของ AAA และเป็นการแก้ปัญหาของการสังเกตอาการที่ไม่ปกติของบริเวณด้านในและบริเวณที่สามารถมองเห็นได้ยาก โดยเฉพาะอย่างยิ่งการตัดแบ่งภาพแบบอัตโนมัติจะช่วยการตรวจสอบในกระบวนการขยายตัวของลิ้มเลือดอย่างรวดเร็วจากฐานข้อมูลขนาดใหญ่ เกิดความสะดวกในการทำวิจัยทางคลินิก และสามารถรับมือกับสถานการณ์ที่จำเป็นต้องมีการตัดแบ่งภาพอย่างรวดเร็ว

## 1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

- 1.2.1 เพื่อนำภาพที่ได้จากการเรียนรู้เชิงลึกแบบสามมิติแบบอัตโนมัติมาสร้างรูปทรงแบบสามมิติ ที่สามารถใช้เพื่อวินิจฉัยโรค และวางแผนการรักษาโรค ด้วยสถาปัตยกรรมยูเน็ตสามมิติ (3D U-NET)
- 1.2.2 สามารถใช้การเรียนรู้เชิงลึกแบบสามมิติ แสดงผลลัพธ์ที่สามารถแบ่งแยกระหว่างหลอดเลือดแดงใหญ่ในช่องท้องและลิ้มเลือดที่ผนังหลอดเลือดได้ ด้วยสถาปัตยกรรมยูเน็ตสามมิติ (3D U-NET)
- 1.2.3 สามารถใช้ฟังก์ชันเสริม (Data Augmentation) ในการเพิ่มจำนวนข้อมูล เพื่อให้ได้ผลลัพธ์ที่สมบูรณ์ เพียงพอต่อการวินิจฉัยโรค

## 1.3 สมมุติฐานของการศึกษา

งานวิจัยนี้จัดทำขึ้นมาเพื่อใช้การเรียนรู้เชิงลึกแบบสามมิติอัตโนมัติด้วยสถาปัตยกรรมยูเน็ตสามมิติช่วยในการทำนายภาพและได้ภาพมีขอบเขตที่ถูกต้อง โดยที่สามารถแยกออกจากอวัยวะอื่นที่อยู่ใกล้เคียงที่มีระดับความขรุขระภาพเอกซเรย์คอมพิวเตอร์หลอดเลือดที่มีการฉีดสารทึบรังสีใกล้เคียงกัน เช่น กระดูกสันหลัง ลำไส้ ไต เป็นต้น เพื่อให้ได้รูปทรงของหลอดเลือดไม่มีอวัยวะอื่นๆ ปรากฏขึ้นมาในภาพ และงานวิจัยจะรวมไปถึงความสามารถในการทำนายพร้อมทั้งลิ้มเลือดที่ผนังหลอดเลือดอีกด้วย ซึ่งเป็นสาเหตุหนึ่งทำให้หลอดเลือดแตกได้

## 1.4 ขอบเขตการวิจัย

งานวิจัยเป็นงานเกี่ยวกับการใช้การเรียนรู้เชิงลึก ด้วยสถาปัตยกรรมยูเน็ตแบบสามมิติ เพื่อทำนายขอบเขตภาพหลอดเลือดแดงใหญ่ส่วนท้องสามแบบ คือ แบบปกติ แบบหลอดเลือดโป่งพอง และหลอดเลือดโป่งพองที่มีลิ้มเลือดที่ผนังหลอดเลือด โดยใช้ภาพเอกซเรย์คอมพิวเตอร์หลอดเลือดแดงใหญ่ส่วนท้องที่มีการฉีดสารทึบรังสี 60 ชุดข้อมูล เป็นไฟล์ภาพ DICOM และนำมาแปลงสกุลภาพเป็นไฟล์ Nifti โดยภาพเก็บข้อมูลมาจากโรงพยาบาลรามาริบัติ มหาวิทยาลัยมหิดล ซึ่งเป็นฐานข้อมูลขนาดใหญ่ โดยมีเอกสารผ่านกระบวนการวิจัยในคน

## 1.5 แผนการดำเนินงานวิจัย

ช่วงเวลาการดำเนินงานวิจัยนี้เริ่มตั้งแต่ภาคการศึกษาที่ 1/2563 สิงหาคม พ.ศ.2563 ถึงภาคการศึกษาที่ 2/2564 เดือนพฤษภาคม พ.ศ. 2565

ตารางที่ 1.1 แผนการดำเนินงานวิจัย

ลำดับ	วิธีดำเนินงานวิจัย	ปีการศึกษา			
		1/2563	2/2563	1/2564	2/2564
		ส.ค. - ธ.ค. 2563	ม.ค. - พ.ค. 2564	ส.ค. - ธ.ค. 2564	ม.ค. - พ.ค. 2565
1.	ศึกษาเรียนรู้เกี่ยวกับการประมวลผลดิจิทัล (Digital Image Processing)				
2.	ศึกษาเรียนรู้เกี่ยวกับการเรียนรู้เชิงลึก (Deep Learning)				
3.	เก็บข้อมูลภาพผู้ป่วย				
4.	ทำตัวอย่างการตัดแบ่งภาพของหลอดเลือดแบบการทำด้วยตัวเอง เพื่อเป็นตัวอย่างการเรียนรู้				
5.	ออกแบบอัลกอริทึมเพื่อใช้ในกระบวนการการเรียนรู้เชิงลึก				
6.	ใช้อัลกอริทึมที่ออกแบบไว้ ทำการฝึกการเรียนรู้เชิงลึก และทดสอบการฝึก				
7.	เก็บข้อมูลผลการทดลอง และบันทึก				
8.	ทำรูปเล่มวิจัย				

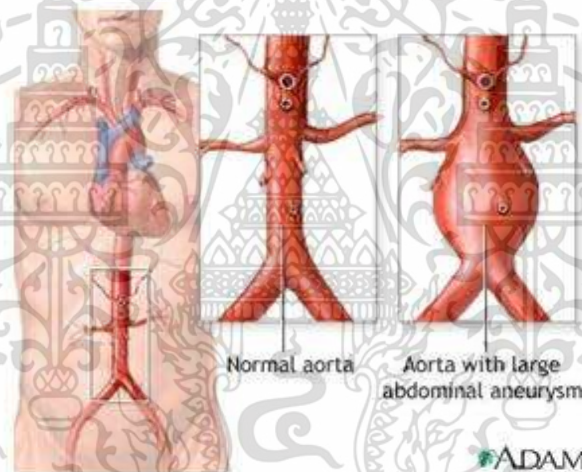
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 2

### ทฤษฎีที่เกี่ยวข้อง

#### 2.1 โรคหลอดเลือดแดงใหญ่ในช่องท้องโป่งพอง (Abdominal Aorta Aneurysm) [1]

เป็นความผิดปกติเฉพาะจุดของหลอดเลือดแดงใหญ่ในช่องท้อง ที่มีขนาดใหญ่กว่าปกติ 1.5 เท่า โดยปกติหลอดเลือดแดงใหญ่ในช่องท้องจะมีขนาดประมาณ 2 cm. ซึ่งถ้าตรวจพบว่าหลอดเลือดแดงใหญ่ในช่องท้องมีขนาดใหญ่กว่าหรือเท่ากับ 3 cm. ก็สงสัยได้ว่าเป็นหลอดเลือดแดงใหญ่โป่งพอง หลอดเลือดแดงใหญ่ในช่องท้องโป่งพองมักพบได้มากในคนอายุ 50 ปี หรือ หญิงอายุ 70 ปีขึ้นไป พบในผู้ชายมากกว่าผู้หญิงและในกลุ่มญาติที่มีประวัติสูบบุหรี่ ความดันโลหิตสูง ส่วนใหญ่ตำแหน่งที่โป่งพองจะอยู่ต่อหลอดเลือดที่ไปเลี้ยงไต (Infrarenal type) ซึ่งสาเหตุหลักมาจาก atherosclerosis ส่วนการโป่งพองอีกชนิดจะเกิดเหนือต่อหลอดเลือดที่ไปเลี้ยงไต (Suprarenal type)



รูปที่ 2.1 ตำแหน่งหลอดเลือดแดงใหญ่ส่วนท้อง ประเภทปกติและโป่งพอง  
(ที่มา : <http://www.lph.go.th/inforlphng/index.php/news/or/mcvt/24-evar>)

#### สาเหตุ

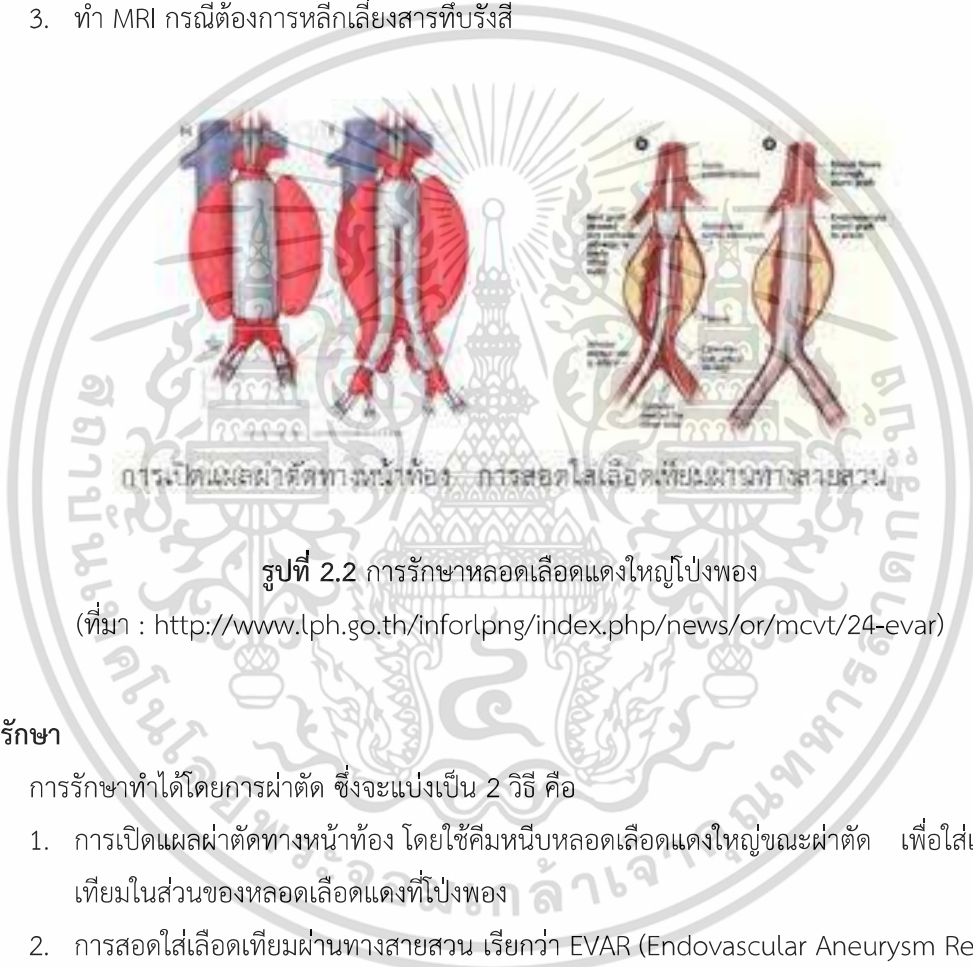
1. จากการพอกของแผ่นไขมันบริเวณผนังหลอดเลือดแดง
2. มีการติดเชื้อมีในหลอดเลือดแดง ทำให้มีการทำลายผนังหลอดเลือดบางส่วนจนเป็นรูทะลุ และขยายใหญ่ขึ้น
3. การบาดเจ็บของหลอดเลือดแดง ทำให้ผนังของหลอดเลือดฉีกขาดมีรูทะลุทำให้เกิดลิ้มเลือดมาออก และมีเนื้อเยื่อพังผืดมาปกคลุม ซึ่งผนังส่วนนี้จะไม่แข็งแรงเมื่อมีแรงดันของเลือดผ่านมาเรื่อย ๆ ก็จะขยาย และโป่งพองขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## การตรวจรักษา

ต้องมีการตรวจพิเศษเพื่อการวินิจฉัย 3 อย่างคือ

1. การถ่ายภาพหลอดเลือดแดงด้วยเครื่องเอกซเรย์คอมพิวเตอร์ (CTA) เป็นการตรวจสอบสภาพของหลอดเลือดแดงโดยฉีดสารทึบรังสีแล้วเอกซเรย์คอมพิวเตอร์ดูพยาธิสภาพของหลอดเลือดใหญ่ในช่องท้อง ปัจจุบันการตรวจพิเศษนี้เป็นการตรวจที่แม่นยำและมีประสิทธิภาพมากที่สุด
2. การฉีดสีเพื่อดูลักษณะของหลอดเลือดแดงต่างๆ ปัจจุบันไม่จำเป็น
3. ทำ MRI กรณีต้องการหลีกเลี่ยงสารทึบรังสี



รูปที่ 2.2 การรักษาหลอดเลือดแดงใหญ่โป่งพอง

(ที่มา : <http://www.lph.go.th/inforlphng/index.php/news/or/mcvt/24-avar>)

## การรักษา

การรักษาทำได้โดยการผ่าตัด ซึ่งจะแบ่งเป็น 2 วิธี คือ

1. การเปิดแผลผ่าตัดทางหน้าท้อง โดยใช้คีมหนีบลอดเลือดแดงใหญ่ขณะผ่าตัด เพื่อใส่เส้นเลือดเทียมในส่วนของหลอดเลือดแดงที่โป่งพอง
2. การสอดใส่เลือดเทียมผ่านทางสายสวน เรียกว่า EVAR (Endovascular Aneurysm Repair)

## การผ่าตัดสอดใส่หลอดเลือดเทียมผ่านสายสวน (EVAR)

เป็นวิธีที่ได้รับความนิยม เนื่องจาก

1. แผลผ่าตัดมีขนาดเล็ก และปวดแผลน้อยกว่าการเปิดแผลผ่าตัดทางหน้าท้อง
2. หลีกเลี่ยงการใช้คีมหนีบลอดเลือดแดงใหญ่ ซึ่งจะทำให้ความดันเลือดแกว่ง เกิดภาวะบีบคั้นหัวใจ
3. เสียเลือดน้อยกว่าการผ่าตัดเปิดหน้าท้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### การเตรียมตัวก่อนผ่าตัดสอดใส่หลอดเลือดเทียมผ่านสายสวน

ผู้ป่วยต้องมีการเตรียมร่างกายให้พร้อมสำหรับการผ่าตัด ที่สำคัญ คือ

- งดน้ำและอาหารอย่างน้อย 6-8 ชั่วโมง
- ทำความสะอาดร่างกายทั่วไปโดยเฉพาะบริเวณขาหนีบทั้ง 2 ข้าง
- ก่อนผ่าตัดผู้ป่วยจะได้รับการดมยาสลบ
- หลังผ่าตัด จะมีแผลประมาณ 3-4 บริเวณขาหนีบ 2 ข้าง

### คำแนะนำเกี่ยวกับการปฏิบัติตัวเมื่อกลับบ้าน

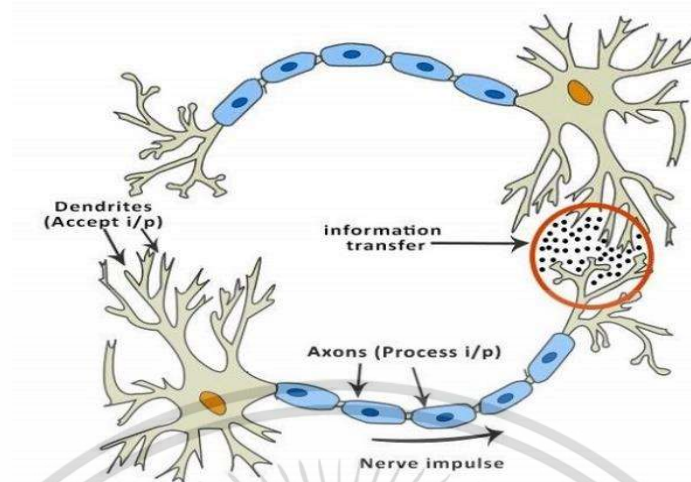
1. การพักผ่อนใช้เวลา 2-3 สัปดาห์
2. การอาบน้ำควรอาบ เมื่อแผลที่ขาหนีบแห้งดีแล้ว
3. การรับประทานอาหารควรเลี่ยงอาหารเค็ม โดยเฉพาะผู้ป่วยที่เป็นความดันโลหิตสูง เพราะอาจทำให้เกิดภาวะหลอดเลือดแดงส่วนอื่นโป่งพองได้
4. รับประทานยาตามแพทย์สั่งให้เคร่งครัด
5. มาตรวจตามแพทย์นัดทุกครั้งอย่างเคร่งครัด

## 2.2 โครงข่ายประสาทเทียมเบื้องต้น [2]

โครงข่ายประสาทเทียม (Artificial Neuron Network: ANN) คือ แบบจำลองทางคณิตศาสตร์ที่พัฒนาขึ้นเพื่อจำลองการทำงานของโครงข่ายประสาทในสมองมนุษย์ โดยที่โครงข่ายประสาทของมนุษย์ประกอบด้วยเซลล์ประสาท (Neuron) และจุดประสานประสาทหรือไซแนปส์ (Synapses) โดยโครงสร้างของการส่งสัญญาณ ประสาทประกอบจากการเชื่อมต่อระหว่างเซลล์ประสาทหลายพันล้านเซลล์ เซลล์ประสาทแต่ละเซลล์ประกอบด้วยแขนงรับสัญญาณประสาทซึ่งเป็นเสมือนหน่วยรับข้อมูลป้อนเข้า เรียกว่า “เดนไดรต์” (dendrites) และส่วนปลายของเซลล์ประสาทในการส่งสัญญาณประสาทซึ่งเป็นเสมือนหน่วยส่งข้อมูลออกของเซลล์ เรียกว่า “แอกซอน” (axon) โดยการส่งสัญญาณประสาทดังกล่าว อาจทำให้เกิดได้ทั้งการกระตุ้นและยับยั้ง ทั้งนี้นอกจากลักษณะดังกล่าวแล้ว วิธีการประมวลผลภายในเซลล์ประสาทแต่ละเซลล์ยังมีการขยายหรือลดขนาดของสัญญาณอีกด้วย โดยสัญญาณจากเดนไดรต์ที่ต่าง ๆ จะรวมกันเข้าสู่เซลล์ประสาท และหากสัญญาณรวมมีความแรงเกินค่าระดับ (threshold) ของเซลล์ประสาทนั้น ๆ เซลล์ประสาทก็จะส่งสัญญาณออกทางแอกซอนต่อไป ลักษณะโดยทั่วไปของโครงสร้างระบบประสาทแสดงดังรูปที่ 2.3

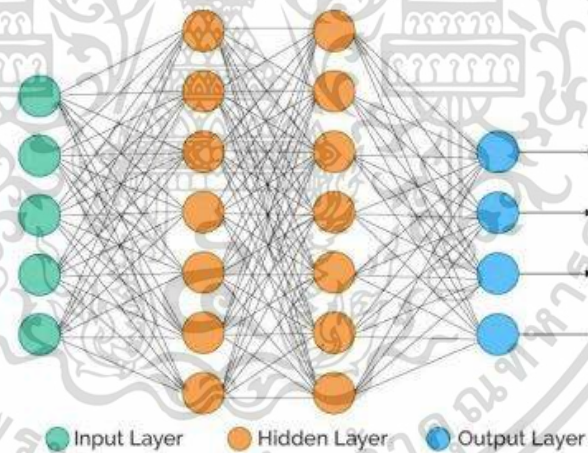
### 2.2.1 องค์ประกอบของโครงข่ายประสาทเทียม [3]

วิธีการสร้างแบบจำลองข้อมูลหรือโมเดลข้อมูล (Data modeling) โดยวิธีโครงข่ายประสาทเทียมอัจฉริยะ เป็นวิธีการที่มีพื้นฐานมาจากการเลียนแบบการทำงานของสมองมนุษย์ซึ่งประกอบด้วยนิวรอน (Neurons) ที่ใช้ในการเชื่อมโยงเพื่อแก้ปัญหาต่าง ๆ โดยมีโครงสร้างการทำงาน ตามรูปที่ 2 (Zou, Han, and So 2009) ซึ่งประกอบด้วยข้อมูลเข้า (Input) 1 ชั้น และข้อมูลออก (Output) 1 ค่าจะเรียกว่าเพอร์เซพตรอน (Perceptron) ดังรูปที่ 2.4



รูปที่ 2.3 โครงสร้างระบบประสาท

(ที่มา : [https://www.tutorialspoint.com/artificial\\_intelligence/artificial\\_intelligence\\_neural\\_networks.htm](https://www.tutorialspoint.com/artificial_intelligence/artificial_intelligence_neural_networks.htm))



รูปที่ 2.4 ส่วนประกอบของโครงข่ายประสาทเทียม

(ที่มา : <https://csit.nu.ac.th/kraisak/ds/ds/chapter07/Chapter07.pdf>)

ส่วนประกอบของโครงข่ายประสาทเทียมจากรูปที่ 2.4 จะเห็นว่าส่วนประกอบของโครงข่ายประสาทเทียมประกอบด้วย 3 ส่วน

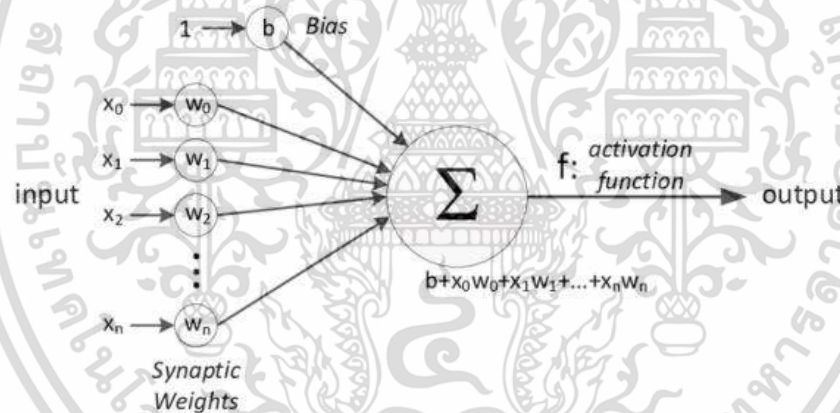
1. ชั้นข้อมูลเข้า (Input Layer) ที่ชั้นนี้จะเป็นข้อมูลเข้า จำนวนของโหนดในชั้นนี้ขึ้นอยู่กับจำนวนของข้อมูลเข้าว่ามีข้อมูลอะไรบ้างที่จะนำเข้ามาคิดในแบบจำลอง เช่น ถ้าข้อมูลของกลุ่มลูกค้าเป็นข้อมูลนำเข้าที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประกอบด้วย อายุ เพศ จังหวัดที่อาศัย รวมทั้งสิ้น 4 อย่าง ดังนั้นชั้นข้อมูลเข้าก็จะมี 4 โหนด ซึ่งอาจจะเรียกปัจจัยที่นำมาวิเคราะห์เหล่านี้ว่าคุณลักษณะ (Feature)

2. ชั้นซ่อน (Hidden Layer) เป็นชั้นที่อยู่ระหว่างกลาง ซึ่งจะมีผลอย่างมากต่อประสิทธิภาพในการเรียนรู้ของแบบจำลอง ซึ่งในชั้นซ่อนนั้นจะมีกี่ชั้นก็ได้ และแต่ละชั้นจะมีจำนวนของนิวรอนจำนวนเท่าไรก็ได้เช่นกัน ซึ่งการเพิ่มชั้นและจำนวนนิวรอนจะส่งผลต่อการทำงานของโมเดล ในส่วนของชั้นซ่อนมีการทำงานเปรียบเสมือนส่วนที่เรียนรู้ข้อมูลเชิงลึก หรือ deep learning นั่นเอง โดยสิ่งสำคัญในชั้นซ่อนอีกประการหนึ่งคือ ทุก ๆ โหนดต้องประกอบด้วยฟังก์ชันแบบไม่เป็นเชิงเส้น ซึ่งจะอธิบายต่อไป

3. ชั้นข้อมูลออก (Output Layer) ชั้นที่จะนำเอาข้อมูลจากการคำนวณไปใช้ และจำนวนของโหนดในชั้นนี้ขึ้นอยู่กับรูปแบบของข้อมูลออกที่จะเอาไปใช้ ตัวอย่างเช่น ถ้างานที่ทำเป็นสมการถดถอย (Regression) ก็กำหนดให้ชั้นข้อมูลออกเป็นแบบ 1 โหนด เพราะต้องการคำตอบเพียงค่าเดียว ถ้าเป็นหลายค่าก็เพิ่มไปตามที่ต้องการ เช่น ในบางงานอาจจะทำนายหาตำแหน่งของภาพในแกน  $x$  และ  $y$  พร้อม ๆ กัน ในกรณีนี้ก็ต้งกำหนดชั้นข้อมูลออกเป็น 2 โหนด เป็นต้น



รูปที่ 2.5 โครงสร้างการทำงานของโครงข่ายประสาทเทียม

(ที่มา : <https://csit.nu.ac.th/kraisak/ds/ds/chapter07/Chapter07.pdf>)

รูปที่ 2.5 ประกอบด้วยข้อมูลเข้า  $x_1, x_2, x_3, \dots, x_n$  ซึ่งเข้าทางปลายประสาทหรือเรียกว่า “Synapse” ที่ปลายของประสาทเทียมนี้จะมีค่าน้ำหนักต่าง ๆ กัน แสดงถึงความแข็งแรงของแต่ละโหนดของปลายประสาท ซึ่งแทนด้วย  $w_1, w_2, w_3, \dots, w_n$  และค่า  $b$  คือค่าความลำเอียง (bias) เพื่อให้ฟังก์ชันกระตุ้น (Activation function) ทำงานได้มีประสิทธิภาพมากยิ่งขึ้น

โครงข่ายประสาทเทียม จะนำข้อมูลเข้าและค่าน้ำหนักแต่ละปลายประสาทมาคูณกัน และนำค่าทั้งหมดมารวมกัน หากค่าที่รวมกันนี้เกินกว่าค่าแบ่งที่กำหนดไว้ (threshold : T) จะทำการส่งต่อข้อมูลไปยังฟังก์ชันกระตุ้น ดังสมการ (1)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$y = f \left[ \left( b + \sum_{i=1}^n (x_i w_i) \right) - T \right] \quad (1)$$

ฟังก์ชันกระตุ้น หรือเรียกอีกชื่อว่า “ฟังก์ชันการส่งต่อ (Transfer function)” เป็นฟังก์ชันในการคำนวณ เพื่อทำนายค่าของข้อมูลออก รูปแบบที่นิยมมากที่สุดและมีประโยชน์คือแบบที่ไม่เป็นฟังก์ชันสมการเส้นตรง (Non-linear function) เนื่องจากปัญหาในโลกความเป็นจริงมีลักษณะเป็นแบบสมการเส้นตรงน้อยมาก (Minsky, Papert, and Bottou 2017) ฟังก์ชันกระตุ้นทำหน้าที่ในการตัดสินใจว่านิวรอนควรจะถูกกระตุ้นหรือไม่ โดยดูค่าผลรวมของข้อมูลเข้าและค่าน้ำหนักฟังก์ชันกระตุ้นจะถูกนำไปใช้ทั้งโหนดซ่อน (Hidden node) และโหนดข้อมูลออก (Output node) ซึ่งทั้งสองโหนดอาจจะใช้ฟังก์ชันกระตุ้นที่เหมือนหรือต่างกันได้ แต่ส่วนมากจะใช้ฟังก์ชันแบบไม่เป็นเชิงเส้น เนื่องจากในโหนดซ่อนจะมีการคำนวณแบบการรวมเชิงเส้น (Linear combination) ถ้าฟังก์ชันกระตุ้นของโหนดซ่อนจะมีการคำนวณแบบเชิงเส้นอีก จะเป็นการทำงานซ้ำซ้อนกับการคำนวณแบบการรวมเชิงเส้นในชั้นข้อมูลออก และจะทำให้ผลลัพธ์เทียบเท่ากับสมการถดถอยลจิสติก ฟังก์ชันกระตุ้นจะมีหลากหลายรูปแบบ ดังต่อไปนี้

1. ฟังก์ชันกระตุ้นค่าแบ่ง (Threshold Activation Function) ฟังก์ชันนี้เรียกอีกชื่อว่าฟังก์ชันไบนารีสเตป (Binary step function) ซึ่งจะพิจารณาค่าข้อมูลเข้าว่ามากกว่าหรือน้อยกว่าค่าแบ่งที่กำหนดไว้ (threshold) หรือไม่ เพื่อส่งค่าต่อไปยังชั้นถัดไป

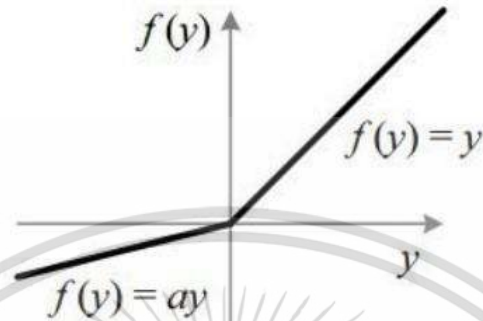
2. ฟังก์ชันกระตุ้นซิกมอยด์ (Sigmoid Activation Function) เป็นฟังก์ชันทางคณิตศาสตร์ที่มีลักษณะเป็นตัวเอส “S-curve” หรือเรียกว่า “Sigmoid curve” จะมีค่าระหว่าง 0 และ 1 ฟังก์ชันนี้จะใช้เมื่อต้องการทำนายความน่าจะเป็น (Probability) ของข้อมูลออก

3. ฟังก์ชันไฮเพอร์โบลิกแทนเจนต์ (Hyperbolic Tangent Function: tanh) มีการทำงานคล้ายฟังก์ชันกระตุ้นซิกมอยด์ แต่มีประสิทธิภาพดีกว่า จะมีค่าระหว่าง  $[-1, 1]$  ข้อดีของฟังก์ชันไฮเพอร์โบลิกแทนเจนต์คือ สามารถแปลงค่าข้อมูลเข้าที่มีค่าเป็นลบมาก ๆ ให้เป็นข้อมูลออกที่ติดลบได้ และข้อมูลที่ค่าเป็นศูนย์จะถูกแปลงเป็นข้อมูลออกที่มีค่าใกล้ศูนย์ (near-zero output)

4. ฟังก์ชันเรกติไฟด์ลินเนี่ยยูนิต (Rectified Linear Units, ReLu) เป็นฟังก์ชันที่นิยมใช้งานมากที่สุด ในโครงข่ายประสาทเทียมแบบบิด (Convolutional Neural Networks: CNN) และโครงข่ายประสาทเทียมอัจฉริยะ (ANN) ฟังก์ชันนี้จะมีค่าอยู่ระหว่าง  $[0, \infty]$  หมายถึงถ้าข้อมูลเข้ามีค่ามากกว่าศูนย์ ข้อมูลออกจะมีค่าเป็นบวก และถ้าข้อมูลเข้ามีค่าศูนย์หรือติดลบ ข้อมูลออกจะมีค่าเป็นศูนย์

นักวิทยาศาสตร์พิสูจน์แล้วว่าฟังก์ชันเรกติไฟด์ลินเนี่ยยูนิตมีประสิทธิภาพดีกว่าฟังก์ชันไฮเพอร์โบลิกแทนเจนต์สูงถึง 6 เท่า (Zou et al. 2009) ข้อเสียของฟังก์ชันเรกติไฟด์ลินเนี่ยยูนิตคืออาจจะเกิดปัญหานิวรอนตาย (dead neurons) กล่าวคือ ฟังก์ชันกระตุ้นไม่ทำงานไม่ว่าข้อมูลเข้ามีค่าเท่าไรก็ตาม จึงมีการพัฒนา ฟังก์ชันเรกติไฟด์ลินเนี่ยยูนิตเป็นฟังก์ชัน “ลึกลับเรกติไฟด์ลินเนี่ยยูนิต (Leaky ReLu)” ซึ่งจะมีค่าจาก  $-\infty$  ถึง  $+\infty$  การเลือกฟังก์ชันกระตุ้นสำหรับโหนดข้อมูลออกต้องเลือกสิ่งที่เข้ากับคำตอบที่เราต้องการ

เช่น เราต้องการทำนายความน่าจะเป็นของการเป็นโรค ซึ่งคำตอบที่เป็นไปได้คือ เป็น (0) และไม่เป็น (1) เราจึงควรเลือกฟังก์ชัน กระตุ้นซิกมอยด์ เพราะฟังก์ชันนี้ให้ผลลัพธ์ระหว่าง 0 ถึง 1



รูปที่ 2.6 ฟังก์ชันลึกลับเรคตีไฟต์ลีนเนี่ยยูนิต

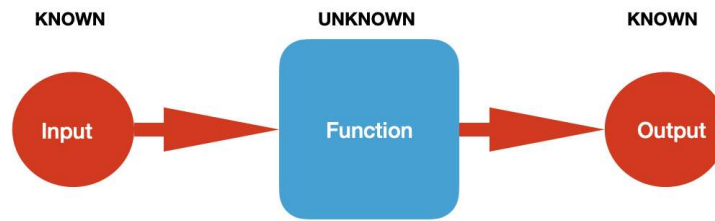
(ที่มา : <https://csit.nu.ac.th/kraisak/ds/ds/chapter07/Chapter07.pdf>)

### 2.3. การเรียนรู้ของโครงข่ายประสาทเทียม (learning of neural network) [4]

โดยทั่วไปการเรียนรู้ของโครงข่ายประสาทเทียม ก็คือการสอนโครงข่ายให้ทำการคำนวณข้อมูลเอาต์พุตพร้อมกับปรับปรุณค่าน้ำหนักโดยใช้ข้อมูลอินพุตที่ป้อนให้กับโครงข่าย โดยอาศัยกระบวนการทำซ้ำ (iterative) สามารถแบ่งเป็น 3 ประเภท คือ

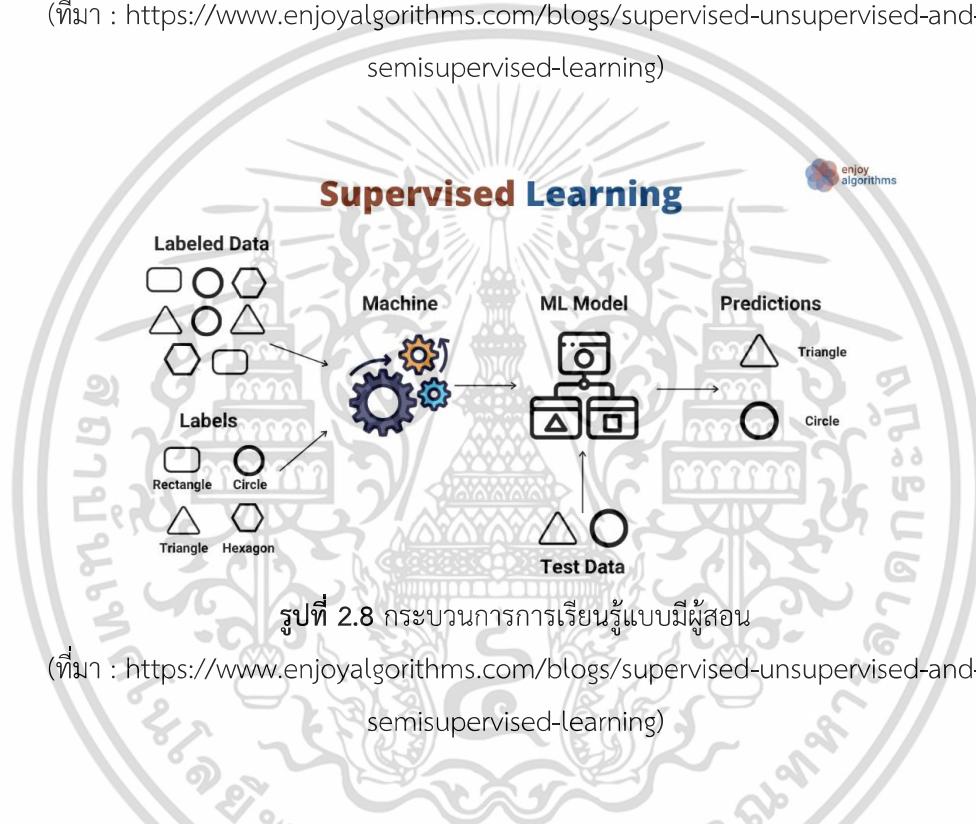
2.3.1 การเรียนรู้แบบมีผู้สอน (Supervised Learning) การเรียนรู้แบบมีผู้สอนนั้นต้องการชุดข้อมูลอินพุตและเอาต์พุตเป้าหมายเป็นชุดฝึกสอนควบคู่ (training pair) โดยปกติการสอนโครงข่ายนั้นจะใช้ชุดฝึกสอนควบคู่กันหลายชุด ในระหว่างการสอนโครงข่ายจะเกิดเอาต์พุตจริงซึ่งแตกต่างจากเอาต์พุตเป้าหมาย ทำให้ได้ค่าความคลาดเคลื่อนหรือค่าความผิดพลาด โดยโครงข่ายจะเรียนรู้ข้อมูลทั้งสองโดยการปรับค่าน้ำหนักเพื่อลดค่าความแตกต่าง หรือค่าความผิดพลาดระหว่างค่าของตัวแปรเอาต์พุตของโครงข่ายกับค่าข้อมูลเอาต์พุตที่ถูกต้องให้น้อยที่สุด การปรับค่าน้ำหนักจะปรับทีละน้อย ๆ โดยกระบวนการทำซ้ำกับข้อมูลที่ละชุดจนกระทั่งค่าน้ำหนักในโครงข่ายลู่เข้าซึ่งทั้งหมดนี้เรียกว่าการเรียนรู้ จากนั้นเมื่อเราป้อนค่าข้อมูลอินพุตล่าสุด ซึ่งเป็นข้อมูลชุดใหม่ก็จะได้ค่าตัวแปรเอาต์พุตของโครงข่าย เมื่อโครงข่ายทำการเรียนรู้แล้วก็จะป้อนข้อมูลอินพุตล่าสุดให้กับโครงข่าย เพื่อที่จะหาค่าของตัวแปรเอาต์พุตซึ่งคือค่าผลการทำนายหรือระบุค่าหรือวลีสำคัญ เป็นต้น ซึ่งในวิทยานิพนธ์นี้จะใช้การเรียนรู้แบบมีผู้สอนเท่านั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.7 โครงสร้างการเรียนรู้แบบมีผู้สอน

(ที่มา : <https://www.enjoyalgorithms.com/blogs/supervised-unsupervised-and-semisupervised-learning>)

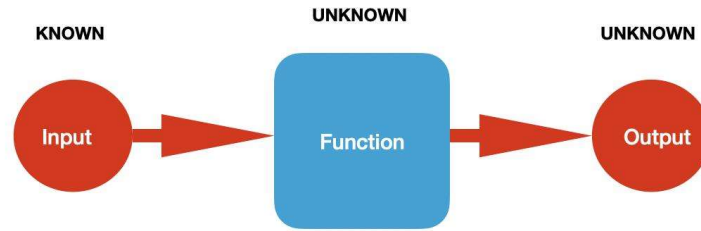


รูปที่ 2.8 กระบวนการการเรียนรู้แบบมีผู้สอน

(ที่มา : <https://www.enjoyalgorithms.com/blogs/supervised-unsupervised-and-semisupervised-learning>)

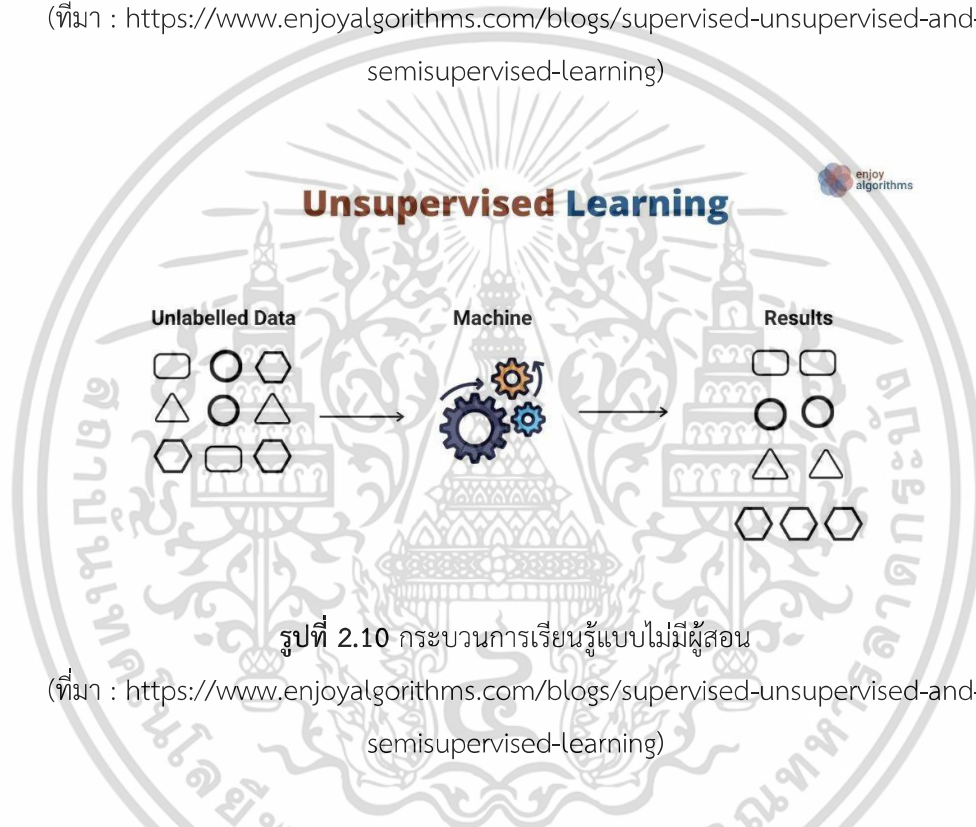
2.3.2 การเรียนรู้แบบไม่มีผู้สอน (unsupervised learning) การเรียนรู้แบบไม่มีผู้สอนนั้นได้ถูกพัฒนาเพื่อให้ใกล้เคียงกับระบบการเรียนรู้ของสมองมนุษย์มากขึ้น โดยจะมีเพียงชุดข้อมูลอินพุตเท่านั้น จากนั้นกระบวนการเรียนรู้จะใช้หลักทางสถิติโดยหาค่าทางสถิติของชุดฝึกสอน และทำการจัดกลุ่มข้อมูลออกเป็นระดับต่าง ๆ โดยโครงข่ายประสาทเทียมจะหาค่าเอาต์พุตเองจากความสัมพันธ์ระหว่างข้อมูลอินพุตและเอาต์พุต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.9 โครงสร้างการเรียนรู้แบบไม่มีผู้สอน

(ที่มา : <https://www.enjoyalgorithms.com/blogs/supervised-unsupervised-and-semisupervised-learning>)

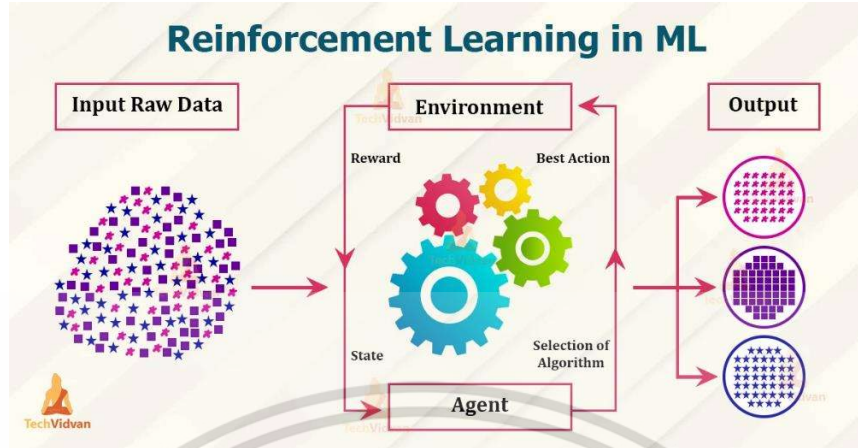


รูปที่ 2.10 กระบวนการเรียนรู้แบบไม่มีผู้สอน

(ที่มา : <https://www.enjoyalgorithms.com/blogs/supervised-unsupervised-and-semisupervised-learning>)

2.3.3 การเรียนรู้เชิงบังคับ (reinforcement learning) การเรียนรู้เชิงบังคับนั้นเป็นการเรียนรู้แบบมีผู้สอนและไม่มีผู้สอน โดยจะใช้การเรียนรู้แบบไม่มีผู้สอนในระหว่างการสอนที่มีเพียงชุดข้อมูลอินพุต และจะใช้การเรียนรู้แบบมีผู้สอนคือเมื่อได้ค่าเอาต์พุตและจะทำการบอกว่าถูกหรือผิดแต่จะไม่บอกว่าเอาต์พุตที่ถูกคืออะไร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.11 กระบวนการเรียนรู้เชิงบังคับ

(ที่มา : <https://techvidvan.com/tutorials/reinforcement-learning/>)

## 2.4 การเรียนรู้เชิงลึก (Deep learning) [5]

การเรียนรู้เชิงลึกเป็นแนวการวิจัยที่เกิดจาก machine learning และปัญญาประดิษฐ์ โดยใช้โครงข่ายประสาทเทียมเชิงลึกเพื่อจำลองกระบวนการเรียนรู้ของสมองมนุษย์ และแยกคุณสมบัติจากฐานข้อมูลขนาดใหญ่ เช่น เสียง ข้อความ รูปภาพ เป็นต้น ในแบบการเรียนรู้ที่ไม่มีผู้สอน โครงข่ายประสาทประกอบด้วยเซลล์ประสาทจำนวนมาก เซลล์ประสาทแต่ละเซลล์ถือได้ว่าเป็นหน่วยประมวลผลข้อมูลขนาดเล็ก เซลล์ประสาทเชื่อมต่อกันในวิธีหนึ่งในการสร้างโครงข่ายประสาทเทียมเชิงลึกทั้งหมด การเกิดขึ้นของโครงข่ายประสาททำให้การประมวลผลภาพแบบ end-to-end เมื่อชั้นซ่อนของโครงข่ายพัฒนาไปหลายชั้น ซึ่งเรียกว่า การเรียนรู้เชิงลึก เพื่อใช้ในการแก้ปัญหาคอมพิวเตอร์วิทัศน์ที่มีความยุ่งยาก โดยจำเป็นต้องมีการกำหนดค่าเริ่มต้นแบบขั้นต่อขั้น (layer by layer) และการแบท (batch) ซึ่งทำให้การเรียนรู้เชิงลึกได้รับความนิยมอย่างมากในการทำวิจัยในด้านคอมพิวเตอร์ การเรียนรู้เชิงลึกส่วนใหญ่จะใช้ในมิติของการลดจำนวนข้อมูล การทำนายตัวเลขที่เขียนด้วยลายมือ การจดจำรูปแบบต่างๆ และในด้านอื่นๆ เช่น การจดจำภาพ การปรับปรุงภาพ การแบ่งส่วนภาพ การติดตามวัตถุ การวิเคราะห์รูปภาพ ที่มีประสิทธิภาพสูงเป็นต้น

## 2.5 โครงข่ายประสาทคอนโวลูชัน (Convolutional Neural Networks : CNN)

โครงข่ายประสาทคอนโวลูชันเป็นแบบจำลองพื้นฐานที่เกิดจากการผสมผสานระหว่างการเรียนรู้เชิงลึกและเทคโนโลยีการประมวลผลภาพ โครงข่ายประสาทที่เป็นตัวแทนหนึ่งในด้านเทคโนโลยีการเรียนรู้เชิงลึก มีความก้าวหน้ามากมายในด้านการวิเคราะห์และประมวลผลภาพ ในชุดคำอธิบายมาตรฐาน (annotation set) ของ ImageNet ซึ่งใช้กันทั่วไปในทางวิชาการและประสบผลสำเร็จอย่างมาก จากการใช้โครงข่ายประสาทเทียมคอนโวลูชัน รวมถึงการทำนายและการจำแนกประเภทของภาพ การทำนายรูปแบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นอนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

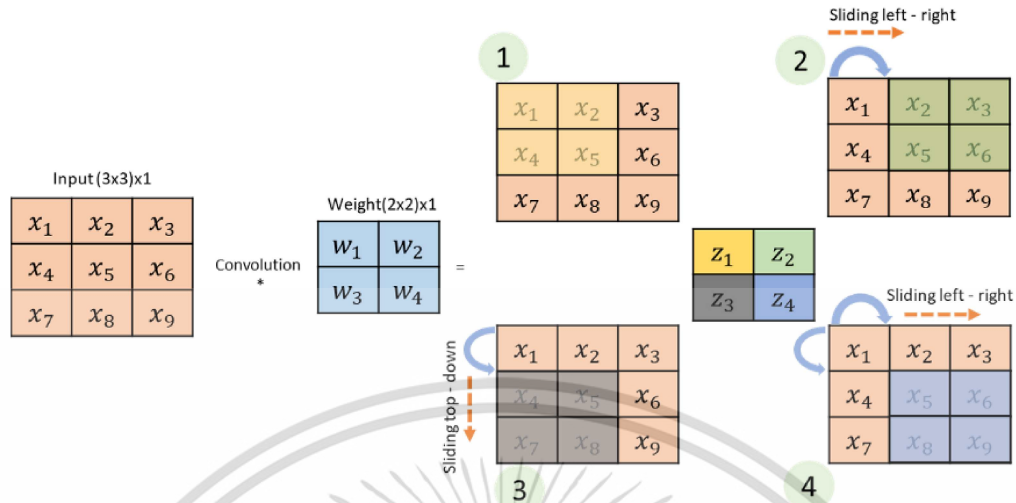
ภาพ เป็นต้น โครงข่ายประสาทเทียมคอนโวลูชันเป็นแบบจำลองเชิงลึกด้วยการเรียนรู้แบบมีผู้สอน แนวคิดพื้นฐานคือการแบ่งปันน้ำหนักของการแมปคุณลักษณะตำแหน่งต่างๆ ของโครงข่ายในชั้นก่อนหน้า และลดจำนวนพารามิเตอร์ โดยใช้ความสัมพันธ์เชิงพื้นที่เพื่อปรับปรุงประสิทธิภาพการฝึกอบรม แนวคิดของ CNN นั้นค่อนข้างเป็นแนวคิดที่ดีมาก แต่สิ่งที่ซับซ้อนของมันคือระบบการคำนวณที่สอดคล้องกับคอนเซปของมันเป็นเองและต้องมีคณิตศาสตร์มารองรับ โดยการคำนวณตามแนวคิดนี้ใช้หลักการเดียวกันกับคอนโวลูชันเชิงพื้นที่ (Spatial Convolution) ในการทำงานด้านการประมวลผลภาพ การคำนวณนี้จะเริ่มจากการกำหนดค่าในตัวกรอง (filter) หรือเคอร์เนล (kernel) ที่ช่วยดึงคุณลักษณะที่ใช้ในการเรียนรู้จำวัตถุออก โดยปกติตัวกรอง/เคอร์เนลอันหนึ่งจะดึงคุณลักษณะที่สนใจออกมาได้หนึ่งอย่าง เราจึงจำเป็นต้องมีตัวกรองหลายตัวกรองด้วยเพื่อหาคุณลักษณะทางพื้นที่หลายอย่างประกอบกัน ใจเดียวหลักของ CNN คือการใช้ Layer ชนิดพิเศษ ที่เรียกว่า Convolution layer ซึ่งทำหน้าที่สกัดเอาส่วนต่างๆ ของภาพออกมา เช่น เส้นขอบของวัตถุต่างๆ เพื่อให้โมเดลสามารถเรียนรู้ลักษณะของภาพได้อย่างมีประสิทธิภาพและแม่นยำ ใน CNN จะใช้ Convolution layer มาประกอบกับ Layer ชนิดอื่น เช่น Pooling layer แล้วนำกลุ่ม Layer ดังกล่าวมาซ้อนต่อกัน โดยอาจเปลี่ยน Hyperparameter บางอย่าง เช่นขนาดของ Filter layer (ซึ่งเป็นส่วนหนึ่งของ Convolution layer) และจำนวน Channel ของ layer วิธีการนำเอาส่วนต่างๆ มาประกอบกันนี้เรียกว่าเป็นโครงสร้าง (Architecture) ของ CNN ซึ่งมีหลายแบบ เช่น LeNet, AlexNet, VGG, ResNet, Inception Network เป็นต้น แต่ตอนนี้เรามุ่งความเข้าใจส่วนประกอบต่างๆ ของ CNN ซึ่งเป็นพื้นฐานที่จะนำมาประกอบกันเสียก่อน

### 2.5.1 Convolutional Layer [6]

เป็น Layer หลักของ CNN ทำหน้าที่รับ Input เข้ามา จากนั้นจะทำการดำเนินการทาง คณิตศาสตร์ เพื่อหาคุณสมบัติที่สำคัญจากรูปภาพการคำนวณจะเริ่มจากการกำหนดค่าใน ตัวกรอง (Filter) หรือ เคอร์เนล (Kernel) ที่ช่วยดึงคุณลักษณะที่ใช้ในการรู้จำวัตถุออกมา หรือที่เรียกว่า Feature Map

**ตัวอย่าง** คำนวณ Convolutional [7]

กำหนด Input ขนาด  $3 \times 3 \times 1$  และ Weight ขนาด  $2 \times 2 \times 1$



รูปที่ 2.12 วิธีการคำนวณคอนโวลูชัน

(ที่มา : <https://medium.com/@hadee2531earvesdrop/convolutional-neural-network-%E0%B8%84%E0%B8%B7%E0%B8%AD%E0%B8%AD%E0%B8%B0%E0%B9%84%E0%B8%A3-42c45f7ec16b>)

ขั้นตอนการคำนวณคือ หาผลรวมของการคูณระหว่าง Input กับ weight โดยใช้ weight ชุดเดิมแล้วสแกนไปทั้ง Input จากซ้ายไปขวา และบนลงล่าง (1)-(4) ผลลัพธ์ (Z) ที่ได้คือ

$$z_1 = w_1x_1 + w_2x_2 + w_3x_4 + w_4x_5$$

$$z_2 = w_1x_2 + w_2x_3 + w_3x_5 + w_4x_6$$

$$z_3 = w_1x_4 + w_2x_5 + w_3x_7 + w_4x_8$$

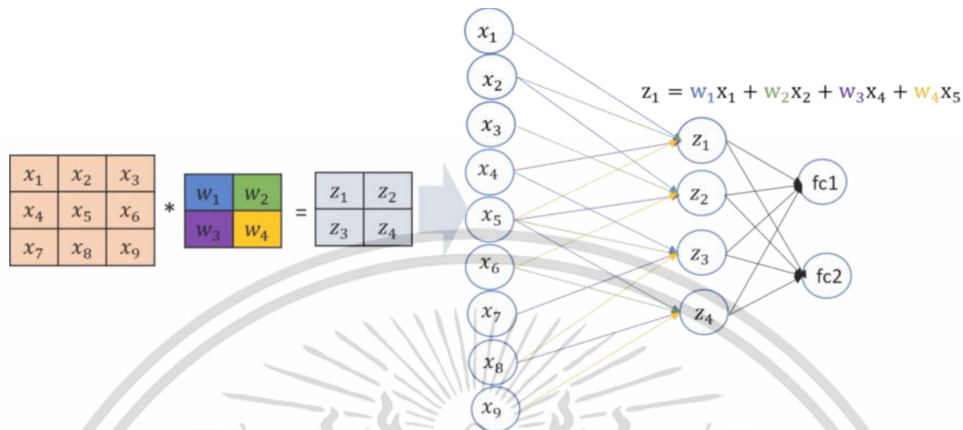
$$z_4 = w_1x_5 + w_2x_6 + w_3x_8 + w_4x_9$$

รูปที่ 2.13 ขั้นตอนการคำนวณคอนโวลูชัน

(ที่มา : <https://medium.com/@hadee2531earvesdrop/convolutional-neural-network-%E0%B8%84%E0%B8%B7%E0%B8%AD%E0%B8%AD%E0%B8%B0%E0%B9%84%E0%B8%A3-42c45f7ec16b>)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หากเขียนในรูปแบบ Neural network จะได้ว่า



รูปที่ 2.14 การเขียนรูปแบบ Neural network

(ที่มา : <https://medium.com/@hadee2531earvesdrop/convolutional-neural-network-%E0%B8%84%E0%B8%B7%E0%B8%AD%E0%B8%AD%E0%B8%B0%E0%B9%84%E0%B8%A3-42c45f7ec16b>)

จะเห็นได้ว่าสิ่งที่แตกต่างจาก Neural network ปกติ นั่นคือ ค่า Weight ที่ใช้คำนวณ  $Z$  นั้นจะเป็นค่าชุดเดียวกันหมด และนี่คือหนึ่งเหตุผลว่า ทำไมถึงใช้ ConvNet เพราะถ้าเราใช้ Neural network ปกติ โมเดลจะต้องเรียนรู้ค่า weight จำนวนมาก

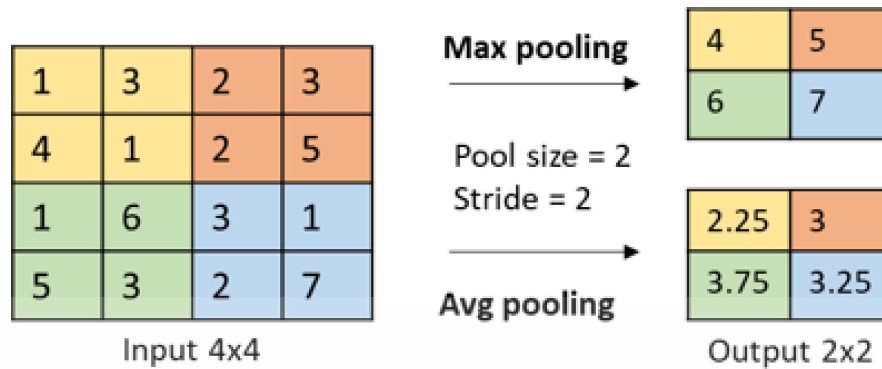
### 2.5.2 Pooling layer

การปรับขนาดข้อมูล แนวคิดคือ แทนที่เราจะใช้ข้อมูลภาพทั้งหมด เราจะแบ่งเป็นพื้นที่แล้วเลือกบางค่าจากพื้นที่นั้น (เหมือนกับลักษณะการมองของคน ที่บางครั้งมองภาพคร่าวๆ ก็บอกได้ว่า คืออะไร) โดยทั่วไป Pooling layer มีสองแบบที่เราสามารถเลือกใช้ได้ คือ

- Max pooling : เลือกค่าสูงสุดจาก Pool size
- Average pooling: ค่าเฉลี่ยของ Pool size

พารามิเตอร์ที่ต้องกำหนด (Hyper parameters) คือ สเตปการสแกน (Stride) และ ขนาดของ Pool ตัวอย่างเช่น มี Input ขนาด 4x4, กำหนดขนาด Pool size =2 และ stride = 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นอนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.15 วิธีการคำนวณ Pooling Layer

(ที่มา : <https://medium.com/@hadee2531earvesdrop/convolutional-neural-network-%E0%B8%84%E0%B8%B7%E0%B8%AD%E0%B8%AD%E0%B8%B0%E0%B9%84%E0%B8%A3-42c45f7ec16b>)

ตัวอย่าง Pool สีเหลือง ผลลัพธ์จะได้ว่า

- Max pooling : เลือกค่าสูงสุดใน pool นั่นคือ 4
- Average pooling: หาค่าเฉลี่ยใน pool นั่นคือ  $(1+3+4+1)/4 = 2.25$

จุดประสงค์ของการทำ Pooling คือ เพื่อลดขนาดของข้อมูล ทำให้โมเดลไม่จำเป็นต้องเรียนรู้ค่า weight จำนวนมากเกินไป

### 2.5.3 Padding [8]

กระบวนการ Convolution มักจะทำให้ Matrix ผลลัพธ์มีมิติเล็กลง ซึ่งถ้าหากเราทำ Convolution หลายๆ ชั้น ภาพสุดท้ายที่ออกมา ก็จะเล็กลงมาก นอกจากนั้น Convolution ยังมีแนวโน้มที่จะทำให้ข้อมูลที่อยู่ตามขอบของภาพนั้นไม่ถูกนำไปคำนวณอย่างเต็มที่เหมือนข้อมูลที่อยู่ส่วนกลางของภาพ เพราะ Filter มีโอกาสจับข้อมูลตามขอบภาพน้อยกว่าตรงกลางภาพ ทางออกต่อปัญหา 2 ข้อนี้ คือการทำ Padding ซึ่งก็คือการขยายขอบของข้อมูล Input ออกทุกด้านเท่าๆ กัน

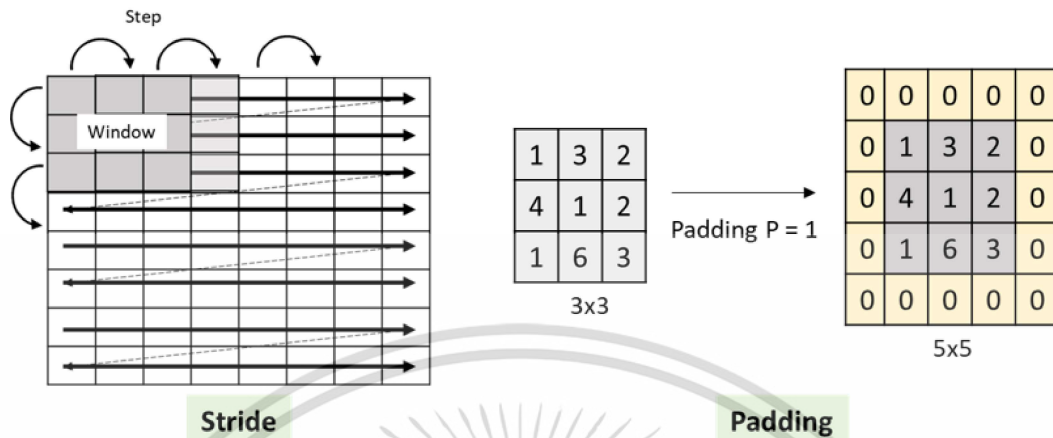
ตัวอย่าง กำหนด Input ขนาด 3x3 [7]

การตั้งค่า Padding = 1 หมายความว่า เราต้องการขยาย Input จาก 3x3 เป็น 5x5 โดยการขยายใช้วิธีเติมค่า 0 ล้อมรอบ Input เดิม เป้าหมายของการกำหนด padding คือ การให้ความสำคัญกับข้อมูลตรงขอบของ Input

### 2.5.4 Stride

Stride แปลว่า "ย่างก้าว" หมายถึงจำนวนช่องที่จะเลื่อนไปในกระบวนการ Convolution แต่ละครั้ง เช่นถ้าเลื่อนทีละช่องการตั้งค่า Stride = 1 หมายความว่า เราจะให้ weight สแกน Input ทีละ 1 step จากซ้ายไปขวา และบนลงล่าง แต่ถ้าเลื่อนทีละ 2 ช่อง ก็คือ Stride = 2 เป้าหมายของ Stride คือ กำหนดความละเอียดของการสแกน ยิ่งสลับการสแกนน้อย คือ ต้องการความละเอียดมาก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นอนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

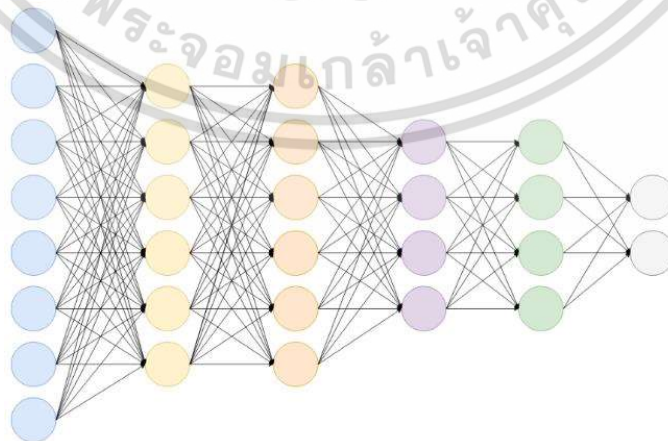


รูปที่ 2.16 พารามิเตอร์ที่ต้องกำหนดใน Convolutional layer

(ที่มา : <https://medium.com/@hadee2531earvesdrop/convolutional-neural-network-%E0%B8%84%E0%B8%B7%E0%B8%AD%E0%B8%AD%E0%B8%B0%E0%B9%84%E0%B8%A3-42c45f7ec16b>)

### 2.5.5 Fully Connected Layer [9]

หลังจากการประกอบกันของชั้นคอนโวลูชันและชั้นการรวมจำนวนหนึ่งแล้ว ในขั้นสุดท้ายของโครงข่ายประสาทเทียมคอนโวลูชัน จะเป็นการเชื่อมโยงเต็มรูปแบบ นั่นคือ ในขั้นนี้จะประกอบด้วยชั้นย่อยๆ ที่มีนิวรอลอยู่จำนวนหนึ่ง โดยที่นิวรอลแต่ละตัวจะมีเส้นเชื่อมกับนิวรอลทุกตัวในชั้นก่อนหน้าและนิวรอลทุกตัวในชั้นถัดไป ทำให้สามารถทำการคำนวณการป้อนไปข้างหน้าและการแพร่กระจายย้อนกลับได้ด้วยวิธีการปกติได้ชั้นการเชื่อมโยงเต็มรูปแบบดังรูปที่ 2.17



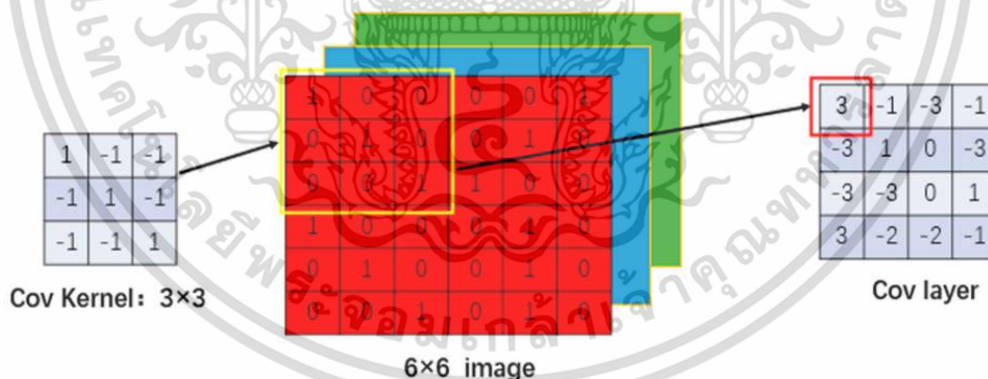
รูปที่ 2.17 Fully connected layer.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(ที่มา : <https://towardsdatascience.com/convolutional-neural-network-17fb77e76c05>)

## 2.6 โครงข่ายประสาทคอนโวลูชันแบบ 2 มิติ (2D CNN) [5]

โครงข่ายประสาทคอนโวลูชันประกอบด้วยชั้น Input ชั้น Output และชั้น hidden หลายชั้น แต่ละชั้น hidden จะฟังก์ชันเฉพาะ เช่น convolution, pooling, และ กระตุ้น (activation) ชั้น input เชื่อมต่อกับ input รูปภาพ และจำนวนเซลล์ประสาทในชั้นนี้คือ พิกเซล (pixel) ของภาพที่ป้อนเข้า ชั้น convolutional บริเวณตรงกลางทำการแยกคุณสมบัติข้อมูลที่ป้อนเข้าผ่านการ convolution เพื่อให้ได้ฟังก์ชันลักษณะ (feature map) ผลของการ convolution ขึ้นอยู่กับการตั้งค่าของพารามิเตอร์คุณสมบัติของ convolution ชั้น pooling ที่อยู่ถัดจากชั้น convolutional ทำให้ filter และเลือกฟังก์ชันลักษณะง่ายขึ้นจากความซับซ้อนในการคำนวณจากโครงข่ายทั้งหมด ผ่านชั้นที่เชื่อมต่ออย่างสมบูรณ์ (fully connect layer) เซลล์ประสาททั้งหมดในชั้นก่อนหน้ามีการเชื่อมต่ออย่างสมบูรณ์ ค่า output ที่ได้จะถูกส่งไปจำแนก เพื่อให้ได้ผลการจำแนก โครงข่ายประสาทเทียมทั่วไปคือ 2D CNN ภาพ input จะเป็นแบบ 2 มิติ และ convolution kernel แบบ 2 มิติ เช่น ResNet, VGG เป็นต้น สมมติว่าขนาดภาพที่ป้อนคือ สูง x กว้าง (H x W) แบบโหมดอาจีบี (convolution kernel จะมีขนาด (c, h, w) เลื่อนบนพื้นที่มิติของภาพ Input โดยที่ c, h, w หมายถึงจำนวนช่อง ความสูง และ ความกว้างของ convolution kernel ตามลำดับ ค่าของภาพและค่าของ (h, w) ถูกป้อนในแต่ละช่องเพื่อดำเนินการ convolution เพื่อให้ได้ค่าเดอะ กระบวนการของ 2D CNN convolution แสดงในรูปที่ 2.18

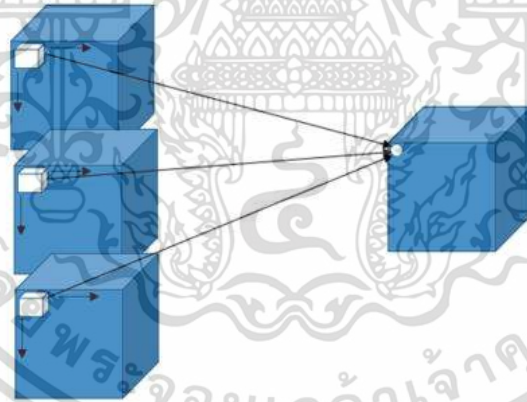


รูปที่ 2.18 โครงข่ายประสาทคอนโวลูชันแบบ 2 มิติ (2D CNN)

(ที่มา : <https://www.mdpi.com/2071-1050/13/3/1224>)

## 2.7 โครงข่ายประสาทคอนโวลูชันแบบ 3 มิติ (3D CNN)

ภาพทางการแพทย์ส่วนใหญ่มักจะเป็นภาพ 3 มิติ เช่น เอกซเรย์คอมพิวเตอร์ (CT) และ เครื่องสร้างภาพด้วยสนามแม่เหล็กไฟฟ้า (MRI) ถึงแม้ว่าภาพ CT ที่เราเห็นมักจะเป็นภาพ 2 มิติ ซึ่งเป็นภาพสไลด์ส่วนหนึ่งของวอลูมเท่านั้น ดังนั้นหากต้องการตัดแบ่งส่วนภาพ (segmentation) เนื้อเยื่อที่เป็นโรค จะต้องใช้ 3D convolution kernel ตัวอย่างเช่น convolution kernel ที่ใช้โครงข่ายสถาปัตยกรรมยูเน็ตแบบสามมิติ (3D U-Net) ในการตัดแบ่งส่วนภาพสามมิติ โดยการเปลี่ยน convolution kernel แบบสองมิติในโครงข่ายสถาปัตยกรรมยูเน็ตให้เป็น convolution kernel แบบสามมิติ ซึ่งเหมาะสมสำหรับการตัดแบ่งส่วนภาพทางการแพทย์แบบสามมิติ โครงข่ายประสาทคอนโวลูชันแบบ 3 มิติ สามารถตัดแบ่งข้อมูลแบบวอลูมที่มีขนาดใหญ่ได้ ที่แสดงภาพตัดขวางแบบสามแกน X, Y และ Z การใช้ข้อมูล 3 มิติในการตัดแบ่งส่วนภาพ จะใช้ประโยชน์จากข้อได้เปรียบของข้อมูลเชิงพื้นที่อย่างเต็มที่ convolution kernel แบบ 3 มิติ จะมีมิติความหนาเพิ่มขึ้นมาจาก kernel 2 มิติ ซึ่งคือจำนวนสไลด์ของภาพทางการแพทย์แบบ 2 มิติ โดยกำหนดให้ภาพแบบ 3 มิติเป็น  $C \times N \times H \times W$  โดยที่ C, N, H และ W การแทนค่าของ จำนวน channel จำนวนชั้นของสไลซ์ ความสูงและความกว้างของ convolution kernel เช่นเดียวกับการดำเนินการคอลโวลูชันแบบ 2 มิติ ค่าที่ได้จะได้รับการเลื่อนบล็อกตามความสูง ความกว้าง และจำนวนชั้นในแต่ละ 채널 กระบวนการของโครงข่ายประสาทคอนโวลูชันแบบ 3 มิติ ดังแสดงในรูปที่ 2.19



รูปที่ 2.19 โครงข่ายประสาทคอนโวลูชันแบบ 3 มิติ (3D CNN)

(ที่มา : <https://www.mdpi.com/2071-1050/13/3/1224>)

## 2.8 การแบ่งส่วนภาพ (Image Segmentation)

การแบ่งส่วนภาพเป็นปัญหาคลาสสิกในการวิจัยการมองเห็นด้วยคอมพิวเตอร์ (computer vision) และได้กลายเป็นฮอตสปอตในด้านการเข้าใจของภาพ การแบ่งส่วนภาพที่เรียกกันจะหมายถึง การแบ่งภาพออกเป็นหลายส่วนที่ไม่ปะติดปะต่อกัน ตามคุณลักษณะต่างๆ เช่น ระดับสีเทา (grayscale) สี (color) พื้นผิว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

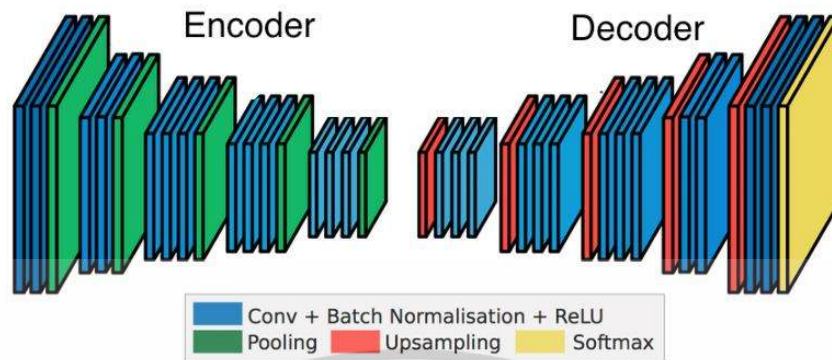
เชิงพื้นที่และรูปร่างเรขาคณิต เพื่อให้คุณลักษณะเหล่านี้แสดงความสอดคล้องหรือคล้ายคลึงกันในพื้นที่เดียวกัน แต่ความแตกต่างระหว่างพื้นที่แสดงให้เห็นความแตกต่างอย่างชัดเจน การแบ่งส่วนภาพ แบ่งออกเป็น semantic segmentation, instance segmentation และ panoramic segmentation ตามการแบ่งส่วนหยาบและส่วนละเอียดที่แตกต่างกัน การแบ่งส่วนภาพทางการแพทย์ถือเป็นงานแบ่งส่วนแบบ semantic segmentation ปัจจุบันได้มีการพัฒนาด้านการวิจัยการแบ่งส่วนภาพมากยิ่งขึ้น เช่น การแบ่งส่วนภาพถ่ายดาวเทียม การแบ่งส่วนภาพทางการแพทย์ การขับขี่แบบอิสระ เป็นต้น ด้วยรูปแบบของโครงสร้างโครงข่ายที่เพิ่มมากขึ้น วิธีการแบ่งส่วนภาพได้รับการปรับปรุงไปละขั้นตอนเพื่อให้ได้ผลการแบ่งส่วนที่แม่นยำมากยิ่งขึ้น อย่างไรก็ตาม สำหรับตัวอย่างการแบ่งส่วนที่แตกต่างกัน ยังไม่มีอัลกอริทึมการแบ่งส่วนแบบตายตัวที่จะเหมาะสมสำหรับทุกภาพ

## 2.9 โครงข่ายประสาทเทียมสำหรับการแบ่งส่วน

โครงข่ายการแบ่งส่วนก็มีการเปลี่ยนแปลงในโครงสร้างโครงข่ายประสาทเทียมแบบทั่วไป ครั้งแรกของโครงข่ายการแบ่งส่วนคือการเปลี่ยนชั้นที่เชื่อมต่ออย่างสมบูรณ์ (fully connected) ในสองชั้นสุดท้ายสำหรับโครงข่ายการจำแนกประเภทเป็นชั้นคอนโวลูชัน แนวทางหลักของโครงข่าย (backbone) การแบ่งส่วนภาพทางการแพทย์ขึ้นอยู่กับโครงสร้างเชิงลึกเช่น VGG และ ResNet ตลอดจนโครงสร้างตัวเข้ารหัส (encoder) – ตัวถอดรหัส (decoder) LeNet และ AlexNet เป็นโครงข่ายรุ่นแรกๆ โครงสร้างโครงข่ายทั้งสองค่อนข้างคล้ายกันและอยู่ในโครงข่ายที่ตื้น AlexNet มีจำนวนพารามิเตอร์ที่มากกว่าโครงข่าย LeNet แนวคิดในการเพิ่มขึ้นหลังจากชั้นคอนโวลูชัน คือยังคงเป็นที่นิยมในขณะนี้ การพัฒนาโครงข่าย VGG ให้เหนือกว่า AlexNet คือการเพิ่มจำนวนชั้นโครงข่ายให้ลึกขึ้น โดยเพิ่ม  $3 \times 3$  convolution kernel หลายชั้นติดต่อกันเพื่อแทนที่ convolution kernel ให้ใหญ่กว่า convolution kernel ใน AlexNet ภายใต้เงื่อนไขของการรองรับขอบเขตเดียวกัน ความลึกของโครงข่ายและเอฟเฟกต์ของการดึงคุณสมบัติขั้นสูง โครงสร้างของ VGG นั้นจะมีเรียบง่ายและไม่ยุ่งยากกว่าโครงข่ายทั้งหมด ใช้ convolution kernel ขนาดเดียวกันและขนาด pooling สูงสุด มาตรวจสอบว่าประสิทธิภาพสามารถพัฒนาได้อย่างต่อเนื่องเจาะลึก โครงสร้างโครงข่าย โครงข่ายทั้งหมดที่กล่าวมาข้างต้นได้รับการฝึกอบรมที่ดีขึ้น ผลกระทบโดยการเพิ่มจำนวนชั้นโครงข่าย แต่สิ่งนี้อาจทำให้เกิดปัญหาได้เช่นกันคือ overfitting และ vanishing gradients

### 2.10 ตัวเข้ารหัส - ตัวถอดรหัส (encoder – decoder) [10]

จากรูปที่ 2.20 สิ่งนี้เรียกว่าโครงสร้างตัวเข้ารหัส-ตัวถอดรหัส โดยชั้นที่สุ่มตัวอย่างนำเข้าไปเป็นส่วนหนึ่งของตัวเข้ารหัสและชั้นที่สุ่มตัวอย่างเป็นส่วนหนึ่งของตัวถอดรหัส เมื่อโมเดลได้รับการฝึกฝนสำหรับการแบ่งภาพ ตัวเข้ารหัสจะส่งออกเทนเซอร์ที่มีข้อมูลเกี่ยวกับวัตถุ ตลอดจนรูปร่างและขนาดของวัตถุ ตัวถอดรหัสจะใช้ข้อมูลนี้และแมพ (map) การแบ่งส่วนขึ้นมา

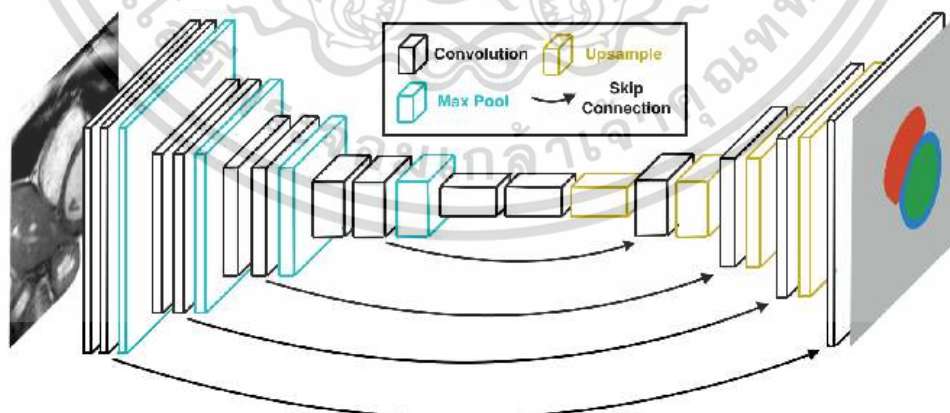


รูปที่ 2.20 สถาปัตยกรรมตัวเข้ารหัส-ตัวถอดรหัส

(ที่มา : <https://medium.com/swlh/image-segmentation-using-deep-learning-a-survey-e37e0f0a1489>)

### 2.11 การเชื่อมต่อตรงข้าม (Skip connections)

หากวางชั้นตัวเข้ารหัสและตัวถอดรหัสซ้อนกันข้อมูลในระดับต่ำอาจสูญหายไป ดังนั้นขอบเขตในแมพการแบ่งส่วนที่สร้างขึ้นโดยตัวถอดรหัสอาจไม่ถูกต้อง เพื่อที่จะชดเชยข้อมูลที่หายไปเราอนุญาตให้ตัวถอดรหัสเข้าถึงคุณสมบัติระดับต่ำที่สร้างโดยชั้นตัวเข้ารหัส ซึ่งทำได้โดยการเชื่อมต่อตรงข้ามการนำออกระดับกลางของตัวเข้ารหัสถูกเพิ่มเชื่อมต่อกับข้อมูลนำเข้าไปยังชั้นระดับกลางของตัวถอดรหัสในตำแหน่งที่เหมาะสม การเชื่อมต่อตรงข้ามจากชั้นก่อนหน้าจะให้ข้อมูลที่จำเป็นแก่ชั้นตัวถอดรหัสซึ่งจำเป็นสำหรับการสร้างขอบเขตที่ต้องการ



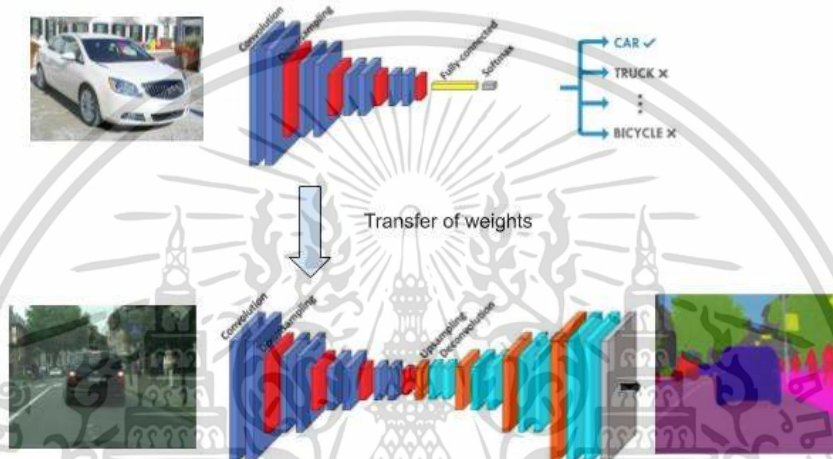
รูปที่ 2.21 การเชื่อมต่อตรงข้าม (Skip connections)

(ที่มา : <https://medium.com/swlh/image-segmentation-using-deep-learning-a-survey-e37e0f0a1489>)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 2.12 การถ่ายทอดการเรียนรู้ (Transfer learning)

แบบจำลองของโครงข่ายประสาทเทียมที่ได้รับการฝึกฝนสำหรับการจำแนกภาพ ประกอบด้วยข้อมูลที่มีความหมายซึ่งสามารถใช้สำหรับการแบ่งส่วนได้เช่นกัน เราสามารถใช้ชั้นคอนโวลูชันของโมเดลที่ผ่านการฝึกอบรมมาแล้วซ้ำในชั้นตัวเข้ารหัสของโมเดลการแบ่งส่วน การใช้ Resnet หรือ VGG ที่ได้รับการฝึกอบรมล่วงหน้าในชุดข้อมูล ImageNet เป็นทางเลือกยอดนิยม



รูปที่ 2.22 การถ่ายทอดการเรียนรู้ (Transfer Learning)

(ที่มา : <https://medium.com/swlh/image-segmentation-using-deep-learning-a-survey-e37e0f0a1489>)

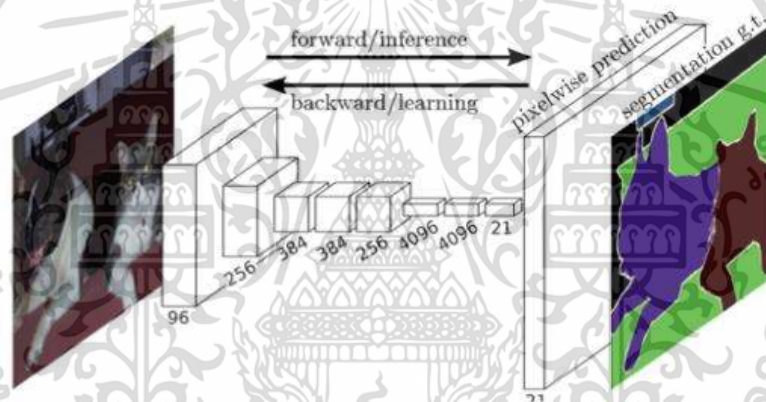
## 2.13 สถาปัตยกรรมการเรียนรู้เชิงลึกสำหรับการแบ่งส่วนภาพ

สถาปัตยกรรมการเรียนรู้เชิงลึกสำหรับการแบ่งส่วนมีมากกว่าร้อยวิธีมากกว่าวิธีการ ตั้งแต่แรกเริ่มจนถึงปี 2019 โดยจัดกลุ่มเป็น 10 หมวดหมู่ (ขึ้นอยู่กับสถาปัตยกรรมที่จำลอง) บางสถาปัตยกรรมที่พบได้ทั่วไปในหลายๆ งาน เช่น มีส่วนเข้ารหัสและส่วนถอดรหัส, การเชื่อมต่อตรงข้าม, การวิเคราะห์หลายมาตราส่วนและการใช้คอนโวลูชันเมื่อไม่นานมานี้ ด้วยเหตุนี้จึงเป็นเรื่องยากที่จะกล่าวถึงผลงานแต่ละขั้นที่ไม่ซ้ำกัน แต่ง่ายสำหรับการจัดกลุ่มขึ้นอยู่กับผลงานทางสถาปัตยกรรมพื้นฐาน จากผลงานที่ผ่านมานอกจากการจัดหมวดหมู่ทางสถาปัตยกรรมของโมเดลเหล่านี้ เราสามารถจัดกลุ่มตามเป้าหมายการแบ่งกลุ่มออกเป็น ความหมาย, อินสแตน, พาโนปิก และหมวดหมู่การแบ่งส่วนเชิงลึก แต่เนื่องจากความแตกต่างอย่างมาก ในแง่ของปริมาณงานในงานเหล่านี้จึงต้องแบ่งตามสถาปัตยกรรมสำหรับการแบ่งกลุ่ม

สถาปัตยกรรมสำหรับการแบ่งส่วนภาพด้วยการเรียนรู้เชิงลึกได้รวบรวมมาดังนี้ [11]

### 2.13.1 Fully Convolutional networks

Fully Convolutional Network (FCN) ประกอบด้วยชั้นคอนโวลูชันเท่านั้นที่คุณสมบัติต่างๆ จะถูกดึงออกมาโดยการรวมคอร์เนลหรือตัวกรองของตัวถ่วงน้ำหนัก ที่ถ่ายภาพขนาดใดก็ได้และสร้างแมปการแบ่งส่วนที่มีขนาดเท่ากัน โดยใช้การเชื่อมต่อตรงข้ามซึ่งส่งต่อให้พีเจอร์แมปในชั้นสุดท้าย สามารถเพิ่มขนาดและรวมเข้ากับพีเจอร์แมปของชั้นก่อนหน้าได้ สิ่งนี้ช่วยให้โมเดลสร้างการแบ่งส่วนที่ถูกต้องและลงรายละเอียดมากยิ่งขึ้น ด้วยการรวมข้อมูลมุ่งมาจากชั้นที่ลึกและหยาบเข้ากับข้อมูลลักษณะที่ปรากฏจากชั้นที่ตื้นและละเอียด โมเดลได้รับการทดสอบบนชุดข้อมูล เช่น PASCAL VOC, NYUDv2 และ SIFT Flow เมื่อเทียบกับงานวิจัยที่ผ่านมา แอปพลิเคชันที่โดดเด่นของ FCN คือการแบ่งส่วนม่านตาและการแบ่งส่วนเนื้องอกในสมอง



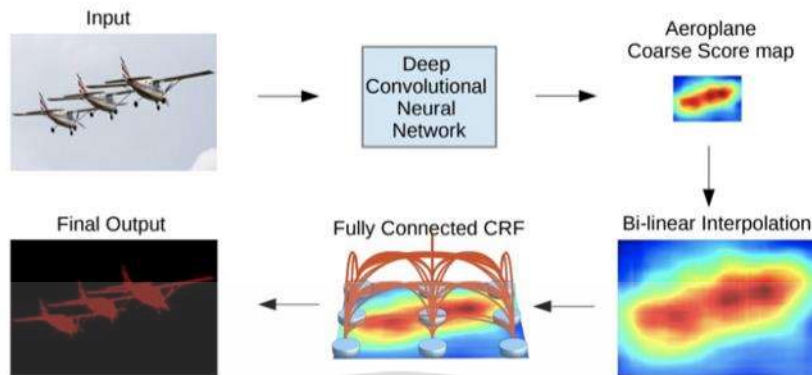
รูปที่ 2.23 Fully Convolutional Network (FCN)

(ที่มา : <https://medium.com/swlh/image-segmentation-using-deep-learning-a-survey-e37e0f0a1489>)

### 2.13.2 คอนโวลูชันโมเดลกับกราฟฟิกโมเดล

จากการวิจัยพบว่าโครงข่ายคอนโวลูชันเชิงลึกมีคุณสมบัติการจำกัดขอบเขตที่ไม่ดี ซึ่งหมายความว่า การตอบสนองที่คอนโวลูชันชั้นสุดท้ายนั้นได้รับการจำกัดขอบเขตไม่เพียงพอที่จะสร้างการแบ่งส่วนของวัตถุได้ไม่แม่นยำ การตอบสนองที่คอนโวลูชันชั้นสุดท้ายจึงถูกเชื่อมต่อแบบเต็มเข้ากับ Conditional Random Field (CRF) ซึ่งได้รับอัตราความแม่นยำสูงกว่าวิธี FCN ก่อนหน้านี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.24 คอนโวลูชันโมเดลกับกราฟิกโมเดล

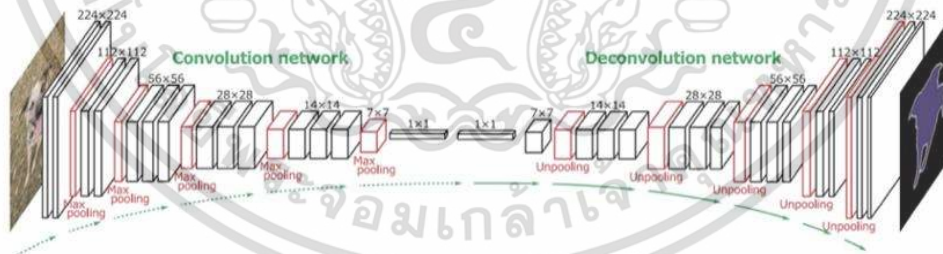
(ที่มา : <https://medium.com/swlh/image-segmentation-using-deep-learning-a-survey-e37e0f0a1489>)

### 2.13.3 โมเดลที่ใช้ตัวเข้ารหัส - ตัวถอดรหัส (encoder – decoder models)

โมเดลที่ใช้ตัวเข้ารหัส-ตัวถอดรหัสสามารถแบ่งออกได้เป็นสองประเภท:

(1) โมเดลตัวเข้ารหัส-ตัวถอดรหัสสำหรับการแบ่งส่วนทั่วไป

ประกอบด้วยสองส่วนคือ ส่วนเข้ารหัส และส่วนถอดรหัส ตัวเข้ารหัสใช้ชั้นคอนโวลูชันในขณะที่ตัวถอดรหัสใช้โครงข่ายดีคอนโวลูชัน (deconvolutional) ซึ่งสร้างแผนที่ของความน่าจะเป็นของคลาสที่พิกเซลไวส ตามการนำเข้าคุณบลักษณะเวกเตอร์ รุ่นยอดนิยมในหมวดนี้ ได้แก่ SegNet และ HRNet



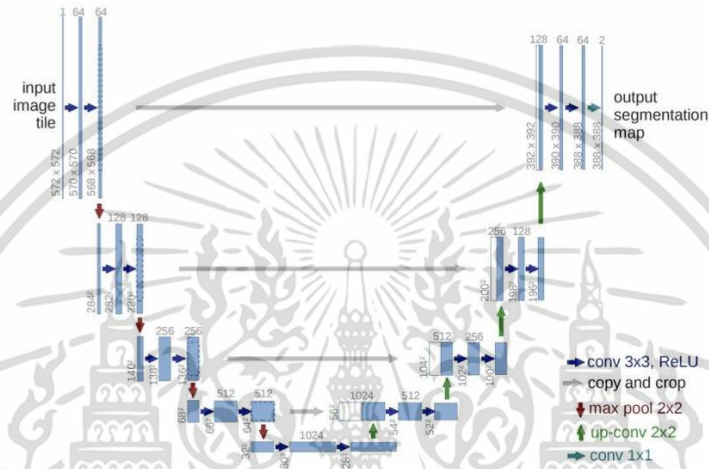
รูปที่ 2.25 โมเดลที่ใช้ตัวเข้ารหัส - ตัวถอดรหัส (encoder – decoder models)

(ที่มา : <https://medium.com/swlh/image-segmentation-using-deep-learning-a-survey-e37e0f0a1489>)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## (2) โมเดลตัวเข้ารหัส-ตัวถอดรหัสสำหรับการแบ่งส่วนภาพทางการแพทย์และชีวการแพทย์

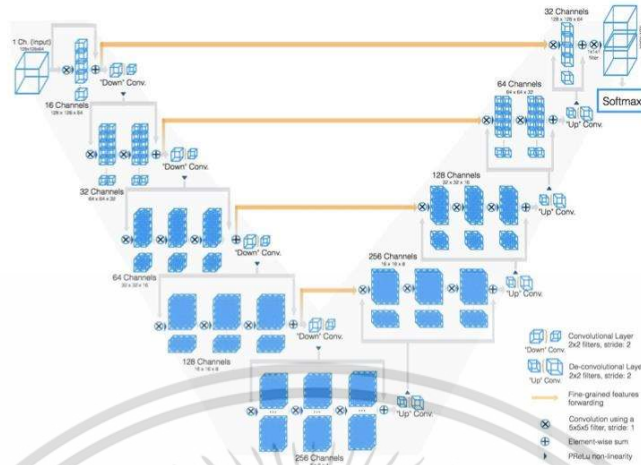
ยูเน็ต (U-Net) และ วีเน็ต (V-Net) เป็นสองสถาปัตยกรรมที่ได้รับความนิยมสูงสุดที่ใช้ในการแบ่งส่วนภาพทางการแพทย์และชีวการแพทย์ โดยทั่วไปแล้วยูเน็ตใช้สำหรับการแบ่งส่วนของภาพจุลทรรศน์ชีวภาพ ใช้เทคนิคการเสริมข้อมูล (data augmentation) เพื่อเรียนรู้จากภาพที่มีคำอธิบายประกอบ สถาปัตยกรรมยูเน็ตประกอบด้วยสองส่วน คือส่วนที่ย่อ (contracting) และส่วนขยาย (symmetric) แบบสมมาตร สำหรับการจับภาพบริบทและเปิดใช้ในงานการแปลที่แม่นยำตามลำดับ



รูปที่ 2.26 โครงสร้างสถาปัตยกรรมยูเน็ต (U-Net)

(ที่มา : <https://medium.com/swlh/image-segmentation-using-deep-learning-a-survey-e37e0f0a1489>)

วีเน็ตเป็นอีกหนึ่งโมเดลยอดนิยมที่ใช้สำหรับการแบ่งส่วนภาพทางการแพทย์แบบ 3 มิติ โดยใช้ฟังก์ชันวัดดูประสงคใหม่สำหรับการฝึกอบรมโมเดล ซึ่งขึ้นอยู่กับค่าสัมประสิทธิ์ไดส์ (Dice coefficient) โมเดลวีเน็ตได้รับการฝึกอบรมกับภาพเอ็มอาร์ไอโวลุ่ม และคาดเดา (predicts) การแบ่งส่วนสำหรับเอ็มอาร์ไอโวลุ่มทั้งหมดในคราวเดียว

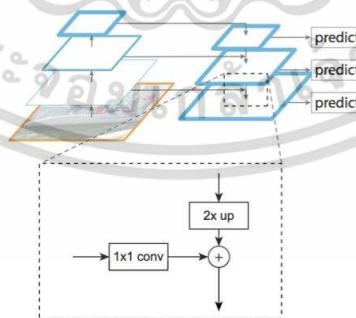


รูปที่ 2.27 โครงสร้างสถาปัตยกรรมวีเน็ต (V-Net)

(ที่มา : <https://medium.com/swlh/image-segmentation-using-deep-learning-a-survey-e37e0f0a1489>)

#### 2.13.4 โมเดลตามโครงข่ายแบบหลายสเกลและพีระมิด

คุณสมบัติโครงข่ายแบบหลายสเกลและพีระมิด (FPN) เป็นโมเดลรุ่นที่ได้รับความนิยมสูงสุดในหมวดนี้ เริ่มแรกได้รับการพัฒนาสำหรับการตรวจจับวัตถุ ต่อมาใช้สำหรับการแบ่งส่วนภาพเช่นกัน ใช้คุณสมบัติสร้างพีระมิดต่างๆ และใช้แนวทางจากล่างขึ้นบน แนวทางจากบนลงล่าง และการเชื่อมต่อด้านข้างเพื่อรวมคุณลักษณะที่มีความละเอียดต่ำและสูงเข้าด้วยกัน จากนั้นจะใช้คอนโวลูชันขนาด  $3 \times 3$  บนการรวมพีระจอร์แมป เพื่อสร้างผลลัพธ์ของแต่ละขั้นตอน สุดท้ายแต่ละชั้นของแนวทางจากบนลงล่างจะสร้างการคาดเดาเพื่อตรวจจับวัตถุ สำหรับการแบ่งส่วนภาพ ผู้เขียนใช้โครงข่ายแบบหลายชั้น (MLP) สองตัวเพื่อสร้างมาสก์



รูปที่ 2.28 โมเดลตามโครงข่ายแบบหลายสเกลและพีระมิด

(ที่มา : <https://medium.com/swlh/image-segmentation-using-deep-learning-a-survey-e37e0f0a1489>)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 2.13.5 โมเดลอาร์ซีเอ็นเอ็น (R-CNN)

โครงข่ายคอนโวลูชันระดับภูมิภาค (RCN) สำหรับการแบ่งส่วนภาพขั้นสูง (สำหรับการแบ่งส่วนอินสแตนซ์) เป็นโมเดลที่ได้รับความนิยมอย่างมากในการแก้ปัญหาการแบ่งส่วนอินสแตนซ์ จะทำงานการตรวจจับวัตถุและการแบ่งส่วนแบบซีเมนติกพร้อมกัน ส่วนขยายที่เร็วมากของอาร์ซีเอ็นเอ็นใช้โครงข่ายข้อเสนอระดับภูมิภาค (RPN) เพื่อแยกพื้นที่ที่น่าสนใจ (RoI) จากนั้นใช้ชั้น RoIPool สำหรับการคำนวณคุณสมบัติจากข้อเสนอเหล่านี้ และสรุปพิกัดขอบเขตและคลาสของวัตถุ

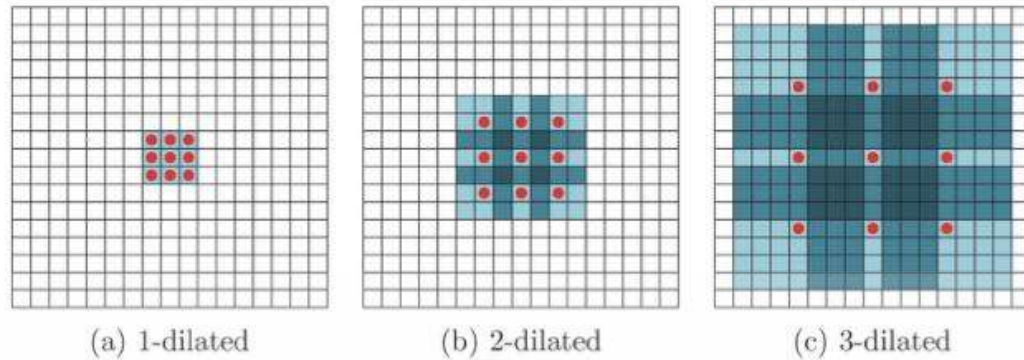


รูปที่ 2.29 โมเดลอาร์ซีเอ็นเอ็น (R-CNN)

(ที่มา : <https://medium.com/swlh/image-segmentation-using-deep-learning-a-survey-e37e0f0a1489>)

### 2.13.6 โมเดลคอนโวลูชันแบบขยายและตระกูล DeepLab

โมเดลคอนโวลูชันแบบขยายพารามิเตอร์ที่เพิ่มเติมจะถูกเพิ่มไปยังชั้นคอนโวลูชัน ซึ่งเรียกว่าอัตราขยาย ซึ่งกำหนดระยะห่างระหว่างค่าถ่วงน้ำหนักของคอร์เนล เป็นที่นิยมมากสำหรับการแบ่งส่วนตามเวลาจริง



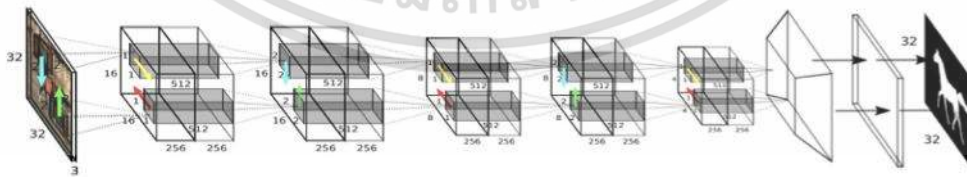
รูปที่ 2.30 โมเดลคอนโวลูชันแบบขยายและตระกูล DeepLab

(ที่มา : <https://medium.com/swlh/image-segmentation-using-deep-learning-a-survey-e37e0f0a1489>)

ในบรรดาตระกูล DeepLab นั้น DeepLab v1, DeepLab v2 และ DeepLab v3 เป็นโมเดลที่ล้ำสมัยสำหรับวิธีการแบ่งส่วนภาพ โดย DeepLab v3+ เป็นรุ่นใหม่ที่ล้ำสุด DeepLab ที่ดีที่สุด โดยใช้ ResNet 101 เป็นแนวทางหลักทำให้ค่าเฉลี่ย IoU ได้มากถึง 70.4% ในการแข่งขัน Cityscapes

### 2.13.7 โมเดลตามโครงข่ายประสาทเทียมที่กัดซ้ำ (RNN)

นอกจากโครงข่ายประสาทเทียม แล้วโครงข่ายประสาทเทียมที่กัดซ้ำยังมีประโยชน์ในการแบ่งส่วนภาพอีกด้วย อาจมีปรับปรุงการประมาณของแมปการแบ่งส่วน โดยสร้างโมเดลการพึ่งพาระยะสั้นและระยะยาวระหว่างพิกเซล ReSeg เป็นโมเดลแล้วโครงข่ายประสาทเทียมที่กัดซ้ำที่ใช้รุ่นแรก สำหรับการแบ่งส่วนภาพและถูกพัฒนามาจาก ReNet ซึ่งใช้ในการจำแนกภาพ โมเดล ReSeg ใช้ชั้นแบบ ReNet ซึ่งวางซ้อนกันบนชั้นคอนโวลูชัน VGG-16 ก่อนการฝึกเทรน ซึ่งจะแยกคุณสมบัติแบบทั่วไปเพื่อทำการแบ่งส่วนภาพ เพื่อกู้คืนความละเอียดของภาพดั้งเดิมในการคาดเดาขั้นสุดท้ายของ ReNet จากนั้นจะถูกตามด้วยชั้นขยายโครงข่ายนี้ต้องใช้ Gated Recurrent Units (GRU) เนื่องจากให้ความสมดุลที่ดีระหว่างการใช้หน่วยความจำและพลังการคำนวณ

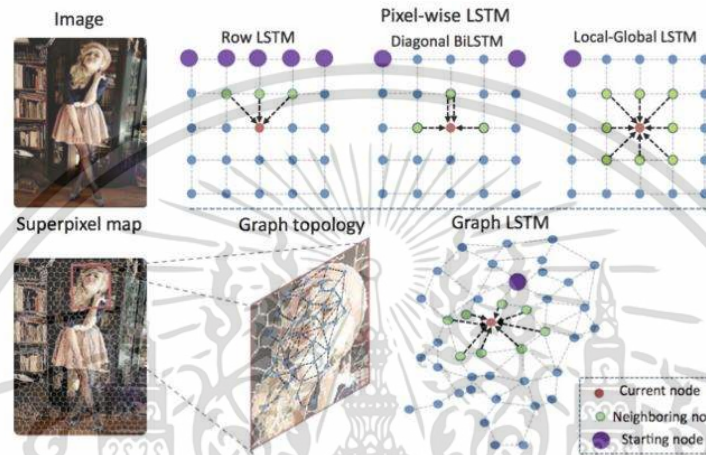


รูปที่ 2.31 โมเดลตามโครงข่ายประสาทเทียมที่กัดซ้ำ (RNN)

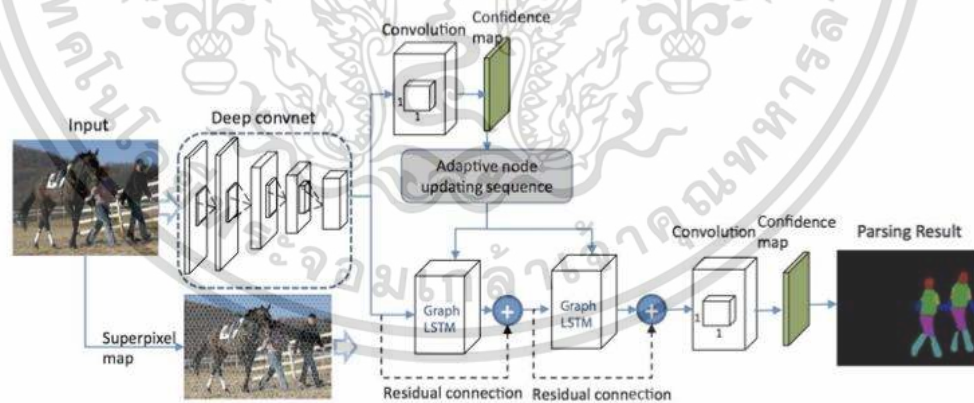
(ที่มา : <https://medium.com/swlh/image-segmentation-using-deep-learning-a-survey-e37e0f0a1489>)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ต่อมาโมเดลการแบ่งส่วนแบบซีเมนทริกอีกรูปแบบหนึ่งเกิดขึ้นจากโครงข่าย Graph LSTM (Graph Long Short-Term Memory) ซึ่งเป็นการทำให้ LSTM มีลักษณะทั่วไปจากข้อมูลมิติเดียวหรือหลายมิติไปจนถึงข้อมูลที่มีโครงสร้างเป็นกราฟทั่วไป ในกรณีนี้ได้ใช้ซูเปอร์พิกเซลที่มีรูปร่างตามอิสระแต่ละโหนดเป็นโหนดที่สอดคล้องกันทางซีเมนทริกและสร้างกราฟแบบไม่มีทิศทางสำหรับภาพ ซึ่งความสัมพันธ์เชิงพื้นที่ของซูเปอร์พิกเซลจะถูกใช้เป็นขอบโดยธรรมชาติ



รูปที่ 2.32 การเปรียบเทียบระหว่างโมเดลกราฟ-LSTM กับโมเดล RNN แบบใช้พิกเซลแบบดั้งเดิม (ที่มา : <https://medium.com/swlh/image-segmentation-using-deep-learning-a-survey-e37e0f0a1489>)



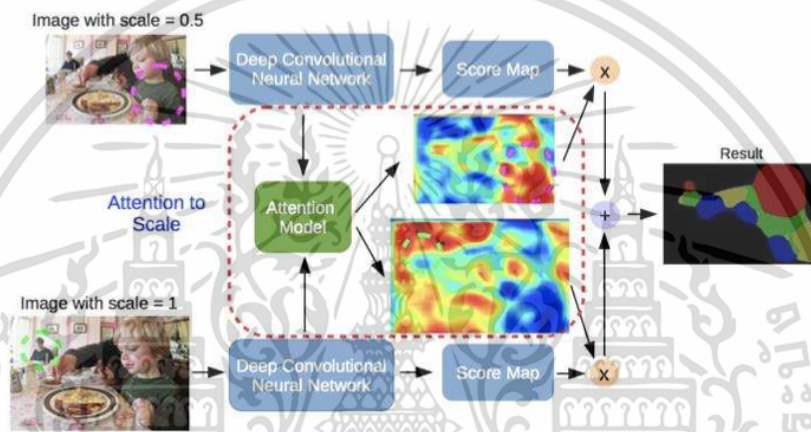
รูปที่ 2.33 แบบจำลองกราฟ LSTM สำหรับการแบ่งส่วน

(ที่มา : <https://medium.com/swlh/image-segmentation-using-deep-learning-a-survey-e37e0f0a1489>)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 2.13.8 โมเดลที่เน้นความสนใจ (Attention-Based Models)

กลไกที่เน้นความสนใจมีประสิทธิภาพดีการใช้แมคพลูแบบเฉลี่ยและแมคพลูสูงสุด และทำให้โมเดลสามารถประเมินความสำคัญของคุณลักษณะในตำแหน่งและสเกลที่แตกต่างได้ ความแตกต่างของโมเดลโครงข่ายประสาทเทียมตรงคอนโวลูชันชนิดแยกประเภทจะได้รับการฝึกให้เรียนรู้คุณลักษณะทางซีเมนทริกที่เป็นตัวอย่างของวัตถุ สถาปัตยกรรมโครงข่ายเน้นความสนใจย้อนกลับ (RAN) จะฝึกโมเดลเพื่อจับคุณลักษณะที่ไม่เกี่ยวข้องกับคลาสเป้าหมาย RAN เป็นโครงข่ายมีสามแขนงที่ดำเนินกระบวนการเรียนรู้โดยตรงและแบบย้อนกลับพร้อมกัน

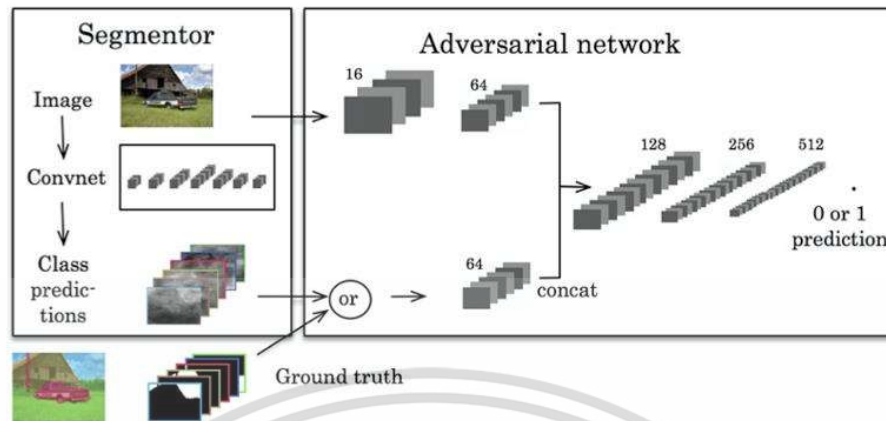


รูปที่ 2.34 โมเดลที่เน้นความสนใจ (Attention-Based Models)

(ที่มา : <https://medium.com/swlh/image-segmentation-using-deep-learning-a-survey-e37e0f0a1489>)

### 2.13.9 โมเดลต้นกำเนิดและการฝึกอบรมด้านตรงข้าม

ตั้งแต่ GAN ถือกำเนิดขึ้น โมเดลนี้ได้รับความนิยมอย่างมากและถูกใช้สำหรับงานแบ่งส่วนภาพด้วย ในแนวทางการฝึกอบรมด้านตรงข้าม โครงข่ายการแบ่งส่วนแบบซีเมนทริกคอนโวลูชัน ได้รับการฝึกอบรมพร้อมกับโครงข่ายด้านตรงข้ามที่นั้นแยกผลเฉลี่ยแมปการแบ่งส่วนจากการสร้างขึ้นโดยโครงข่ายการแบ่งส่วน วิธีการนี้ได้แสดงความแม่นยำที่ดีขึ้นในชุดข้อมูล Stanford Background และ PASCAL VOC 2012



รูปที่ 2.35 โมเดลต้นกำเนิดและการฝึกอบรมด้านตรงข้าม

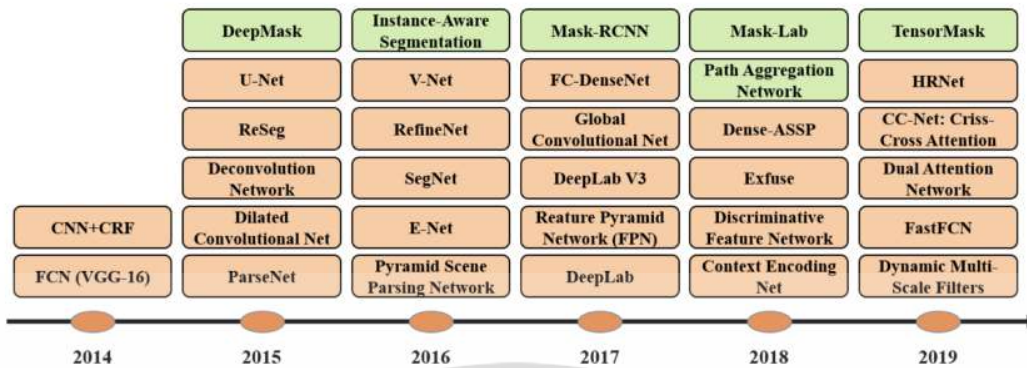
(ที่มา : <https://medium.com/swlh/image-segmentation-using-deep-learning-a-survey-e37e0f0a1489>)

#### 2.13.10 โมเดลโครงข่ายประสาทเทียมพร้อมโมเดลแอคทีฟคอนทัว

FCNs พร้อมกับโมเดลแอคทีฟคอนทัว (ACMs) เพิ่งได้รับความสนใจและเป็นการวิจัยที่กำลังดำเนินอยู่ แนวทางที่เกี่ยวข้องกับการกำหนดฟังก์ชันการสูญเสียใหม่ที่ได้รับแรงบันดาลใจจากหลักการ ACM ต่างๆ ในขณะที่แนวทางอื่นๆ ใช้ ACM เป็นเพียงตัวประมวลผลหลังของเอาต์พุตของ FCN และความพยายามหลายอย่างพยายามเรียนรู้ร่วมกันในระดับปานกลางโดยการฝึกอบรม FCN ล่วงหน้า ตัวอย่างหนึ่งคือ ACM post-processor สำหรับงานแบ่งส่วนความหมายของภาพธรรมชาติโดยที่ ACM ที่ตั้งค่าระดับไว้ถูกนำมาใช้เป็น RNN

#### 2.13.11 โมเดลอื่นๆ

นอกเหนือจากโมเดลที่กล่าวถึงข้างต้นแล้ว ยังมีโมเดลอื่นๆ ที่ใช้การเรียนรู้เชิงลึกสำหรับการแบ่งส่วนภาพ เช่น Context Encoding Network (EncNet), RefneNet, Seednet, Feedforward-Net เป็นต้น รูปที่ 2.36 ต่อไปนี้แสดงช่วงเวลาของงานที่ใช้การเรียนรู้เชิงลึกที่ได้รับความนิยมสูงสุดบางส่วนสำหรับการแบ่งกลุ่มซีเมนทริก ตลอดจนการแบ่งกลุ่มอินสแตนซ์ตั้งแต่ปี 2014



รูปที่ 2.36 ช่วงเวลาของอัลกอริทึมการแบ่งส่วนภาพ 2 มิติ

(ที่มา : <https://medium.com/swlh/image-segmentation-using-deep-learning-a-survey-e37e0f0a1489>)

## 2.14 วิธีการทางการแพทย์สำหรับการแบ่งส่วนภาพทางการแพทย์

กระบวนการแบ่งส่วนมีการใช้งานมากมายในด้านการแพทย์ มีวิธีการทางการแพทย์ที่มีมาแตกต่างกัน ก่อนเข้าสู่กระบวนการแบ่งส่วน วิธีการเหล่านี้จะกล่าวถึงดังนี้

### 2.14.1 Computed Tomography Scan (CT scan) [12]

CT Scan คือ การตรวจวินิจฉัยโรคด้วยเครื่องเอกซเรย์คอมพิวเตอร์ ซึ่งแพทย์จะฉายรังสีเอกซเรย์ตามร่างกายบริเวณที่ต้องการตรวจ แล้วใช้คอมพิวเตอร์สร้างเป็นภาพฉายลักษณะและอวัยวะภายในร่างกาย เพื่อประกอบการวินิจฉัยหาความผิดปกติของร่างกายต่อไป โดยวิธีการนี้จะได้ภาพที่มีความละเอียดสูงกว่าการเอกซเรย์แบบธรรมดา และสามารถใช้ตรวจอวัยวะภายในร่างกายได้เกือบทุกส่วน



รูปที่ 2.37 เครื่องเอกซเรย์คอมพิวเตอร์ (CT Scan)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(ที่มา : <https://www.pobpad.com/ct-scan>)

### ประเภทของ CT scan

- เอกซเรย์คอมพิวเตอร์แบบภาพตัดขวางพื้นฐาน (Conventional CT Scan) เครื่องเอกซเรย์คอมพิวเตอร์จะหมุนเป็นวงรอบตัวผู้ป่วย เพื่อให้รังสีเอกซเรย์ผ่านตัวผู้ป่วย 1 รอบ จะได้ภาพ 1 ภาพ โดยเตียงจะเลื่อนไปที่ละตำแหน่ง ภาพที่ได้จะถูกนำมาสร้างเป็นภาพตัดขวางของอวัยวะได้ที่ละภาพ
- เอกซเรย์คอมพิวเตอร์แบบเกลียว (Spiral/Helical CT Scan) เครื่องเอกซเรย์คอมพิวเตอร์จะหมุนเป็นวงอย่างต่อเนื่องรอบตัวผู้ป่วย เมื่อรังสีเอกซเรย์ผ่านตัวผู้ป่วยจะได้ภาพหลายภาพ ซึ่งใช้เวลาน้อยกว่าแบบพื้นฐาน และได้ภาพที่มีความแม่นยำสูงกว่า

### ข้อจำกัดในการทำ CT Scan

- การตั้งครรภ์หากไม่ใช่ในกรณีฉุกเฉินเร่งด่วน แพทย์จะให้ผู้ป่วยที่ตั้งครรภ์ใช้วิธีการอื่นในการตรวจวินิจฉัย เพื่อเลี่ยงไม่ให้รังสีเอกซเรย์ส่งผลกระทบต่อทารกในครรภ์ได้
- ในผู้ป่วยที่เป็นเด็ก ผู้ปกครองควรปรึกษาและรับคำแนะนำจากแพทย์ถึงการเตรียมตัวก่อนให้เด็กเข้าเครื่อง CT Scan หากเด็กยังเล็กมาก หรือรู้สึกตื่นกลัวมาก แพทย์จะให้ยาระงับประสาทเพื่อบรรเทาอาการก่อนทำ CT Scan
- ผู้ป่วยควรทำ CT Scan เมื่อมีเหตุจำเป็นอันสมควรเมื่อผู้ป่วยมีอาการป่วยที่ปรากฏเท่านั้น โดยจะไม่ใช้ CT Scan เพื่อการคัดกรองเบื้องต้น เพราะเพิ่มความเสี่ยงต่อปัญหาสุขภาพจากการได้รับรังสีเอกซเรย์ และเพิ่มความวิตกกังวลเกินจำเป็น

### ข้อดีของการทำ CT Scan

- สามารถสแกนตรวจอวัยวะในร่างกายส่วนใหญ่ได้
- ไม่ทำให้เกิดความเจ็บปวดขณะตรวจสแกน
- ภาพที่ได้มีรายละเอียดสูงกว่าการทำอัลตราซาวด์ ทำให้แพทย์สามารถวินิจฉัยได้แม่นยำขึ้น
- สามารถสแกนได้รวดเร็วกว่าเมื่อเปรียบเทียบกับสแกนแบบ MRI

### ข้อเสียของการทำ CT Scan

- อาจมีวัตถุแปลกปลอมที่รบกวนการแปลผลเอกซเรย์ เช่น เครื่องประดับต่าง ๆ
- ในขณะที่สแกน ต้องมีการกลั้นหายใจ ซึ่งผู้ป่วยบางรายไม่สามารถปฏิบัติได้
- ในการสแกนสมองด้วย CT Scan อาจถูกกระดูกส่วนกะโหลกศีรษะบัง ทำให้แปลผลคลาดเคลื่อนได้
- ในตำแหน่งอื่น ๆ ที่มีกระดูกอยู่จำนวนมาก เช่น บริเวณกระดูกสันหลัง อาจเกิดการบดบังอวัยวะส่วนที่ต้องการตรวจวินิจฉัย จึงทำให้ภาพที่ได้จาก CT Scan มีโอกาสแปลผลคลาดเคลื่อนได้
- CT Scan ใช้รังสีปริมาณมาก อาจเสี่ยงต่อการเกิดปัญหาสุขภาพจากการได้รับรังสีมากเกินไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 2.14.2 Magnetic resonance imaging (MRI) [13]

Magnetic Resonance Imaging หรือ MRI คือเครื่องสร้างภาพด้วยสนามแม่เหล็กไฟฟ้า ใช้ในการตรวจวินิจฉัยรอยโรคของผู้ป่วย เพื่อนำมาใช้ในการรักษาและติดตามผลการรักษา อาศัยคุณสมบัติความเป็นแม่เหล็กของไฮโดรเจนอะตอม (Hydrogen,H) ซึ่งเป็นส่วนประกอบหลักที่อยู่ภายในร่างกายมนุษย์ เช่น โมเลกุลของน้ำ (H<sub>2</sub>O) เป็นต้น เมื่อผู้รับการตรวจเข้าไปอยู่ภายใต้สนามแม่เหล็กไฟฟ้า เครื่องจะส่งสัญญาณคลื่นวิทยุที่มีความถี่จำเพาะ (Radiofrequency) เข้าไปกระตุ้นระบบอวัยวะที่จะตรวจ เมื่ออวัยวะนั้นๆ ถูกกระตุ้นจะมีการเปลี่ยนแปลงระดับพลังงานตามขบวนการทางฟิสิกส์ ที่เรียกว่า การกำทอน (Resonance) หลังจากหยุดกระตุ้นไฮโดรเจนอะตอมภายในร่างกายมีการคายพลังงาน จะมีอุปกรณ์รับสัญญาณที่ได้ออกมา จากนั้นแปลงเป็นสัญญาณภาพบนจอภาพ

เครื่อง MRI จะเป็นเทคโนโลยีทันสมัยและมีราคาแพง ในปัจจุบันมีการนำมาใช้อย่างแพร่หลาย เนื่องจากเป็นการตรวจวินิจฉัยที่ให้ความถูกต้องและแม่นยำสูง เนื่องจากให้ความแตกต่างของเนื้อเยื่อได้ดี ทำให้หลายระบบ สามารถใช้ตรวจได้ทุกระบบของร่างกาย โดยเฉพาะอย่างยิ่งในส่วนของระบบสมองและกระดูกสันหลัง การตรวจ MRI จะดีกว่าเมื่อเทียบกับการตรวจด้วยเอกซเรย์คอมพิวเตอร์ (Computed Tomography, CT) อีกทั้งมีเทคนิคการตรวจพิเศษหลายแบบ เช่น การตรวจสำหรับผู้ป่วยที่มีภาวะสมองขาดเลือดแบบเฉียบพลัน การตรวจหาระดับชีวเคมีเพื่อแยกชนิดของก้อนเนื้อและการตรวจหลอดเลือดทั่วร่างกาย เป็นต้น MRI ยังมีประโยชน์มากในการตรวจผู้ป่วยเด็ก เนื่องจากไม่มีอันตรายจากรังสี ที่เกิดจากการใช้เอกซเรย์

MRI  
3 tesla  
1.5 tesla



รูปที่ 2.38 เครื่องสร้างภาพด้วยสนามแม่เหล็กไฟฟ้า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(ที่มา : <https://www.rama.mahidol.ac.th/aimc/th/content/09122017-1414-th>)

### ข้อดี ของการตรวจด้วย เครื่องเอ็มอาร์ไอ [14]

- สามารถจำแนกคุณสมบัติของเนื้อเยื่อที่แตกต่างกันได้ชัดเจน ทำให้แพทย์วินิจฉัยโรคได้ถูกต้องแม่นยำมากยิ่งขึ้น
- ใช้ได้ดีกับส่วนที่เป็นเนื้อเยื่อ เช่น เนื้อเยื่อสมอง, เส้นประสาทไขสันหลัง, กล้ามเนื้อ เอ็นยึดกระดูก และกล้ามเนื้อ, เส้นเลือด, ไชกระดูก และข้อต่อกระดูกต่างๆ เป็นต้น
- สามารถตรวจหาสิ่งผิดปกติในร่างกายได้ตั้งแต่ระยะแรกๆ
- ตรวจได้ทุกระนาบโดยไม่ต้องขยับผู้ป่วย
- ไม่มีรังสีเอกซเรย์ที่เป็นอันตรายต่อร่างกาย

### ข้อพึงระวังก่อนเข้ารับการตรวจเอ็มอาร์ไอ

เนื่องจากห้องตรวจเอ็มอาร์ไอ จะมีสนามแม่เหล็กความเข้มสูงตลอดเวลา ซึ่งมีผลต่อการทำงานของเครื่องมือที่ไวต่อแม่เหล็ก และดึงดูดวัตถุที่เป็นโลหะทุกชนิดจึงมีข้อพึงระวังก่อนเข้ารับการตรวจด้วยเครื่องเอ็มอาร์ไอ ดังนี้

- ห้ามตรวจในผู้ป่วยที่มีการฝังโลหะในร่างกาย เช่น ผู้ป่วยที่ผ่าตัดติดคลิปอุดหลอดเลือดในโรคเส้นเลือดโป่งพอง, ผู้ป่วยใส่ลิ้นหัวใจเทียม หรือข้อเทียม, ผู้ป่วยที่ใส่เครื่องกระตุ้นหัวใจ, ผู้ป่วยที่ใส่ Stent หลอดเลือดหัวใจหรือผู้ป่วยที่ผ่าตัดใส่อวัยวะเทียมภายในหู เป็นต้น
- ควรหลีกเลี่ยงการตรวจเอ็มอาร์ไอ ในผู้ป่วยที่กลัวที่แคบ (Claustrophobic) เพราะเครื่องเอ็มอาร์ไอ มีลักษณะเป็นโพรงให้เพียงผู้ป่วยเคลื่อนเข้าไปได้
- ผู้ที่เข้ารับการตรวจเอ็มอาร์ไอ จะต้องนำโลหะต่างๆ ออกจากตัว เช่น ที่หนีบผม, ฟันปลอม, ต่างหู, เครื่องประดับ, บัตรเอทีเอ็ม หรือบัตรเครดิต กุญแจ หรือ นาฬิกา เป็นต้น
- ผู้ป่วยต้องถอดเหล็กดัดฟัน, ไม่ใช่ช้อยแฮนด์ และมาสคาร่า เพราะมีส่วนประกอบของโลหะ ซึ่งจะทำให้เกิดสิ่งแปลกปลอมบนภาพได้
- ถึงแม้ว่าจะไม่มีรายงานผลข้างเคียงต่อทารกในครรภ์ก็ตาม แต่ถ้าไม่มีความจะเป็นจริงๆ จะไม่ตรวจในผู้ป่วยที่ตั้งครรภ์ในระยะ 3 เดือนแรก

### 2.14.3 เอกซเรย์ระบบดิจิตอล (DR) [15]

เครื่องดีอาร์มีหลักการการทำงาน คือ เมื่อฉายรังสีเอกซจากหลอดเอกซเรย์ผ่านผู้ป่วยไปกระทบแผ่นรับภาพที่อยู่ตรงกันข้ามทำด้วยสารกึ่งตัวนำ ซึ่งทำหน้าที่ดูดกลืนพลังงานรังสีผ่านมาแล้วแปลงพลังงานที่ได้รับเป็นสัญญาณไฟฟ้าและแปลงเป็นสัญญาณดิจิตอล ส่งระบบคอมพิวเตอร์ เพื่อสร้างเป็นภาพดิจิตอลสำหรับใช้ในการตรวจวินิจฉัยโรคต่อไป

การถ่ายภาพด้วยซีอาร์ มีหลักการการทำงานคล้ายการเอกซเรย์ทั่วไป คือ การฉายรังสีจากหลอดเอกซเรย์ของเครื่องเอกซเรย์ทั่วไป หรือเครื่องเอกซเรย์เคลื่อนที่ผ่านอวัยวะของผู้ป่วยหรือผู้รับบริการเข้ามา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นอนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กระแทกแผ่นรับภาพ มีลักษณะคล้ายกับตลับใส่ฟิล์มหรือที่เรียกกันทั่วไปว่า คาสเซต สำหรับการถ่ายภาพรังสีทั่วไป ภายในคาสเซตมีแผ่นพลาสติกบาง ฉาบด้วยสารเรืองแสงที่จะทำหน้าที่เป็นตัวเก็บพลังงานรังสีที่ตกกระทบไว้ภายในแผ่นสารเรืองแสงชั่วระยะเวลาหนึ่ง จากนั้นเมื่อต้องการสร้างภาพจะต้องทำการกระตุ้นสารเรืองแสงที่ค้างอยู่ในแผ่นรับภาพด้วยเครื่องอ่านภาพโดยเฉพาะ ภายในจะมีอุปกรณ์ที่ผลิตลำแสงเลเซอร์เพื่อทำให้สารเรืองแสงที่รับพลังงานรังสีเอกซพลดปล่อยพลังงานที่เก็บไว้ออกมาในรูปของแสง จากนั้นจะใช้หลอดโฟโตมิัลติไฟเออร์ ทำหน้าที่แปลงพลังงานแสงที่ปลดปล่อยออกมาให้เป็นสัญญาณไฟฟ้าและแปลงสัญญาณอนาล็อกเป็นดิจิทัลด้วยและผ่านขบวนการประมวลข้อมูล เพื่อสร้างภาพและแสดงภาพด้วยระบบคอมพิวเตอร์ แผ่นรับภาพที่ถูกอ่านพลังงานแล้วสามารถนำกลับมาใช้งานได้อีก โดยการลบล้างพลังงานรังสีที่ตกค้างอยู่ในแผ่นรับภาพออกให้หมดด้วยแสงสว่างที่อยู่ในเครื่องอ่านภาพ ทำให้สามารถนำคาสเซตนั้นกลับไปใช้งานใหม่ครั้งต่อไป ซึ่งอายุการใช้งานของสารเรืองแสงจะอยู่ที่ประมาณ 10,000 ถึง 30,000 ครั้งต่อแผ่น ขึ้นกับชนิดของสารเรืองแสงที่ใช้แต่ละชนิด

#### ข้อดี

- ลดปริมาณรังสีที่ใช้ในการตรวจแก่ผู้ป่วยและบุคลากรที่เกี่ยวข้อง เนื่องจากเครื่องในปัจจุบันสามารถใช้ปริมาณรังสีไม่สูงมาก
- สามารถปรับภาพและแสดงภาพได้ในลักษณะต่างๆ ช่วยให้เห็นรายละเอียดของอวัยวะต่างๆ ได้ชัดเจนขึ้น ช่วยลดจำนวนภาพถ่ายทางรังสีที่ไม่ได้คุณภาพลง
- ลดปริมาณฟิล์ม เนื่องจากแสดงผ่านจอภาพและบันทึกเข้าระบบคอมพิวเตอร์ ช่วยลดพื้นที่ในการเก็บฟิล์ม
- ลดค่าใช้จ่าย ประหยัดค่าแรง ลดขั้นตอนในการทำงาน เนื่องจากส่งข้อมูลถึงผู้ใช้โดยตรง ไม่ต้องเสียเวลาค้นหาฟิล์มเก่า กรณีที่ฟิล์มหาย หาไม่เจอ ก็ไม่ต้องถ่ายภาพใหม่

#### ข้อเสีย

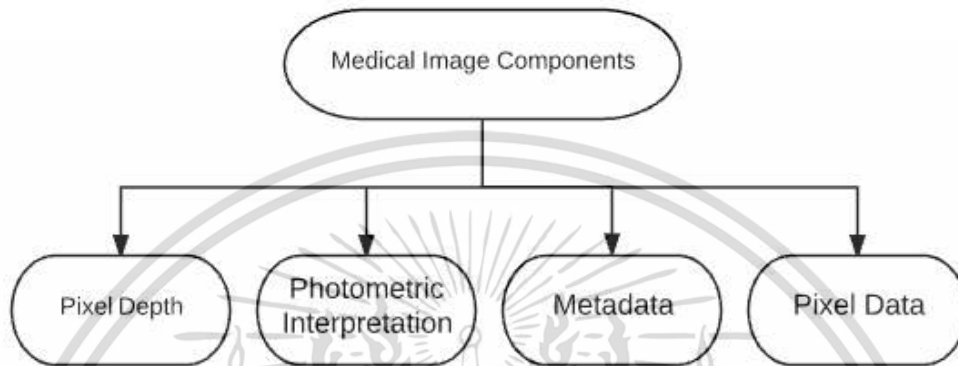
- เครื่องและอุปกรณ์ประกอบที่ใช้มีราคาแพงในระยะแรกของการลงทุน
- ต้องใช้งานด้วยความระมัดระวังในการทำงาน เนื่องจากหากอุปกรณ์ได้รับความกระทบกระเทือนมากจะทำให้อุปกรณ์ชำรุดได้ง่าย
- ใช้งานความรู้เรื่องระบบคอมพิวเตอร์ เพื่อสามารถดูแลบำรุงหรือแก้ไขปัญหาได้ง่ายขึ้น

## 2.15 รูปภาพและส่วนประกอบของภาพทางการแพทย์ [16]

ภาพทางการแพทย์เป็นตัวแทนของโครงสร้างภายในหรือการทำงานของภูมิภาคกายวิภาค ซึ่งอยู่ในรูปของอาร์เรย์ขององค์ประกอบภาพที่เรียกว่า พิกเซล 2 มิติ หรือ ว็อกเซล 3 มิติ เป็นการแสดงแบบไม่ต่อเนื่องที่เกิดจากการสุ่มตัวอย่างหรือกระบวนการสร้างภาพที่จับคู่ค่าตัวเลขกับตำแหน่งแมป จำนวนพิกเซลที่ใช้เพื่ออธิบายขอบเขตการมองเห็นของวิธีการที่ได้มาบางอย่างคือการแสดงออกของรายละเอียดที่สามารถ

อธิบายลักษณะทางกายวิภาคหรือฟังก์ชันได้ ค่าตัวเลขของพิกเซลที่แสดงออกมานั้นขึ้นอยู่กับวิธีการสร้างภาพ โปรโตคอลการตั้งค่า การสร้างภาพ และสุดท้ายคือการประมวลผลภายหลัง

### ส่วนประกอบภาพทางการแพทย์



รูปที่ 2.39 แผนผังแสดงส่วนประกอบภาพทางการแพทย์  
(ที่มา : <https://www.kdnuggets.com/2017/07/medical-image-analysis-deep-learning-part-4.html>)

รูปภาพทางการแพทย์มี 4 องค์ประกอบหลัก คือ ความลึกของพิกเซล การระบุความเข้มแสง ข้อมูลเมตา และข้อมูลพิกเซล องค์ประกอบเหล่านี้มีหน้าที่กำหนดขนาดและความละเอียดของภาพ

#### 2.15.1 ความลึกของพิกเซลหรือความลึกบิตหรือความลึกของสี

คือจำนวนบิตที่ใช้เข้ารหัสข้อมูลของแต่ละพิกเซล ตัวอย่างเช่น ภาพขนาด 8 บิตสามารถมีค่าที่ไม่ซ้ำกันได้ 256 ค่าซึ่งมีช่วงตั้งแต่ 0 ถึง 255

Bit depth	Range of values that each cell can contain
1 bit	0 to 1
2 bit	0 to 3
4 bit	0 to 15
Unsigned 8 bit	0 to 255
Signed 8 bit	-128 to 127
Unsigned 16 bit	0 to 65535
Signed 16 bit	-32768 to 32767
Unsigned 32 bit	0 to 4294967295
Signed 32 bit	-2147483648 to 2147483647
Floating-point 32 bit	$-3.402823466e+38$ to $3.402823466e+38$
Unsigned 64 bit	0 to 18446744073709551616

รูปที่ 2.40 ตารางแสดงค่าจำนวนบิตที่ใช้เข้ารหัสข้อมูลของแต่ละพิกเซล  
(ที่มา : <https://www.kdnuggets.com/2017/07/medical-image-analysis-deep-learning-part-4.html>)

### 2.15.2 การระบุความเข้มแสง

เป็นวิธีการระบุข้อมูลพิกเซลสำหรับการแสดงภาพที่ถูกต้องว่าเป็นภาพขาวดำหรือภาพสี เพื่อที่จะระบุว่าใช้ข้อมูลสีหรือไม่ในค่าพิกเซลของภาพที่เก็บไว้ แนะนำแนวคิดของตัวอย่างต่อพิกเซลหรือที่เรียกว่าจำนวนช่องสัญญาณ ภาพขาวดำมีหนึ่งตัวอย่างต่อพิกเซลและไม่มีข้อมูลสีเก็บไว้ในภาพ ระดับสีภาพจะมีเฉดสีเทาจากดำไปถึงขาว จำนวนเฉดสีเทาขึ้นอยู่กับจำนวนบิตที่ใช้จัดเก็บตัวอย่าง ซึ่งในกรณีนี้ตรงกับความลึกของพิกเซล ภาพรังสีวิทยาทางคลินิก เช่น ภาพเอกซเรย์คอมพิวเตอร์ (CT) และภาพเรโซแนนซ์แม่เหล็ก (MRI) มีการตีความโพโตเมตริกในระดับสีเทา ภาพทางเวชศาสตร์นิวเคลียร์ เช่น เอกซเรย์ปล่อยโพซิตรอน (PET) และเอกซเรย์ปล่อยโฟตอนเดี่ยว (SPECT) โดยทั่วไปจะแสดงด้วยแมปสีหรือเพรทสี

### 2.15.3 ข้อมูลเมตา

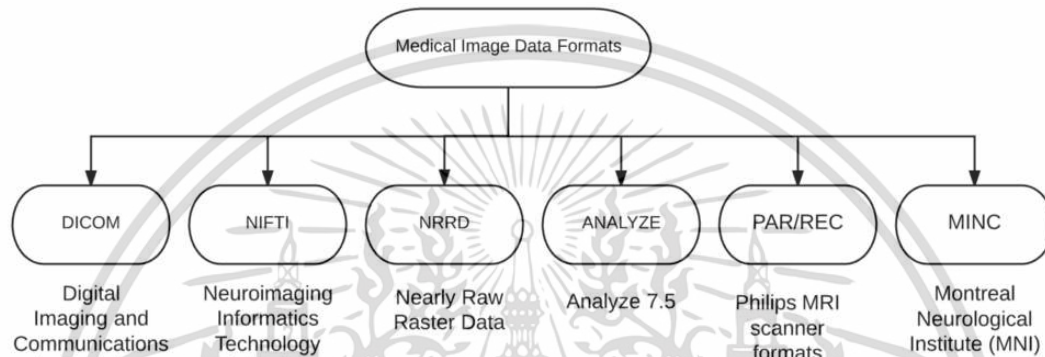
คือข้อมูลที่อธิบายรูปภาพอาจดูแปลกแต่ในรูปแบบไฟล์ใดๆ ก็ตาม มักมีข้อมูลที่เกี่ยวข้องกับรูปภาพ นอกเหนือจากข้อมูลพิกเซลเสมอ ข้อมูลนี้เรียกว่าข้อมูลเมตาโดยทั่วไปจะเก็บไว้ที่ไฟล์ตั้งต้นเป็นไฟล์นำ และประกอบด้วยมิติเมตริกซ์ของภาพ ความละเอียดเชิงพื้นที่ ความลึกของพิกเซล และการระบุความเข้มแสงเป็นอย่างน้อย

### 2.15.4 ข้อมูลพิกเซล

นี่คือส่วนที่เก็บค่าตัวเลขของพิกเซล ตามประเภทข้อมูลข้อมูลพิกเซลจะถูกจัดเก็บเป็นจำนวนเต็มหรือตัวเลขทศนิยมโดยใช้จำนวนไบต์ขั้นต่ำที่จำเป็นในการแสดงค่า

### 2.16 รูปแบบไฟล์ภาพทางการแพทย์

มี 6 รูปแบบที่เป็นที่นิยมสำหรับภาพทางรังสีวิทยา



รูปที่ 2.41 แผนผังแสดงรูปแบบไฟล์ภาพทางการแพทย์

(ที่มา : <https://www.kdnuggets.com/2017/07/medical-image-analysis-deep-learning-part-4.html>)

#### 2.16.1 DICOM (การถ่ายภาพดิจิทัลและการสื่อสารในการแพทย์)

DICOM ย่อมาจาก Digital Imaging and Communications in Medicine DICOM เป็นมาตรฐานที่สร้างขึ้นโดยสมาคมผู้ผลิตไฟฟ้าแห่งชาติ (NEMA) กำหนดมาตรฐานในการจัดการ จัดเก็บ พิมพ์ และส่งข้อมูลในการถ่ายภาพทางการแพทย์ นี่คือนิยามของไฟล์ที่ได้จากเครื่องสแกนหรือ PACS ของโรงพยาบาล (ระบบจัดเก็บรูปภาพและการสื่อสาร) ประกอบด้วยรูปแบบไฟล์และโปรโตคอลการสื่อสารของโครงข่าย ที่ใช้ TCP/IP เพื่อสื่อสารระหว่างเอนทิตีที่สามารถรับภาพและข้อมูลผู้ป่วยในรูปแบบ DICOM ไฟล์ DICOM ประกอบด้วยส่วนข้อมูลรายละเอียดและข้อมูลรูปภาพในไฟล์เดียวกัน (\*.dcm) ขนาดของส่วนข้อมูลรายละเอียดขึ้นอยู่กับจำนวนข้อมูลรายละเอียดที่มีให้ ส่วนข้อมูลรายละเอียดประกอบด้วยข้อมูลเช่น รหัสผู้ป่วย ชื่อผู้ป่วย รูปแบบเครื่องมือ และข้อมูลอื่นๆ นอกจากนี้ยังกำหนดจำนวนเฟรมที่มีอยู่และความละเอียดใดๆ จะต้องใช้โปรแกรมสำหรับดูภาพเพื่อแสดงภาพสำหรับการได้มาของไฟล์ภาพในครั้งเดียวจะมีไฟล์ DICOM จำนวนมาก

### 2.16.2 NIFTI (Neuroimaging Informatics Technology Initiative)

เดิมที NIFTi ถูกสร้างขึ้นสำหรับภาพทางระบบประสาทรูปแบบ NIFTi ได้รับการสร้างขึ้นโดย Neuroimaging Informatics Technology Initiative (NIFTI) เพื่อทดแทนรูปแบบ ANALYZE 7.5 ที่มีต้นกำเนิดในด้านการสร้างภาพระบบประสาท แต่สามารถใช้ในด้านอื่นได้เช่นกัน ลักษณะสำคัญคือรูปแบบประกอบด้วยคำจำกัดความของพิกัดที่เกี่ยวข้องสองรายการ ซึ่งเกี่ยวข้องกับดัชนีแต่ละวอคเซล (i,j,k) กับตำแหน่งเชิงพื้นที่ (x,y,z)

### 2.16.3 NRRD

รูปแบบ Nrrd ที่มีความยืดหยุ่นประกอบด้วยไฟล์ส่วนข้อมูลนำและไฟล์รูปภาพในไฟล์เดียวที่สามารถแยกหรือรวมกันได้ ส่วนข้อมูลนำ Nrrd แสดงข้อมูลแรสเตอร์ N มิติได้อย่างถูกต้องสำหรับการสร้างภาพทางวิทยาศาสตร์และการประมวลผลภาพทางการแพทย์ National Alliance for Medical Image Computing (NA-MIC) ได้พัฒนาวิธีการใช้รูปแบบ Nrrd เพื่อแสดงปริมาณ Diffusion Weighted Images (DWI) และ Diffusion Tensor Images (DTI) สามารถอ่านข้อมูล Nrrd DWI และ Nrrd DTI ลงใน 3D Slicer ได้ เพื่อยืนยันด้วยสายตาว่าการวางแผนของเทนเซอร์นั้นสอดคล้องกับกายวิภาคของระบบประสาทที่คาดหวัง

### 2.16.4 MINC

MINC ย่อมาจาก Medical Imaging NetCDF Toolkit การพัฒนารูปแบบไฟล์ MINC เริ่มต้นในปี 1992 ที่ Montreal Neurological Institute (MNI) ขณะนี้ทีมงานที่กำลังดำเนินการอยู่ที่ศูนย์การถ่ายภาพสมองของแมคกิลล์ (BCI) รูปแบบ Minc เวอร์ชันแรก (Minc1) ใช้รูปแบบมาตรฐาน Network Common Data (NetCDF) Minc2 เปลี่ยนจาก NetCDF เป็น Hierarchical Data Format เวอร์ชัน 5 (HDF5) HDF5 รองรับประเภทข้อมูลได้ไม่จำกัด และได้รับการออกแบบมาสำหรับ I/O ที่ยืดหยุ่นและมีประสิทธิภาพ และสำหรับข้อมูลที่มีปริมาณมากและซับซ้อน คุณสมบัติที่เพิ่มเข้ามาเหล่านี้ช่วยให้ Minc2 ทำงานกับชุดข้อมูลขนาดใหญ่และซับซ้อนได้

## 2.17 เมตริกและการประเมินชุดข้อมูลการแบ่งกลุ่ม

### เมตริกการประเมินผล

การประเมินคุณภาพของอัลกอริทึมจำเป็นต้องมีตัวบ่งชี้วัตถุประสงค์ที่ถูกต้อง ในทางการแพทย์ อัลกอริทึมการแบ่งส่วนคำอธิบายประกอบที่ว่าวัดด้วยมือของแพทย์มักจะใช้เป็นมาตรฐานอ้างอิง เรียกว่า กราวทฤษฎ (GT) หรือ มาสก์ (Mask) ผลลัพธ์อื่นๆ ของอัลกอริทึมการแบ่งส่วนคือผลการทำนาย (prediction) การประเมินการแบ่งส่วนของภาพทางการแพทย์คือ แบ่งออกเป็นวิธีการประเมินแต่ละพิกเซลและตามการซ้อนทับกัน

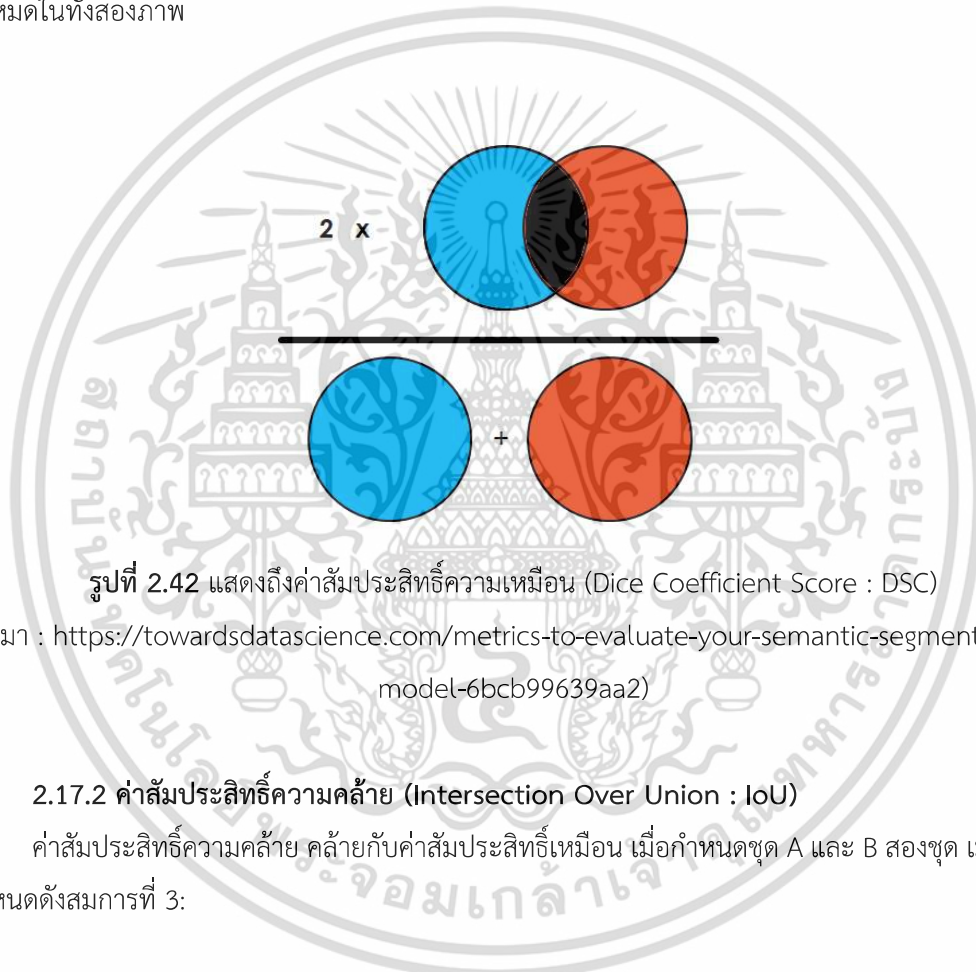
#### 2.17.1 ค่าสัมประสิทธิ์ความเหมือน (Dice Coefficient Score : DSC) [17]

ค่าสัมประสิทธิ์ความเหมือนเป็นฟังก์ชันสำหรับประเมินความเหมือน มักจะใช้เพื่อคำนวณความเหมือนหรือการซ้อนทับระหว่างสองตัวอย่าง และนำมาใช้ในการวัดประสิทธิภาพป้อยที่สุด ช่วงค่าของ

สัมประสิทธิ์คือ 0 ถึง 1 ยิ่งค่าเข้าใกล้ 1 มากเท่าไรผลการทำนายการแบ่งส่วนก็จะยิ่งดีขึ้นเท่านั้น เมื่อกำหนด A และ B สองชุด เมตริกถูกกำหนดดังสมการที่ 2:

$$DSC = 2 \frac{|A \cap B|}{|A| + |B|} \quad (2)$$

จากรูปที่ 2.42 อธิบายได้ว่าค่าสัมประสิทธิ์ความเหมือนคือ 2 คูณด้วยพื้นที่ที่ซ้อนทับหารด้วยจำนวนพิกเซลทั้งหมดในทั้งสองภาพ



รูปที่ 2.42 แสดงถึงค่าสัมประสิทธิ์ความเหมือน (Dice Coefficient Score : DSC)

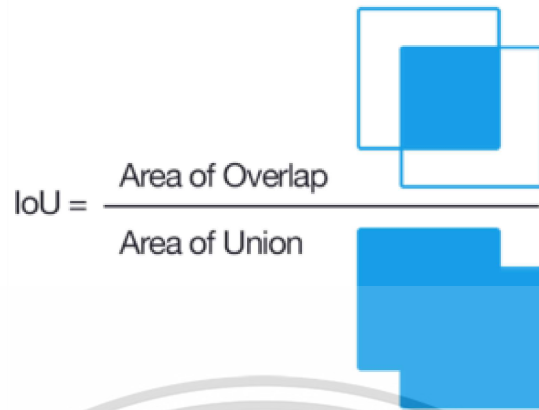
(ที่มา : <https://towardsdatascience.com/metrics-to-evaluate-your-semantic-segmentation-model-6bcb99639aa2>)

### 2.17.2 ค่าสัมประสิทธิ์ความคล้าย (Intersection Over Union : IoU)

ค่าสัมประสิทธิ์ความคล้าย คล้ายกับค่าสัมประสิทธิ์ความเหมือน เมื่อกำหนดชุด A และ B สองชุด เมตริกถูกกำหนดดังสมการที่ 3:

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (3)$$

ค่าสัมประสิทธิ์ความคล้าย คือพื้นที่ที่ทับซ้อนกันระหว่างการแบ่งส่วนที่ทำนายไว้และพื้นที่จริงทั้งหมด หารด้วยพื้นที่ที่รวมกันระหว่างการแบ่งส่วนที่ทำนายไว้และพื้นที่จริง ดังที่แสดงในรูปที่ 2.43 เมตริกนี้มีค่าตั้งแต่ 0 ถึง 1 โดย 0 หมายถึงไม่มีการทับซ้อนกัน และ 1 หมายถึงการแบ่งส่วนที่ทับซ้อนกันอย่างสมบูรณ์



รูปที่ 2.43 แสดงถึงค่าสัมประสิทธิ์ความคล้าย (Intersection Over Union : IoU)

(ที่มา : <https://towardsdatascience.com/metrics-to-evaluate-your-semantic-segmentation-model-6bcb99639aa2>)

### 2.17.3 ค่าความแม่นยำของพิกเซล (Accuracy)

ค่าความแม่นยำของพิกเซล คือจำนวนพิกเซลที่จำแนกประเภทได้อย่างถูกต้อง ในมาส์การแบ่งส่วนที่สร้างขึ้น และเป็นเมตริกที่ง่ายที่สุดในการใช้ประเมินประสิทธิภาพ แต่อาจไม่ได้ค่าประสิทธิภาพของโมเดลจริงๆ ปัญหาเกี่ยวกับเมตริกความแม่นยำของพิกเซล นั้นคือเกิดความเอนเอียงเสมอเมื่อชุดข้อมูลมีคลาสที่ไม่สมดุลอย่างมาก อาจจะได้รับความแม่นยำประมาณ 90% แต่ความเป็นจริงประสิทธิภาพเชิงคุณภาพอาจจะแย่งลง

เมื่อพิจารณาถึงความแม่นยำของพิกเซลต่อคลาส เรากำลังประเมินไบนารีมาส์เป็นหลัก ค่าผลบวกจริง (TP) แสดงถึงพิกเซลที่ทำนายได้อย่างถูกต้องว่าเป็นของคลาสที่กำหนดก็คือมาส์การแบ่งส่วนที่สร้างขึ้น ในขณะที่ผลลบจริง (TN) แสดงถึงพิกเซลที่กำหนดไว้อย่างถูกต้องว่าไม่ได้อยู่ในคลาสที่กำหนดดังสมการที่ 4:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (4)$$

ในบางครั้งเมตริกนี้อาจให้ผลลัพธ์ที่ทำให้เข้าใจผิด เมื่อการแสดงคลาสมีขนาดเล็กภายในภาพ เนื่องจากการวัดจะมีความเอนเอียงในการวัดประสิทธิภาพว่ามีกระบวนการระบุคลาสที่เป็นเชิงลบได้ดีเพียงใด

### 2.17.4 ฟังก์ชันการสูญเสีย (Loss functions) [18]

อัลกอริทึมการเรียนรู้เชิงลึกใช้หลัก Stochastic Gradient Descent (SGD) ในการเพิ่มประสิทธิภาพและการเรียนรู้วัตถุประสงค์ เพื่อเรียนรู้เป้าหมายได้อย่างถูกต้องและรวดเร็ว การแนะนำฟังก์ชันการสูญเสียมีรากฐานมาจากการเรียนรู้ของเครื่อง การเรียนรู้โดยที่ฟังก์ชันการสูญเสียเหล่านี้ได้มีพื้นฐานของการกระจายตัวของสัญลักษณ์ ตัวอย่างเช่น ไบนารีครอสเอนโทรปี ได้มาจากการแจกแจงแบบ Bernoulli ใน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทความนี้เรามีเน้น Semantic Segmentation ดังนั้นจำนวนคลาสในระดับพิกเซลคือ จำกัดไว้ที่ 2 และจะพูดถึงการสูญเสียที่นิยมใช้กันเป็นส่วนใหญ่และเข้าใจการนำไปใช้

### 2.17.5 Binary Cross-Entropy [19]

มีวัตถุประสงค์เพื่อใช้กับการจำแนกประเภทไบนารีโดยมีค่าเป้าหมายเป็น 0 หรือ 1 โดยจะคำนวณความแตกต่างระหว่างการแจกแจงความน่าจะเป็นที่เกิดขึ้นจริงและที่คาดการณ์ไว้สำหรับการทำนายคลาส 1 คะแนนจะลดลงและค่าที่สมบูรณ์แบบคือ 0

เอนโทรปีข้ามหมายถึงการวัดความแตกต่างระหว่างการแจกแจงความน่าจะเป็นสองครั้งสำหรับตัวแปรสุ่มหรือชุดของเหตุการณ์ที่กำหนด มีการใช้กันอย่างแพร่หลายสำหรับวัตถุประสงค์ในการจำแนกประเภท และเนื่องจากการแบ่งส่วนเป็นการจำแนกในระดับพิกเซล จึงทำงานได้ดี

Binary Cross-Entropy ถูกกำหนดดังสมการ 5:

$$Loss = -\frac{1}{\text{output size}} \sum_{i=1}^{\text{output size}} y_i \cdot \log \hat{y}_i + (1 - y_i) \cdot \log(1 - \hat{y}_i) \quad (5)$$

ที่นี้  $\hat{y}$  คือค่าที่ทำนายโดยแบบจำลองการทำนาย

### 2.17.6 Dice Loss [20]

หลังจากที่คุณเข้าใจความหมายของค่าสัมประสิทธิ์ความเหมือนแล้ว Dice Loss ก็เข้าใจง่ายเช่นกัน ค่าที่ดีที่สุดคือค่าความเหมือนที่ใกล้กับ 1 และสำหรับค่าการสูญเสียจะต้องเป็นค่าที่น้อยมากซึ่งจะใช้ในการแก้ไขน้ำหนักในการย้อนกลับ

ดังนั้นเราจึงสามารถสร้างสมการเล็กๆ โดยใช้ค่าสัมประสิทธิ์ Dice เพื่อให้ได้ค่าน้อยๆ แทนค่าที่ใกล้กับ 1 กำหนดดังสมการ 6:

$$DL(y, \hat{p}) = 1 - \left( \frac{2y\hat{p}+1}{y+\hat{p}+1} \right) \quad (6)$$

เมื่อ 1 ถูกเพิ่มในตัวเศษและตัวส่วนเพื่อให้แน่ใจว่าฟังก์ชันไม่ได้ถูกกำหนดในกรณีมีขอบเขต เช่น เมื่อ

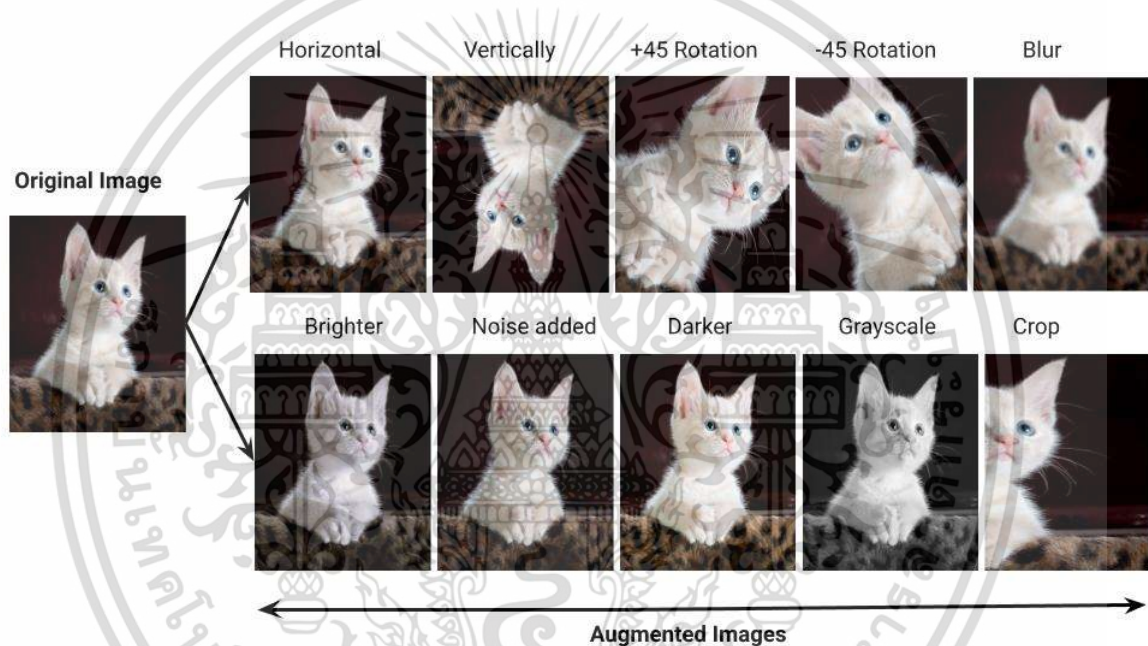
$$y = \hat{p} = 0$$

## 2.18 การขยายข้อมูล (Data Augmentation) [21]

กระบวนการของการขยายข้อมูลถูกนำมาใช้อย่างกว้างขวางในการเพิ่มประสิทธิภาพการฝึกการเรียนรู้ของโครงข่ายประสาทเทียมที่เลือกใช้ เนื่องจากข้อมูลที่ใช้ในการฝึกการเรียนรู้มีขนาดจำกัด หรือมี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จำนวนข้อมูลน้อย เนื่องจากโครงข่ายประสาทเทียมต้องการข้อมูลที่ต้องใช้ประกอบในการเรียนรู้จำนวนมาก เพื่อการเรียนรู้ที่มีประสิทธิภาพ โดยพื้นฐานแล้วกระบวนการขยายข้อมูลนี้จะเพิ่มขนาดของชุดข้อมูลการฝึกอบรมผ่านชุดฟังก์ชันการแปลงภาพ ช่วยลดปัญหาการเกิด Over Fitting ของโมเดล สามารถแก้ได้ด้วยการเพิ่มจำนวนข้อมูลในการฝึกอบรม ดังนั้นในบางกรณีเราสามารถสร้างข้อมูลขึ้นมาเองได้ โดยเฉพาะข้อมูลรูปภาพเดิมที่สามารถใช้เทคนิค เช่นการหมุนภาพ การเลื่อนภาพ การกลับภาพ เปลี่ยนสี ทำภาพให้มืดหรือสว่างขึ้น หรือใส่สัญญาณรบกวน (Noise) ลงไป จะทำให้ได้รูปภาพแบบต่าง ๆ ไม่จำกัด เป็นต้น ซึ่งนอกจากเป็นการขยายจำนวนข้อมูลได้แล้ว การขยายข้อมูลยังช่วยเพิ่มความหลากหลายของภาพในชุดข้อมูลภาพเราได้



รูปที่ 2.44 แสดงถึงชนิดการขยายข้อมูล (Data Augmentation)

(ที่มา : <https://www.tagxdata.com/571/>)

การขยายข้อมูลนั้นจะต้องไม่เหมือนข้อมูลเดิมจากต้นฉบับมากเกินไป เพราะโมเดลนั้นอาจจะไปเรียนรู้สิ่งที่ไม่สำคัญบนภาพและอาจจะจำสิ่งที่ไม่จำเป็นไป ซึ่งจะทำให้การทำนายผลนั้นใช้ไม่ได้ จะใช้กันอย่างไรแพร่หลายในข้อมูลรูปภาพ แต่ก็ปัจจุบันมีการศึกษาเกี่ยวกับการขยายข้อมูลในข้อมูลแบบอื่น ๆ เช่น ตาราง เสียงพูด และ ข้อความ NLP เช่น เปลี่ยนชื่อตัวละครจากชื่อหนึ่งเป็นอีก ชื่อหนึ่งเปลี่ยนสลับคำศัพท์ที่มีความหมายเหมือนกัน

## 2.19 Google Colab GPU [22]

Google Colaboratory คือ Jupyter notebook ดัดแปลงที่รันอยู่บนคราวด์ และไม่จำเป็นต้องติดตั้งโปรแกรมใด ๆ ก่อนใช้งาน ซึ่งทุกคนสามารถใช้งานได้ เพียงแค่มีบัญชี Google Drive เพื่อใช้ในการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จัดเก็บตัวโค้ดเท่านั้น โดยภาษา Python เป็นภาษาหลักที่ใช้ในการเขียนและรันงานบนเว็บ Colab นี้ เป็นบริการคลาวด์อีกหนึ่งบริการจาก Google Research เป็น IDE ที่อนุญาตให้ผู้ใช้เขียนซอร์สโค้ดในตัวแก้ไข และเรียกใช้จากเบราว์เซอร์ โดยเฉพาะอย่างยิ่ง รองรับภาษาการเขียนโปรแกรม Python และเน้นงานแมชชีนเลิร์นนิง การวิเคราะห์ข้อมูล โครงการการศึกษา เป็นต้น

เห็นได้ชัดว่า Google Colaboratory ไม่มีทรัพยากรที่จำกัดและไม่รับประกันว่าจะใช้งานได้ฟรี แต่จะแตกต่างกันไปตามการใช้งานที่มอบให้กับระบบ หากคุณต้องการลบข้อจำกัดเหล่านี้และรับเพิ่มเติม คุณจะต้องจ่ายค่าบริการ Colab Pro หรือ Pro + ในการใช้งานที่มีประสิทธิภาพสูง

สิ่งสำคัญที่ควรทราบคือเมื่อคุณเข้าถึง Colab ด้วยบัญชีของคุณ สิ่งที่คุณได้รับคือเครื่องเสมือนที่คุณสามารถเรียกใช้โค้ดของคุณ โดยแยกออกจากผู้ใช้และทรัพยากรอื่นๆ ดังนั้น คุณสามารถกู้คืนเครื่องเสมือนกลับเป็นสถานะดั้งเดิมได้หากคุณมีปัญหา นี่ยังหมายความว่าหากคุณรันโค้ดบางตัวใน VM และคุณปิดเบราว์เซอร์ เครื่องจะถูกลบออกหลังจากไม่มีการใช้งานเป็นระยะเวลาหนึ่งเพื่อเพิ่มทรัพยากรให้ว่าง อย่างไรก็ตาม คุณจะมีโน้ตบุ๊กใน Google Drive หากคุณบันทึกไว้ หรือคุณสามารถดาวน์โหลดได้ในเครื่อง (รูปแบบโอเพ่นซอร์ส Jupyter .ipynb) เมื่อเข้าถึง Google Colaboratory คุณจะพบ สภาพแวดล้อมที่เป็นมิตรใช้งานง่ายและใช้งานง่าย อันที่จริง มีดัชนีพร้อมเอกสารประกอบและความช่วยเหลือ ตลอดจนตัวอย่างบางส่วนเพื่อเริ่มดำเนินการตามขั้นตอนแรกของคุณ แก้ไขโค้ดที่สร้างไว้แล้วและดำเนินการทดสอบ

#### ข้อดีและข้อเสีย

- เร็วด้วย CPU และ GPU ที่ค่อนข้างดี
- ใช้งานง่าย แค่วางโค้ดแล้วกดรันได้อย่างง่าย
- สนับสนุนคำสั่ง Bash
- สามารถปิด Notebook ได้ โดยเบื้องหลังยังคงรันงานของเราอยู่
- รองรับการเชื่อมต่อกับ Google Drive ทำให้สามารถ Import & Export งานที่รันได้สะดวกมากขึ้น
- พัฒนาแอปพลิเคชันการเรียนรู้เชิงลึกโดยใช้ไลบรารียอดนิยม เช่น Keras, TensorFlow, PyTorch และ OpenCV

#### ข้อเสีย

- จะหลุดเมื่อไหร่ก็ได้ โค้ดที่รันควรมีไฟเจอร์ที่สามารถรันงานต่อจากเดิมได้
- ถ้า GPU เต็ม ก็ต้องรอสักพัก ตอนไหนที่จะใช้ได้ก็ไม่รู้อีกด้วย

## 2.20 3D slicer [23]

3D Slicer เป็นซอฟต์แวร์เปิดฟรี (เลขที่ใบอนุญาตรูปแบบ BSD) ซึ่งเป็นแพลตฟอร์มเครื่องมือที่ยืดหยุ่นสำหรับการวิเคราะห์ภาพและการสร้างภาพ 3D Slicer ถูกขยายเพื่อให้สามารถพัฒนาเครื่องมือประมวลผลทั้งแบบโต้ตอบและแบบกลุ่มสำหรับการใช้งานที่หลากหลาย 3D Slicer รับการจากการลงทะเบียนรูปภาพ การประมวลผล DTI (diffusion tractography) อินเทอร์เฟซไปยังอุปกรณ์ภายนอก สำหรับการสนับสนุนคำแนะนำเกี่ยวกับรูปภาพ และใช้งาน GPU สำหรับสร้างรูปทรงรวมถึงความสามารถ

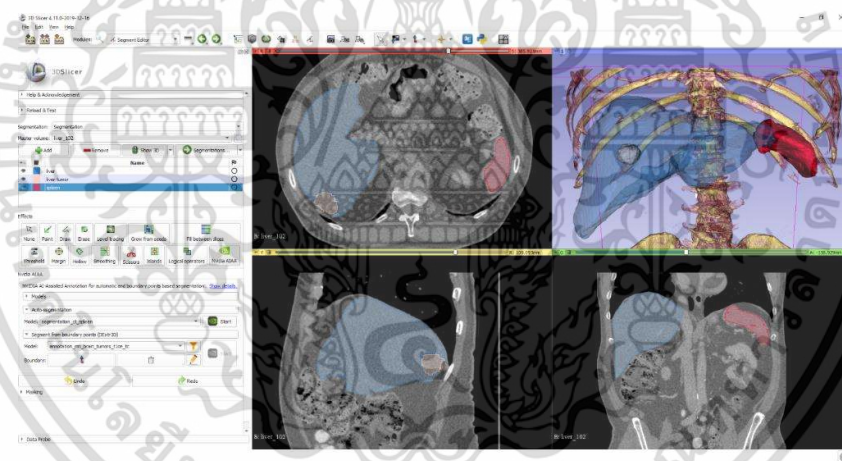
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อื่นๆ อีกมากมาย 3D Slicer ยังมีองค์กรแบบโมดูลาร์ที่อนุญาตให้เพิ่มฟังก์ชันการทำงานใหม่และมีคุณสมบัติทั่วไปจำนวนหนึ่งที่ไม่อยู่ในเครื่องมือการแข่งขัน ความสามารถในการสร้างภาพแบบอินเทอร์แอคทีฟของ 3D Slicer รวมถึงความสามารถในการแสดงรูปภาพเป็นสไลด์ สร้างภาพพื้นผิวจากรูปภาพต้นแบบ และการแสดงผลแบบปริมาตรด้วยความเร่งของฮาร์ดแวร์ 3D Slicer ยังสนับสนุนชุดรูปภาพต้นแบบที่มีส่วนประกอบที่หลากหลาย

#### ความสามารถของตัวแบ่งส่วนข้อมูลรวมถึง

- การจัดการรูปภาพเป็นรูปแบบ DICOM และการอ่าน/เขียนรูปแบบอื่นๆ ที่หลากหลาย
- การแสดงภาพเชิงปริมาตรของภาพแบบวอคเซล
- การแบ่งส่วนภาพอัตโนมัติ
- การวิเคราะห์และการสร้างภาพของข้อมูลการถ่ายภาพเทนเซอร์แบบกระจาย
- การติดตามอุปกรณ์สำหรับขั้นตอนแนะนำด้วยภาพ

ตัวแบ่งส่วนข้อมูลได้รับการคอมไพล์เพื่อใช้บนแพลตฟอร์มคอมพิวเตอร์หลายรูปแบบรวมถึง Windows, Linux และ macOS



รูปที่ 2.45 แสดงการแบ่งส่วนภาพด้วยโปรแกรม 3D slicer

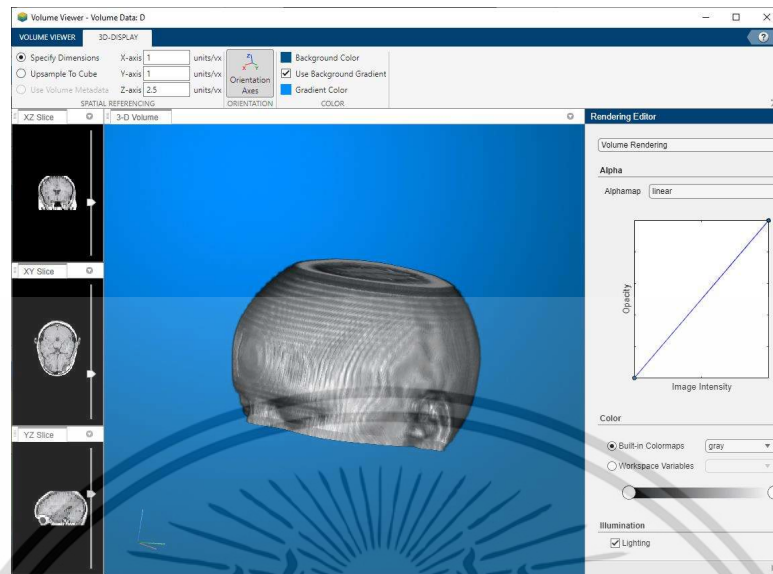
(ที่มา : <https://discourse.slicer.org/t/ai-assisted-segmentation-extension/9536>)

## 2.21 MATLAB [24]

### 2.23.1 Volume Viewer

แอปพลิเคชัน Volume Viewer ช่วยให้คุณดูข้อมูลปริมาตร 3 มิติและข้อมูลการแบ่งภาพต้นแบบเชิงปริมาตร 3 มิติ เมื่อใช้แอปนี้สามารถดูข้อมูลเป็นปริมาณหรือแบบระนาบ สามารถดูข้อมูลที่มีความเข้มสูงหรือพื้นผิว เมื่อใช้คอมพิวเตอร์ Rendering Editor คุณสามารถปรับความทึบเพื่อดูโครงสร้างใดก็ได้ที่ข้อมูลที่คุณต้องการดูและทำให้โครงสร้างเหล่านั้นโปร่งใสในโทรศัพท์ข้อมูลที่ไม่ต้องการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.46 แอปพลิเคชัน Volume Viewer จากโปรแกรม Matlab  
(ที่มา : <https://www.mathworks.com/help/images/ref/volumeviewer-app.html>)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 3

# วิธีดำเนินงานวิจัย

### 3.1 คำอธิบายชุดข้อมูล

ในการดำเนินงานวิจัยนี้เราใช้ภาพเอกซเรย์คอมพิวเตอร์ที่มีการฉีดสารทึบรังสี ใช้ภาพจากการสแกนในเฟส Arterial คือเฟสที่มีสารทึบรังสีไหลเวียนในหลอดเลือดแดงเรียกว่า CTA ซึ่งประกอบด้วยด้วย 3 กลุ่มอาการ คือ หลอดเลือดที่มีเส้นผ่านศูนย์กลางขนาดปกติ หลอดเลือดที่มี AAA และ AAA ที่มี ILT ที่ได้มาจากส่วนหนึ่งของกระบวนการทางการแพทย์ที่เป็นมาตรฐานและจากการรักษาผู้ป่วย ภาพเอกซเรย์คอมพิวเตอร์ทางการแพทย์ CTA มีรูปแบบเป็นลักษณะชุดข้อมูลจำนวนมาก ที่จัดทำโดยคณะแพทยศาสตร์โรงพยาบาลรามาธิบดี มหาวิทยาลัยมหิดล การศึกษาวิจัยนี้ได้รับการตรวจสอบและอนุมัติโดยคณะกรรมการจริยธรรมการวิจัยของมนุษย์เลขที่ใบอนุญาติ COA. MURA2021/575 อนุมัติเมื่อวันที่ 14 กรกฎาคม ค.ศ. 2021 ชุดข้อมูลนี้เป็นกรนำข้อมูลย้อนหลังของเวชระเบียนที่มีอยู่ตั้งแต่ปี ค.ศ. 2019 ถึง 2022 และประกอบด้วยภาพจากเอกซเรย์คอมพิวเตอร์ CTA ของผู้ป่วยที่มีการตรวจเกี่ยวกับการมีข้อบ่งชี้เกี่ยวกับการโป่งพองของเส้นเลือดแดงใหญ่ส่วนท้องก่อนที่จะได้รับการผ่าตัด และรักษาเป็นพยาธิสภาพหลัก จำนวน 60 ราย โดยจำแนกออกเป็นหลอดเลือดแดงปกติ 8 ราย หลอดเลือดที่มีภาวะ AAA 14 รายและ AAA ที่มี ILT 38 รายตามลำดับ ในการดำเนินการวิจัยนี้การเก็บข้อมูลจะไม่รวมผู้ป่วยที่มีการตรวจ CTA พบกลุ่มอาการที่ไม่เกี่ยวกับการโป่งพองของหลอดเลือด ได้แก่ Aortic Dissection หรือภาวะหลอดเลือดแดงใหญ่ฉีกขาด และหลอดเลือดใหญ่ทรวงอกโป่งพอง เป็นต้น การสร้างภาพเอกซเรย์คอมพิวเตอร์หลอดเลือด CTA ชุดข้อมูลจะมีลักษณะเป็นโวลูม(Volumes) ไฟล์คือไฟล์ภาพที่มีการซ้อนทับกันของภาพ 2 มิติหลายๆ ภาพหรือหลายๆ สไลด์ และได้มาจากเครื่องสแกนจากฮาร์ดแวร์ที่หลากหลายในผู้ผลิตที่แตกต่างกัน โดยมาจากผู้ผลิตดังนี้ Philips จากอัมสเตอร์ดัม ประเทศเนเธอร์แลนด์ และ GE Medical System จากชิคาโก รัฐอิลลินอยส์ ประเทศสหรัฐอเมริกา เป็นเครื่องเอกซเรย์คอมพิวเตอร์ความเร็วสูงประเภท Multi-Detector CT (MDCT) หรือ Multi-slice CT (MSCT) สามารถสร้างภาพได้หลายภาพต่อการหมุน 1 รอบ 360 องศา ของหลอดเลือดเอกซเรย์ซึ่งสามารถสร้างภาพที่ให้รายละเอียดของภาพสูง โดยมีขนาดของพิกเซลในระนาบ x และ y ประมาณ 0.625 - 0.74 มิลลิเมตร ในแนวระนาบ z ประมาณ 0.44 - 0.8 มิลลิเมตร โดยขนาดภาพ CTA จะมีขนาดเมตริกซ์ไซนในระนาบ x และ y เท่ากับ 512 x 512 ขนาดความหนาของชุดภาพจะเป็นไปตามความยาวของภาพในบริเวณที่สนใจ โดยแต่ละสไลด์จะแสดงบริเวณหลอดเลือดและเนื้อเยื่อรอบ ๆ ในระนาบรูปแบบ 2 มิติ โดยที่เส้นเลือดแดงใหญ่ส่วนทรวงอกจะถูกตัดออกจากงานวิจัยนี้

### 3.2 การเตรียมชุดข้อมูล

ชุดข้อมูลภาพ CTA แบบโวลูมดังรูป 3.1 เป็นภาพทางการแพทย์โดยไฟล์ภาพจะเป็น DICOM ไฟล์ กระบวนการประมวลผลภาพก่อน (Preprocessing) เข้ากระบวนการฝึกด้วยการเรียนรู้เชิงลึก เราได้แปลงไฟล์ภาพ

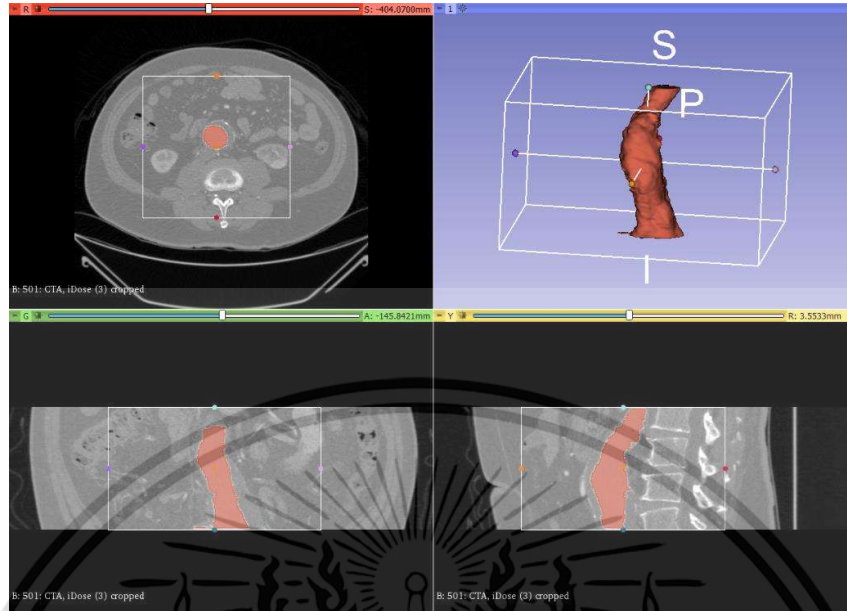
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากต้นฉบับเป็น NIFTI ไฟล์ ที่มีระดับความเทาของภาพ 16 บิต เนื่องจากการวิจัยนี้เป็นการฝึกด้วยรูปแบบภาพสามมิติ 3D volume การใช้ DICOM ไฟล์นั้นกระบวนการสร้างอัลกอริทึมค่อนข้างซับซ้อน จากนั้นจะกำหนดขอบเขตบริเวณที่สนใจ Region of Interest (ROI) จากไฟล์ภาพโวลุ่มต้นแบบบริเวณที่สนใจในการวิจัยนี้คือจะต้องครอบคลุมบริเวณเส้นเลือดแดงใหญ่ส่วนท้องจะต้องแสดงในภาพที่เราจัดทำขึ้น เพื่อใช้ในการฝึกอบรม และทดสอบโมเดล เพื่อช่วยลดการใช้หน่วยความจำของ GPU ที่มากเกินไปจนเกิดอาการค้าง และเพื่อเน้นขอบเขตข้อมูลที่สำคัญที่สนใจเท่านั้น ในการตัดแบ่งภาพของภาพหลอดเลือดแดงใหญ่ส่วนท้อง การกำหนดขอบเขตบริเวณที่สนใจนี้ ยังลดสัญญาณรบกวนจากบริเวณภาพอวัยวะที่อยู่ภายนอกความสนใจได้อีกด้วย เช่น ลำไส้ ไต เป็นต้น รวมทั้งลดความไม่เสถียรของหน่วยความจำ GPU เนื่องจากปริมาณอินพุตมีขนาดเล็กลงอย่างมากช่วยให้การเรียนรู้ด้วยหน่วยความจำ GPU นั้นทำได้เร็วขึ้น ขนาดของภาพหลังจากกำหนดขอบเขตที่สนใจแล้ว ไฟล์ภาพจะมีขนาดลดลงเหลือขนาดเท่ากับ  $256 \times 256 \times 256$  voxels ทั้งสามแกน เพื่อลดเวลาการฝึกอบรมและความต้องการหน่วยความจำ โดยกระบวนการดังกล่าวข้างต้นทั้งหมดดำเนินการด้วยโปรแกรม 3D Slicer ที่เป็นซอฟต์แวร์แบบกึ่งอัตโนมัติ สำหรับใช้ในการเตรียมข้อมูลในงานวิจัยสำหรับการตัดแบ่งภาพต่างๆ อย่างแพร่หลาย แต่ต้องมีการวาดการตัดออกด้วยการวาดด้วยมือของผู้วิจัยร่วมกับนักรังสีการแพทย์เพื่อความถูกต้องของข้อมูลร่วมด้วย และใช้โปรแกรม 3D Slicer นี้ทำชุดข้อมูลการตัดแบ่งภาพต้นแบบหรือการทำมาส์กที่ใช้ในการฝึกการเรียนรู้เชิงลึก ด้วยการใช้นิยาม Thresholding คือการจับระดับความเทาของแต่ละอวัยวะที่ต้องการแสดงในรูปที่ 3.2 ซึ่งเป็นวิธีการอีกทางหนึ่งของการแบ่งส่วนภาพแบบกึ่งอัตโนมัติของภาพ CT แสแกนด้วยโปรแกรมสำเร็จ แต่เพราะการแบ่งส่วนด้วยตนเองนั้นจะให้ออกมาได้ดีแต่ใช้เวลาดำเนินการนาน ในขณะที่การแบ่งส่วนกึ่งอัตโนมัติจะเร็วกว่าเล็กน้อยแต่ความถูกต้องอาจจะไม่ถูกต้องเสมอไปจะต้องแก้ไขด้วยการวาดด้วยมืออีกครั้งหนึ่ง ด้วยเหตุผลนี้ การพัฒนาระบบการแบ่งส่วนแบบอัตโนมัติจะสามารถทดแทนวิธีการข้างต้นได้ทั้งหมด สามารถทำได้และเป็นวิธีการอีกทางเลือกที่ยอดเยียมสำหรับการตัดแบ่งส่วนภาพที่ถูกต้องและใช้เวลาในการทำนายภาพไม่นานเกินไป



รูปที่ 3.1 แสดงภาพเอกซเรย์คอมพิวเตอร์แบบโวลุ่ม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.2 แสดงการใช้ 3D Slicer เป็นซอฟต์แวร์แบบกึ่งอัตโนมัติทำการตัดแบ่งภาพบริเวณที่สนใจ

### 3.3 การเตรียมชุดข้อมูล CT Images

จากวิธีการเตรียมชุดข้อมูลในหัวข้อก่อนหน้านี้ ชุดข้อมูลที่ใช้สำหรับการฝึกอบรมในการศึกษานี้จะต้องได้รับการตรวจสอบโดยนักรังสีการแพทย์ชำนาญการที่มีประสบการณ์มากกว่า 5 ปีในการแปลผลภาพเอกซเรย์คอมพิวเตอร์หลอดเลือดแดงใหญ่ส่วนท้องเพื่อให้แน่ใจว่าข้อมูลมีคุณภาพและสามารถนำมาทำชุดข้อมูลในการฝึกอบรมและทดสอบโมเดลได้ โดยภาพจะถูกลดขนาดลงให้มีขนาด  $256 \times 256 \times 256$  voxels จากขนาดภาพต้นฉบับคือ  $512 \times 512 \times$  จำนวนสไลด์ หลังจากที่ได้กำหนดขอบเขตบริเวณที่สนใจด้วยโปรแกรม 3D Slicer ซึ่งไฟล์ชุดข้อมูลนี้จะเป็นภาพเอกซเรย์คอมพิวเตอร์แบบไวolum เมื่อกระบวนการลดขนาดแต่ละชุดข้อมูลสำเร็จจะทำการเอาพุ่มไฟล์ภาพจากต้นฉบับที่เป็น DICOM ไฟล์เป็น NIFTI ไฟล์ ด้วยโปรแกรม 3D Slicer ในขั้นตอนเดียว ซึ่งภาพที่ได้จะเป็นภาพเอกซเรย์คอมพิวเตอร์ในส่วนของหลอดเลือดแดงใหญ่ส่วนท้องของสามกลุ่มอาการเรียกว่า ภาพเอกซเรย์คอมพิวเตอร์ต้นแบบ โดยมีจำนวนชุดข้อมูลภาพ 60 ชุดข้อมูล โดยจำแนกเป็น 3 กลุ่มอาการของหลอดเลือดดังที่ได้กล่าวมาข้างต้น ภาพที่นำเข้ากระบวนการมีขนาด  $256, 256, 256$  มี channel คือ 1

### 3.4 การเตรียมชุดข้อมูลแบบตัวอย่างการฝึกอบรม Ground truth annotations

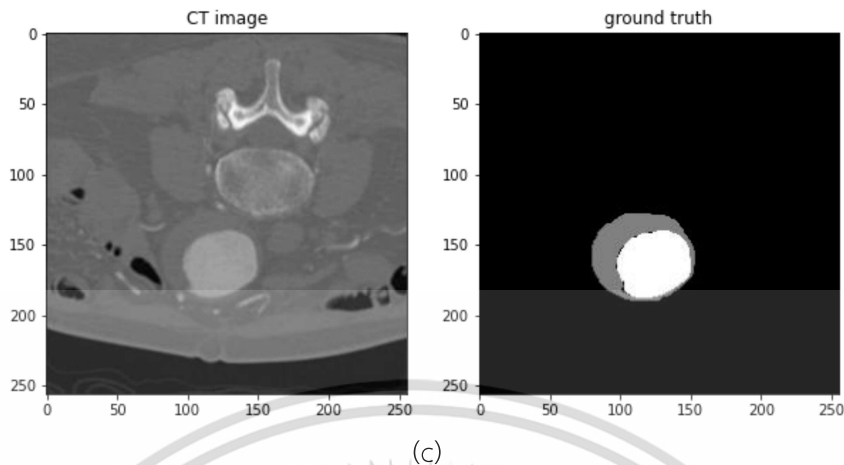
ชุดข้อมูลการตัดแบ่งภาพต้นแบบในการฝึกอบรม Ground truth annotations โดยใช้ 3D Slicer เป็นซอฟต์แวร์แบบกึ่งอัตโนมัติทำการตัดแบ่งภาพ ด้วยการใช้เทคนิค Thresholding ที่มีมาพร้อมกับโปรแกรม เพื่อทำการลบอวัยวะที่ไม่ต้องการออกจากภาพ แต่ถ้ายังมีอวัยวะที่ไม่ได้สนใจหลงเหลืออยู่ในภาพจะต้องลบออกด้วยการตัดออกด้วยมือ โดยนักรังสีการแพทย์ที่มีความชำนาญการเป็นผู้ควบคุมเพื่อให้ได้ชุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูลที่ถูกต้อง และเหลือแต่หลอดเลือดแดงใหญ่ที่เราสนใจ โดยที่พื้นหลังที่ไม่ได้สนใจนั้นจะเป็นเฉดสีดำ เพื่อนำมาฝึกให้โมเดลเรียนรู้และใช้ประเมินเทียบประสิทธิภาพของการคาดเดาของโมเดล นอกจากนี้ชุดข้อมูลแบบตัวอย่างการฝึกอบรมที่สร้างขึ้น จะมีภาพที่แยกระหว่างลูเมนของหลอดเลือดแดงและลิ้มเลือดที่ผนังหลอดเลือด ILT การฝึกและทดสอบโมเดลด้วย ในการแบ่งส่วนของ ILT ด้วยโปรแกรมอัตโนมัติเหล่านี้ มักไม่ประสบความสำเร็จเนื่องจากความยากลำบากในการเปรียบเทียบระหว่างเนื้อเยื่อที่ไม่มีสารทึบรังสีแทรกเข้าเนื้อเยื่อของ ILT และในกรณีเหล่านี้ การแบ่งส่วนแบบการวาดด้วยมือทั้งหมดนั้นเป็นสิ่งจำเป็นโดยที่เราเน้นพื้นที่ที่น่าสนใจในแต่ละส่วน ซึ่งจะมีสามคลาสที่แสดงถึงลูเมนของหลอดเลือดเป็นเฉดสีขาว ลิ้มเลือดที่ผนังหลอดเลือดเป็นเฉดสีเทา และพื้นหลังภาพที่ไม่ได้สนใจเป็นเฉดสีดำ ชุดข้อมูลภาพจะแสดงในภาพตัดตามขวางเพื่อสร้างเป็นมาตรฐานอ้างอิงสำหรับการแบ่งภาพ ดังแสดงในรูปที่ 3.3



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.3 แสดงภาพเอกซเรย์คอมพิวเตอร์หลอดเลือด CT image และ ชุดข้อมูลแบบตัวอย่างการฝึกรูปแบบ (Ground truth) (a) หลอดเลือดขนาดปกติ (b) หลอดเลือดแดงที่มีการโป่งพอง และ (c) หลอดเลือดแดงโป่งพองร่วมกับมีลิ้มเลือดหนาที่ผนังหลอดเลือด

### 3.5 วิธีดำเนินงานวิจัย

#### 3.5.1 ความจำเป็นของคอมพิวเตอร์

ในงานนี้การทดลองดำเนินการโดยใช้ไลบรารีการเรียนรู้เชิงลึกโดยใช้ไลบรารีของ Keras และ TensorFlow เป็นหลักด้วยระบบความเร่ง GPU รุ่น NVIDIA-SMI-Persistence-M เป็นแพลตฟอร์มที่สนับสนุนคุณลักษณะการทำงานร่วมกันของ Google การใช้งานทั้งหมดดำเนินการบน Google Colab Notebook GPU ในรูปแบบการทำงานของ Jupyter notebook งานวิจัยนี้เราใช้ Colab Pro เพื่อเพิ่มความเร็วของ GPU ในการฝึกรูปแบบโมเดล

#### 3.5.2 Importing files and data preprocessing.

การนำเข้าข้อมูลภาพและแปลงไฟล์ภาพเป็นอาร์เรย์ (Arrays) ผ่านอัลกอริทึมไพธอน (python) ด้วยการใช้ไลบรารี (library) ซึ่งไลบรารีเปรียบเสมือนโปรแกรมสำเร็จรูปที่เก็บฟังก์ชันการทำงานที่เฉพาะทาง โดยมีชุดคำสั่งนำเข้าไลบรารีดังภาพที่ 3.4 ประโยชน์ของไลบรารีสำเร็จรูปที่มากับโปรแกรมก็คือผู้ใช้งานไม่จำเป็นต้องสร้างไลบรารีขึ้นมาใหม่เองทั้งหมด แต่สามารถนำไลบรารีที่ได้มีการพัฒนาไว้อยู่แล้วมาใช้งานได้เลย เช่น ฟังก์ชัน print และ for ไลบรารียอดนิยมที่ผู้เขียนไพธอนทั่วไปใช้ในการสร้างโมเดลแบบการเรียนรู้ของเครื่อง เช่น Scikit-Learn, TensorFlow, Keras, Pandas, Matplotlib, NumPy, PyTorch เป็นต้น โดยเริ่มจากการนำเข้าชุดข้อมูลภาพทั้งหมด และใช้คำสั่งแปลงไฟล์ภาพเป็นพิกเซลอาร์เรย์ ซึ่งอาร์เรย์เป็นคุณสมบัติหลักของ NumPy มีลักษณะคล้ายกับ list ยกเว้น สมาชิกทุกตัวในอาร์เรย์จะต้องเป็นข้อมูลชนิดเดียวกัน โดยทั่วไปแล้วข้อมูลที่เก็บจะเป็นตัวเลขเช่น int หรือ float โดยมีชุดคำสั่งแรกดังภาพที่ 3.5 และแสดงผลของชุดข้อมูลออกมาว่าชุดข้อมูลนั้นจับคู่การเรียนรู้ได้ตรงกันหรือไม่เพราะถ้าการจับคู่ภาพและ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

มาส์กสำหรับการเรียนรู้ไม่ตรงกัน ก็จะทำให้การเรียนรู้ที่ผิดและทำนายผลออกมาได้ไม่ถูกต้อง การแสดงผลภาพบางส่วนดังแสดงในภาพที่ 3.6

```

from __future__ import print_function
from numpy.random import randint
from skimage.transform import resize
from skimage.io import imsave
from skimage.io import imread
import random
import SimpleITK as sitk
import matplotlib.pyplot as plt
import glob
import numpy as np
import numpy as np
import nibabel as nib
import itertools
import os
from numpy.random import randint
from matplotlib import pyplot as plt

```

รูปที่ 3.4 อัลกอริทึมการนำเข้าไลบรารี (library)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

data_path = './'

train_files, mask_files = get_filenames()

train_data_path = os.path.join(data_path, 'volume/')
mask_data_path = os.path.join(data_path, 'mask/')

train_arrays_list = []
mask_arrays_list = []
print(len(train_files), 'train')
print(len(mask_files), 'mask')
for image in range(len(train_files)):
    print(train_data_path+train_files[image])
    img_train = sitk.ReadImage(train_data_path+train_files[image])
    out_img = sitk.GetArrayFromImage(img_train)
    train_arrays_list.append(out_img)
    print(mask_data_path+mask_files[image])
    img_mask = sitk.ReadImage(mask_data_path+mask_files[image])
    out_mask = sitk.GetArrayFromImage(img_mask)
    mask_arrays_list.append(out_mask)

img_vol = train_arrays_list[50]
mask_vol = mask_arrays_list[50]

im_vol = img_vol[100, :, :]
print('Shape of volume is '+str(im_vol.shape))

im_mask = mask_vol[100, :, :]
print('Shape of mask is '+str(im_mask.shape))

fig = plt.figure()

ax1 = fig.add_subplot(1, 2, 1)
ax1.imshow(im_vol)

ax2 = fig.add_subplot(1, 2, 2)
ax2.imshow(im_mask)

ax1.set_title('CT image')
ax2.set_title('CT mask')

plt.show()

```

รูปที่ 3.5 อัลกอริทึมการนำเข้าชุดข้อมูลและการแปลงภาพเป็นอาร์เรย์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

60 train
60 mask
./volumee/IMG_01.nii
./maskk/mask_01.nii
./volumee/IMG_02.nii
./maskk/mask_02.nii
./volumee/IMG_03.nii
./maskk/mask_03.nii
./volumee/IMG_04.nii
./maskk/mask_04.nii
./volumee/IMG_05.nii
./maskk/mask_05.nii
./volumee/IMG_06.nii
./maskk/mask_06.nii
./volumee/IMG_08.nii
./maskk/mask_08.nii
./volumee/IMG_09.nii
./maskk/mask_09.nii
./volumee/IMG_10.nii
./maskk/mask_10.nii

```

รูปที่ 3.6 แสดงการนำเข้าชุดข้อมูลบางส่วน

### 3.6 การสร้างแพตช์สำหรับเก็บชุดข้อมูลจากการทำฟังก์ชันเสริมข้อมูล (Patchwise Augmentation)

ขั้นตอนต่อมาเราจะสร้างสคริปต์เพื่อใช้สร้างแพตช์ (Patch) โวลุ่มจากข้อมูลตั้งต้น การสร้างแพตช์มีวัตถุประสงค์เพื่อหลีกเลี่ยงความซ้ำซ้อนของข้อมูลการฝึกอบรมของภาพทั้งหมด และเนื่องจากอินพุตเป็นข้อมูลแบบ 3 มิติ จำเป็นต้องใช้หน่วยความจำ GPU จำนวนมากจึงจำเป็นที่เราจะต้องใช้เทคนิค 3D patch-based เพื่อลดปริมาณอินพุตของข้อมูลตั้งต้นทั้งหมด เพื่อให้การเสริมข้อมูลได้รูปร่างอินพุตที่สอดคล้องกันสำหรับโครงข่ายประสาทเทียมและยังปรับข้อมูลให้เป็นมาตรฐานด้วยการเพิ่มมิติที่ 4 เพื่อให้บรรลุเป้าหมายนี้ จำเป็นต้องมีข้อมูลอินพุตจำนวนมาก เราใช้การเสริมข้อมูลเพื่อให้ได้ข้อมูลปริมาณที่มาก โดยใช้ชุดคำสั่งการสร้างแพตช์ดังภาพที่ 3.7 ก่อนเพื่อใช้เป็นทีเก็บชุดข้อมูล

โดยกระบวนการดำเนินการนั้น เราจะต้องย่อขนาดชุดข้อมูลจากชุดแรกเริ่ม เพื่อลดปริมาณอินพุต (resize) เป็นขนาด  $128 \times 128 \times 128$  voxel จากนั้นจะนำเข้าสู่กระบวนการเสริมข้อมูลด้วยการเปลี่ยนแปลงรูปร่างของข้อมูลต้นฉบับและบันทึกข้อมูลในแพตช์ที่เราสร้างไว้ โดยเราจะใช้การเสริม 4 ประเภท คือ axis y flip, shear, random crop, และ gray-scale variation ด้วยอัลกอริทึมโดยจะเป็นการเสริมข้อมูลแบบเต็มรูปแบบไม่ใช่แบบสุ่ม เพื่อให้ได้ชุดข้อมูลจำนวนมากพอและเหมาะสมสำหรับการทำนายภาพ โดยการทดลองของเราจะเริ่มจากการเสริมข้อมูลในปริมาณที่น้อยในการฝึกการเรียนรู้และทดสอบ ไปจนถึงปริมาณที่มากและสามารถทำนายข้อมูลได้ความแม่นยำสูง โดยใช้อัลกอริทึมที่นำเสนอตั้ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 3.8 เป็นคำสั่งสร้างข้อมูลใหม่โดยให้ชุดข้อมูลที่ต่างจากชุดข้อมูลเดิม และเปลี่ยนคำสั่งโดยใช้คำสั่งดังภาพที่ 3.9 – 3.12 เป็นชุดคำสั่งสำหรับสร้างแพตช์การเสริมข้อมูล ให้มีลักษณะภาพที่ต่างจากชุดข้อมูลต้นแบบเดิม การใช้นั้นสามารถนำไปใช้ร่วมกับชุดคำสั่งข้างต้นและบันทึกลงแพตช์ที่เราสร้างไว้ เพื่อเก็บเป็นชุดข้อมูลใหม่ที่จะนำไปใช้ฝึกการเรียนรู้เชิงลึกและใช้ทดสอบประสิทธิภาพของโมเดลเราด้วย ซึ่งวิธีนี้จะช่วยให้หน่วยความจำ GPU ไม่สิ้นเปลืองและเหลือไปถึงขั้นตอนฝึกการเรียนรู้เชิงลึกด้วย และไม่เสี่ยงต่อความแปรปรวนของหน่วยความจำของแพลตฟอร์ม

```

data_path = './'

def get_filenames():
    train_data_path = os.path.join(data_path, 'volumes/')
    mask_data_path = os.path.join(data_path, 'mask/')
    train_dirs = os.listdir(train_data_path)
    mask_dirs = os.listdir(mask_data_path)
    return train_dirs, mask_dirs

def create_patch_dataset():
    patch_folder_scan = './patch_train_dataset/'
    patch_folder_mask = './patch_train_mask_dataset/'
    if not os.path.exists(patch_folder_scan):
        os.makedirs(patch_folder_scan)
    if not os.path.exists(patch_folder_mask):
        os.makedirs(patch_folder_mask)
    train_files, mask_files = get_filenames()
    train_data_path = os.path.join(data_path, 'volumes/')
    mask_data_path = os.path.join(data_path, 'mask/')

```

รูปที่ 3.7 อัลกอริทึมการสร้างแพตช์เพื่อเก็บชุดข้อมูล

```

for image in range(len(train_files)):
    print(train_data_path+train_files[image])
    print(mask_data_path+mask_files[image])
    img_train = sitk.ReadImage(train_data_path+train_files[image])
    out_img = sitk.GetArrayFromImage(img_train)

    img_mask = sitk.ReadImage(mask_data_path+mask_files[image])
    out_mask = sitk.GetArrayFromImage(img_mask)

    if out_img.shape[0]-1 > 128 and out_img.shape[1]-1 > 128:
        patch_origin = randint(0, ((out_img.shape[0]-1)-128))
        patch_train = out_img[patch_origin:patch_origin+128,
                               patch_origin:patch_origin+128,
                               patch_origin:patch_origin+128]
        patch_train = patch_train.astype('float32')
        patch_train /= 255. # scale masks to [0, 1]
        patch_train = np.expand_dims(patch_train, axis = 3)
        train_arrays_list.append(patch_train)
        np.save('./patch_train_dataset/train_patch_rdc_'+str(image), (patch_train)

        patch_mask = out_mask[patch_origin:patch_origin+128,
                               patch_origin:patch_origin+128,
                               patch_origin:patch_origin+128]
        patch_mask = patch_mask.astype('float32')
        patch_mask /= 255. # scale masks to [0, 1]
        patch_mask = np.expand_dims(patch_mask, axis = 3)
        mask_arrays_list.append(patch_mask)
        np.save('./patch_train_mask_dataset/mask_patch_rdc_'+str(image), patch_mask)
    else:

```

รูปที่ 3.8 ชุดคำสั่งการสร้างข้อมูลรูปแบบต่างๆ

```

import numpy as np
import nibabel as nib
import itertools
import os
import scipy.ndimage as ndi
def resize_data(data):
    initial_size_x = data.shape[0]
    initial_size_y = data.shape[1]
    initial_size_z = data.shape[2]
    new_size_x = 128
    new_size_y = 128
    new_size_z = 128
    delta_x = initial_size_x / new_size_x
    delta_y = initial_size_y / new_size_y
    delta_z = initial_size_z / new_size_z
    new_data = np.zeros((new_size_x, new_size_y, new_size_z))
    for x, y, z in itertools.product(range(new_size_x),
                                     range(new_size_y),
                                     range(new_size_z)):
        new_data[x][y][z] = data[int(x * delta_x)][int(y * delta_y)][int(z * delta_z)]
    return new_data

```

รูปที่ 3.9 ชุดคำสั่งการลดขนาดภาพ (resize)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
def grayscale_variation(voxels):
    # Introduce a random global increment in gray-level value of volume.
    im_min = np.min(voxels)
    im_max = np.max(voxels)
    mean = np.random.normal(0, 0.05)
    smp = np.random.normal(mean, 0.05, size = np.shape(voxels))
    voxels = voxels + im_max*smp
    voxels[voxels <= im_min] = im_min # Clamp to min value
    voxels[voxels > im_max] = im_max # Clamp to max value
    return voxels
```

รูปที่ 3.10 ชุดคำสั่งการเปลี่ยนการเทา (Grayscale variation)

```
def flip3D(X):
    """
    Flip the 3D image respect one of the 3 axis chosen randomly
    """
    choice = np.random.randint(1)
    if choice == 0: # flip on y
        X_flip = X[:, ::-1, :]
    return X_flip
```

รูปที่ 3.11 ชุดคำสั่งการกลับด้านภาพ แนว x , y (Flip)

```

def shear(
    x, shear=(0.1, 0.1), is_random=False, row_index=0, col_index=1, channel_index=2
, fill_mode='nearest', cval=0.,order=1):
    if len(shear) != 2:
        raise AssertionError
    if isinstance(shear, tuple):
        shear = list(shear)
    if is_random:
        shear[0] = np.random.uniform(-shear[0], shear[0])
        shear[1] = np.random.uniform(-shear[1], shear[1])

    shear_matrix = np.array([[1, shear[0], 0], \
                             [shear[1], 1, 0], \
                             [0, 0, 1]])

    h, w = x.shape[row_index], x.shape[col_index]
    transform_matrix = transform_matrix_offset_center(shear_matrix, h, w)
    x = affine_transform(x, transform_matrix, channel_index, fill_mode, cval, order
)
    return x

```

รูปที่ 3.12 ชุดคำสั่งการเลื่อนภาพบนล่าง (Shear)













### 3.7 การโหลดไฟล์จากแพตช์ data augmentation

ไฟล์ภาพที่เราบันทึกลงแพตช์นั้น จะเป็น NPY ไฟล์ ดังภาพ 3.13 ซึ่งไฟล์ NPY เป็นไฟล์อาร์เรย์ NumPy ที่สร้างขึ้นโดยชุดซอฟต์แวร์ Python พร้อมกับติดตั้งไลบรารี NumPy มันมีอาร์เรย์ที่บันทึกในรูปแบบไฟล์ NumPy (NPY) ไฟล์ NPY เก็บข้อมูลทั้งหมดที่จำเป็นในการสร้างอาร์เรย์ใหม่บนคอมพิวเตอร์เครื่องใดก็ได้ซึ่งรวมถึงข้อมูล dtype และรูปร่าง NumPy เป็นไลบรารีภาษาการเขียนโปรแกรม Python ที่ให้การสนับสนุนอาร์เรย์และเมทริกซ์ขนาดใหญ่

- สามารถส่งออกอาร์เรย์ไปยังไฟล์ NPY โดยใช้ np.save ('filename.npy', array)
- สามารถโหลดอาร์เรย์ในไฟล์ NPY โดยใช้ np.load ('filename.npy')

โดยเราจะเปิดไฟล์ด้วยชุดคำสั่งดังแสดงในภาพ 3.14 โดยไฟล์จะเรียงลำดับตามตัวอักษร และลำดับเลขที่เราตั้งชื่อก่อนบันทึกเข้าแพตช์

My Drive > Colab Notebooks > My Drive > Colab Notebooks >

Name ↑	Name ↑
 train_patch_flip_0.npy	 mask_patch_flip_0.npy
 train_patch_flip_1.npy	 mask_patch_flip_1.npy
 train_patch_flip_2.npy	 mask_patch_flip_2.npy
 train_patch_flip_3.npy	 mask_patch_flip_3.npy
 train_patch_flip_4.npy	 mask_patch_flip_4.npy
 train_patch_flip_5.npy	 mask_patch_flip_5.npy

รูปที่ 3.13 ไฟล์แพตช์ที่ใช้เก็บข้อมูลภาพการจากใช้ฟังก์ชันการเสริมข้อมูล

```
def load_patch_dataset():
    train_files, mask_files = get_filenames()

    data_path = './'

    # paths to folders with patches
    train_data_path = os.path.join(data_path, 'patch_train_dataset/')
    mask_data_path = os.path.join(data_path, 'patch_train_mask_dataset/')

    train_arrays_list = []
    for image in range(len(train_files)):
        out = np.load(train_data_path+train_files[image])
        train_arrays_list.append(out)
    mask_arrays_list = []
    for image in range(len(mask_files)):
        out = np.load(mask_data_path+mask_files[image])
        mask_arrays_list.append(out)

    return train_arrays_list, mask_arrays_list
```

รูปที่ 3.14 อัลกอริทึมเรียกภาพจากไฟล์แพตช์

### 3.8 การแบ่งข้อมูลสำหรับใช้ในการฝึกการเรียนรู้ และทดสอบโมเดล

โดยปกติแล้วในการแบ่งข้อมูลของการทำโมเดลในการเรียนรู้เชิงลึกเราจะเห็นได้ว่าเราสามารถแบ่งข้อมูลออกเป็น 2 ส่วนหลักๆ ดังรูปที่ 3.15 นี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.15 แสดงการแบ่งสัดส่วนข้อมูลสำหรับฝึกการเรียนรู้ และทดสอบโมเดล

**Training set (labeled data):** Training set เป็นข้อมูลที่มีทั้งรูปภาพต้นแบบ feature และ label ที่ใช้ในการสอนโมเดลของเราให้มีความสามารถในการแบ่งแยกและจดจำคุณลักษณะ โดยไอดีง่าย ๆ คือเป็นการบอกโมเดลว่า หากเราเจอ feature แบบนี้ควรจำแนกข้อมูลของเราให้มี label แบบนี้ เป็นต้น

**Testing set (unlabeled data):** เป็นชุดข้อมูลที่ใช้ทดสอบว่าโมเดลของเราหากไม่เคยเห็นข้อมูลมาก่อน จะมีประสิทธิภาพทำนายได้ดีแบบ training set หรือไม่

เราจะใช้ชุดคำสั่งแบ่งชุดข้อมูลออกเป็นสองชุด คือ ชุดข้อมูลสำหรับการฝึกการเรียนรู้ และชุดข้อมูลสำหรับการทดสอบ โดยใช้ฟังก์ชันที่ใช้เพื่อรับค่าที่มีองค์ประกอบเป็นอาร์เรย์อินพุต เป็นฟังก์ชันไลบรารีทางคณิตศาสตร์ NumPy ที่คืนค่าพื้นฐานขององค์ประกอบอาร์เรย์ด้วยชนิดข้อมูลแบบ float โดยการใช้งานเราจะกำหนดจำนวนของชุดข้อมูลที่ใช้สำหรับฝึกการเรียนรู้ก่อนว่าเราต้องการใช้กี่เปอร์เซ็นต์ของข้อมูลที่มีอยู่ในแพตช์ที่เราสร้างไว้ ซึ่งจำนวนชุดข้อมูลในแพตช์ของเราจะมีทั้งหมด 300 ชุด จากเดิม 60 ชุด ที่ยังไม่ได้ทำการเพิ่มชุดข้อมูลด้วยการใช้ฟังก์ชันเสริมโดยที่เราเสริมชุดข้อมูลขึ้นมาจากชุดข้อมูลทั้งหมด 4 รูปแบบ จะเป็นข้อมูลจากการเสริมข้อมูล 240 ชุด รวมกับชุดข้อมูลเดิม 60 ชุด จากการบันทึกลำดับของข้อมูลในแพตช์นั้นจะเรียงตามอักขระที่เรากำหนด จะทำให้สะดวกแก่การแบ่งชุดข้อมูลให้ได้ตามที่เรต้องการ โดยชุดคำสั่งที่ใช้จะเป็นการแบ่งจากข้อมูลตั้งต้นไปจนถึงเปอร์เซ็นต์ที่เราที่กำหนดคือ 95 เปอร์เซ็นต์ 285 ชุดข้อมูล และชุดข้อมูลที่เหลือ 5 เปอร์เซ็นต์เป็นชุดข้อมูลที่ใช้สำหรับการทดสอบโมเดล 15 ชุดข้อมูล โดยสามารถใช้ชุดคำสั่งดังแสดงในรูปที่ 3.16 ใช้สำหรับในการแบ่งชุดข้อมูลทั้งชุดข้อมูลต้นแบบและชุดมาสก์ต้นแบบการเรียนรู้

```
def split_data():
    train_arrays_list, mask_arrays_list = load_patch_dataset()

    alpha = np.int16(np.floor(len(train_arrays_list) * 0.95))
    end = len(train_arrays_list)

    train_img = np.array(train_arrays_list[0:alpha])
    train_mask = np.array(mask_arrays_list[0:alpha])

    test_img = np.array(train_arrays_list[alpha:end])
    test_mask = np.array(mask_arrays_list[alpha:end])

    return train_img, train_mask, test_img, test_mask
```

รูปที่ 3.16 อัลกอริทึมใช้สำหรับการแบ่งข้อมูลสำหรับการฝึกการเรียนรู้ และทดสอบโมเดล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 3.9 การนำข้อมูลเข้าโครงข่ายประสาทเทียมเพื่อทำการฝึกการเรียนรู้

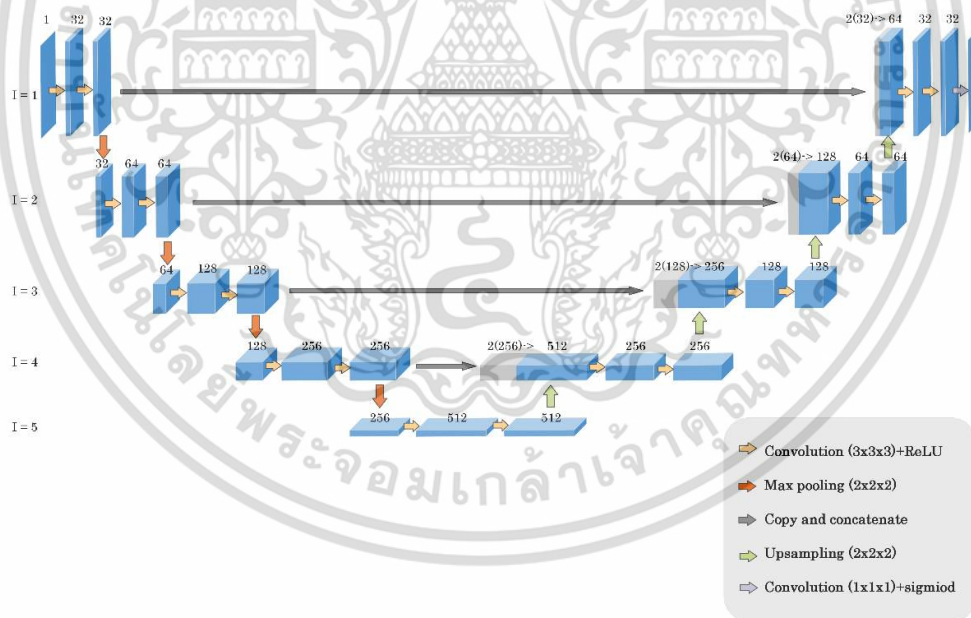
ในงานนี้เรานำเสนอผลลัพธ์ของวิธีการอัตโนมัติเต็มรูปแบบสำหรับการตัดแบ่งภาพของ AAA ด้วยคอนโวลูชันแบบสามมิติด้วยการเรียนรู้เชิงลึกโดยใช้สถาปัตยกรรมยูเน็ตสามมิติ แนวทางการใช้ยูเน็ตมีความสำคัญอย่างมากสำหรับการแบ่งภาพทางการแพทย์ส่วนใหญ่ โดยสามารถพิสูจน์ข้อเท็จจริงที่ว่าสถาปัตยกรรมนี้มีความสามารถเพียงใด ขึ้นอยู่กับปัจจัยที่แตกต่างกันและรูปแบบที่นำมาใช้ทั้งโครงข่ายภายในของยูเน็ตทั้งหมด ดังนั้น ชุดข้อมูลภาพทั้งหมดจะถูกสุ่มตัวอย่าง และโครงข่ายสถาปัตยกรรมการเรียนรู้เชิงลึกจะได้รับการเรียนรู้ด้วยชุดข้อมูลภาพตัดขวางในแนวระนาบ axial แบบสามมิติ โดยมีลำดับขั้นตอนดังต่อไปนี้ นำเข้าชุดข้อมูลด้วยขนาด  $128 \times 128 \times 128$  voxels ในระนาบ x, y และ z ตามลำดับ ชุดคำสั่งการอินพุตขนาดภาพดังภาพที่ 3.19 เพื่อลดความต้องการใช้หน่วยความจำที่สิ้นเปลือง และช่วยลดเวลาในการฝึกอบรมเพราะการใช้ชุดข้อมูลแบบสามมิติจะใช้เวลาในการฝึกอบรมมากกว่าชุดข้อมูลแบบสองมิติ ในการเรียนรู้โครงข่ายสถาปัตยกรรมปัญญาประดิษฐ์คอนโวลูชันแบบสามมิติ (3D CNN) มีความสามารถเรียนรู้จากการป้อนชุดข้อมูลเข้าด้วยขนาดแรกเริ่มเท่าไรและสามารถทำนายข้อมูลออกมาเป็นขนาด voxels เท่าเดิมได้ในขั้นสุดท้ายของการเรียนรู้ โดยจะอธิบายในส่วนของการสถาปัตยกรรมโครงข่ายประสาทเทียม ในส่วนนี้เราจะอธิบายสถาปัตยกรรมโครงข่ายประสาทเทียมเป็นวิธีการหลักที่มีผลต่อการวิเคราะห์การเรียนรู้ ข้อมูลของโครงข่ายประสาทเทียมต่อป้อนข้อมูลและการทำนายผลของโครงข่าย วิธีการเสริมข้อมูลเป็นหลักการทางวิทยาศาสตร์ที่เป็นความน่าสนใจหลักของบทความนี้

สำหรับการสร้างสถาปัตยกรรมยูเน็ตสามมิติ เราจะใช้เฟรมเวิร์กการเรียนรู้เชิงลึกของ TensorFlow ตามที่กล่าวไว้แล้ว ดังนั้น เราจะนำเข้าไลบรารี TensorFlow เพื่อจุดประสงค์นี้ ชุดคำสั่งดังรูป 3.18 เช่นเดียวกับเฟรมเวิร์ก Keras ซึ่งตอนนี้เป็นส่วนสำคัญของโครงสร้างโมเดล TensorFlow จากความเข้าใจก่อนหน้านี้เกี่ยวกับสถาปัตยกรรมยูเน็ต เราทราบว่า การนำเข้าที่จำเป็นบางอย่างรวมถึงเลเยอร์ Convolutional, เลเยอร์ Maxpool, เลเยอร์ Input และฟังก์ชันกระตุ้น ReLU สำหรับโครงสร้างการสร้างแบบจำลองพื้นฐาน จากนั้นเราจะใช้เลเยอร์เพิ่มเติมบางอย่าง เช่น เลเยอร์ Conv3DTranspose ซึ่งจะทำให้การ Upsampling สำหรับบล็อกตัวถอดรหัสที่เราต้องการ นอกจากนี้เราจะใช้เลเยอร์ Concatenate เพื่อรวมด้วยการเชื่อมต่อตรงข้าม

### 3.10 สถาปัตยกรรมโครงข่ายประสาทเทียมยูเน็ตสามมิติ (3D U-Net architecture)

หลังจากนำเข้าไลบรารีที่จำเป็นแล้ว เราสามารถสร้างสถาปัตยกรรมยูเน็ตต่อไปได้ สามารถทำได้ในคลาสเดียวที่สมบูรณ์ โดยกำหนดพารามิเตอร์และค่าทั้งหมดตามลำดับและดำเนินการต่อจนกว่าจะถึงจุดสิ้นสุดหรือทำบล็อกซ้ำสองถึงสามบล็อก ซึ่งการทดลองนี้จะใช้วิธีหลังเนื่องจากสะดวก ผู้ใช้งานเป็นส่วนใหญ่ โดยเสริมบล็อกช่วยไม่ก็บล็อก เราจะใช้บล็อกวนซ้ำสามบล็อกตามที่แสดงในการแสดงสถาปัตยกรรมยูเน็ต ได้แก่ บล็อก convolution บล็อกตัวเข้ารหัส (encoder) และบล็อกตัวถอดรหัส (decoder) บล็อกทั้งสามนี้ในการช่วยเราสามารถสร้างสถาปัตยกรรมยูเน็ตได้

สถาปัตยกรรมโครงข่ายประสาทเทียมสถาปัตยกรรมยูเน็ตสามมิติ ประกอบด้วยสองส่วนหลัก (1) ส่วนการเข้ารหัส (encoder) ฝั่งซ้ายซึ่งงานหลักคือการวิเคราะห์ภาพทั้งหมดและ (2) ส่วนถอดรหัส (decoder) ฝั่งขวาเป็นส่วนขยายที่สร้างการตัดแบ่งส่วนภาพแบบความละเอียดสูงออกมา แต่ละชั้นในส่วน ของ encoder ประกอบไปด้วย Convolution Block ขนาด  $3 \times 3 \times 3$  จำนวน 2 ชั้นแต่ละชั้นตามด้วยฟังก์ชัน กระตุ้น rectified linear unit (ReLU) และหลังจากชั้นก่อนหน้าจะทำการ down-sampling ด้วยชั้น Max Pooling ขนาด  $2 \times 2 \times 2$  เพื่อดึงคุณลักษณะ (features) ออกมาจากภาพ จะเห็นว่าการดึงคุณลักษณะนี้ เกิดขึ้นหลายระดับตั้งแต่ high-resolution features ไปจนถึง low-resolution features ในทำนอง เดียวกันในส่วนของ decoder ประกอบไปด้วยการทำ up-sampling ด้วยการ up-convolution ขนาด  $2 \times 2 \times 2$  โดย stride เท่ากับ 2 สำหรับ ดำเนินการ up-sampling และตามด้วยการทำ Convolution Block ขนาด  $3 \times 3 \times 3$  จำนวน 2 ชั้นแต่ละชั้นตามด้วยฟังก์ชันกระตุ้น rectified linear unit (ReLU) ในขั้นสุดท้าย Convolution Block แบบ  $1 \times 1 \times 1$  จะลดจำนวนช่องสัญญาณเพื่อให้เท่ากับจำนวนข้อมูลนำเข้ามานอกจากนี้เรากำลังหลีกเลี่ยงปัญหาคอขวดทั้งในส่วนของตัวเข้ารหัสและตัวถอดรหัส โดยใช้ค่าสองเท่าของ ช่องก่อนที่จะทำการ max pooling แต่ละครั้ง ภาพประกอบของสถาปัตยกรรมยูเน็ตแบบสามมิติ ที่อธิบาย ไว้ก่อนหน้านี้แสดงในรูปที่ 3.17 และชุดคำสั่งในการสร้างอัลกอริทึมดังภาพที่ 3.20



รูปที่ 3.17 แผนผังสถาปัตยกรรมยูเน็ตสามมิติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

from __future__ import print_function
import os
import keras.models as models
from skimage.transform import resize
from skimage.io import imsave
import numpy as np
import tensorflow as tf
from keras.models import Model
from keras.layers import Input, concatenate, Conv3D, MaxPooling3D, Conv3DTr
anspose, AveragePooling3D, ZeroPadding3D
from keras.callbacks import ModelCheckpoint, CSVLogger
from keras import backend as K
from keras.regularizers import l2
from tensorflow.keras.optimizers import Adam

```

รูปที่ 3.18 โลบรารีสำหรับสร้างชุดคำสั่งโครงข่ายประสาทเทียม

```

K.set_image_data_format('channels_last')

img_rows = 128
img_cols = 128
img_depth = 128

```

รูปที่ 3.19 คำสั่งการอินพุตขนาดภาพ 3 ระนาบ

```

def get_unet():
    inputs = Input(shape=(img_depth, img_rows, img_cols, 1))
    conv1 = Conv3D(32, (3, 3, 3), activation='relu', padding='same')(inputs)
    conv1 = Conv3D(32, (3, 3, 3), activation='relu', padding='same')(conv1)
    pool1 = MaxPooling3D(pool_size=(2, 2, 2))(conv1)

    conv2 = Conv3D(64, (3, 3, 3), activation='relu', padding='same')(pool1)
    conv2 = Conv3D(64, (3, 3, 3), activation='relu', padding='same')(conv2)
    pool2 = MaxPooling3D(pool_size=(2, 2, 2))(conv2)

    conv3 = Conv3D(128, (3, 3, 3), activation='relu', padding='same')(pool2)
    conv3 = Conv3D(128, (3, 3, 3), activation='relu', padding='same')(conv3)
    pool3 = MaxPooling3D(pool_size=(2, 2, 2))(conv3)

    conv4 = Conv3D(256, (3, 3, 3), activation='relu', padding='same')(pool3)
    conv4 = Conv3D(256, (3, 3, 3), activation='relu', padding='same')(conv4)
    pool4 = MaxPooling3D(pool_size=(2, 2, 2))(conv4)

    conv5 = Conv3D(512, (3, 3, 3), activation='relu', padding='same')(pool4)
    conv5 = Conv3D(512, (3, 3, 3), activation='relu', padding='same')(conv5)

    up6 = concatenate([Conv3DTranspose(256, (2, 2, 2), strides=(2, 2, 2), padding='same')(conv5), conv4), axis=4)
    conv6 = Conv3D(256, (3, 3, 3), activation='relu', padding='same')(up6)
    conv6 = Conv3D(256, (3, 3, 3), activation='relu', padding='same')(conv6)

    up7 = concatenate([Conv3DTranspose(128, (2, 2, 2), strides=(2, 2, 2), padding='same')(conv6), conv3), axis=4)
    conv7 = Conv3D(128, (3, 3, 3), activation='relu', padding='same')(up7)
    conv7 = Conv3D(128, (3, 3, 3), activation='relu', padding='same')(conv7)

    up8 = concatenate([Conv3DTranspose(64, (2, 2, 2), strides=(2, 2, 2), padding='same')(conv7), conv2), axis=4)
    conv8 = Conv3D(64, (3, 3, 3), activation='relu', padding='same')(up8)
    conv8 = Conv3D(64, (3, 3, 3), activation='relu', padding='same')(conv8)

    up9 = concatenate([Conv3DTranspose(32, (2, 2, 2), strides=(2, 2, 2), padding='same')(conv8), conv1), axis=4)
    conv9 = Conv3D(32, (3, 3, 3), activation='relu', padding='same')(up9)
    conv9 = Conv3D(32, (3, 3, 3), activation='relu', padding='same')(conv9)

    conv10 = Conv3D(1, (1, 1, 1), activation='sigmoid')(conv9)

```

รูปที่ 3.20 ชุดคำสั่งการสร้างสถาปัตยกรรมยูเน็ตสามมิติ

### 3.11 เริ่มฝึกชุดข้อมูลที่เตรียมไว้สำหรับฝึกการเรียนรู้

เมื่อเราได้สร้างชุดคำสั่งสำหรับการฝึกการเรียนรู้ด้วยการคอนโวลูชันหลายชั้นแล้ว ในขั้นตอนถัดไป เราจะกำหนดการแบ่งระหว่างชุดข้อมูลการฝึกอบรมและชุดข้อมูลสำหรับการตรวจสอบ (validation set) ตามลำดับ เราได้ดำเนินการในขั้นตอนนี้เพื่อให้แน่ใจว่าไม่มีความผิดพลาดเกิดขึ้นระหว่างองค์ประกอบที่มีความสมบูรณ์ในชุดฝึกอบรมและชุดทดสอบ ข้อมูลจะแบ่งออกเป็นสองส่วนเพื่อไม่ให้โมเดลได้รับรู้ข้อมูลสำหรับการทดสอบโมเดล ในชุดข้อมูลการตรวจสอบนั้นเราจะดำเนินการสุ่มจากชุดข้อมูลทั้งหมดในชุดสำหรับฝึกการเรียนรู้ และเราสามารถรับตัวอย่างแบบสุ่มสำหรับทั้งฝึกการเรียนรู้และชุดข้อมูลการตรวจสอบความถูกต้อง เมื่อเราเริ่มฝึกและทดสอบโครงข่ายประสาท สิ่งสำคัญคือต้องแยกข้อมูลการฝึกออกจากการตรวจสอบความถูกต้อง เพื่อที่จะได้ชุดตรวจสอบความถูกต้องของโมเดลที่ไม่ใช่ชุดข้อมูลเดียวกันกับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ที่ใช้ในการฝึก เพื่อช่วยลดการเกิด overfitting เราจะใช้พารามิเตอร์ validation\_split = 0.10 คือการแบ่งชุดข้อมูลการตรวจสอบออกจากชุดฝึกอบรวม 10 เปอร์เซ็นต์ เมื่อปรับแต่งโมเดลของเราเพื่อแยกข้อมูลออกเป็นชุดการฝึกและชุดการตรวจสอบความถูกต้องโดยอัตโนมัติ และใช้เพื่อตรวจสอบความถูกต้องของโมเดลของเราโดยการทดลองของเรา จากนั้นเราจะเรียกค่าการฝึกอบรวมและค่าการตรวจสอบที่ละรายการและจัดเก็บไว้ในตัวแปรที่เกี่ยวข้อง

จากนั้นเราได้กำหนดฟังก์ชันการวัดสำหรับวัดการสูญเสีย และวัดประสิทธิภาพของโมเดลที่ใช้ในการฝึกการเรียนรู้ เราจะใช้ค่าสัมประสิทธิ์ Dice coefficient score เป็นตัววัดประสิทธิภาพ ฟังก์ชันการสูญเสียเป็น Dice coefficient loss โดยชุดคำสั่งที่ใช้สำหรับค่าวัดสัมประสิทธิ์และวัดการสูญเสียดังภาพที่ 3.21 และโมเดลถูกคอมไพล์ด้วยเครื่องมือเพิ่มประสิทธิภาพ Adam Optimizer ด้วยอัตราการการเรียนรู้เริ่มต้น 0.0001 โดยมีขนาด batch size เป็น 1 เราฝึกการเรียนรู้ของโมเดลจำนวน 200 ครั้ง

```
def dice_coefficient(y_true, y_pred):
    intersection = K.sum(K.abs(y_true * y_pred), axis=[-3,-2,-1])
    dn = K.sum(K.square(y_true) + K.square(y_pred), axis=[-3,-2,-1]) + 1e-8
    return K.mean(2 * intersection / dn, axis=[0,1])

def loss_gt(e=1e-8):
    def loss_gt_(y_true, y_pred):
        intersection = K.sum(K.abs(y_true * y_pred), axis=[-3,-2,-1])
        dn = K.sum(K.square(y_true) + K.square(y_pred), axis=[-3,-2,-1]) + e
        return 1 - K.mean(2 * intersection / dn, axis=[0,1])
    return loss_gt_
```

รูปที่ 3.21 อัลกอริทึมค่าสัมประสิทธิ์ Dice coefficient ฟังก์ชันการสูญเสีย Dice coefficient loss

หลังจากกำหนดทุกอย่างแล้ว เราจะทำการใช้ชุดคำสั่งเพื่อรวบรวมชุดคำสั่งต่างๆ ที่เราได้สร้างไว้เพื่อติดตั้งข้อมูลเพื่อใช้ในการฝึกเรียนรู้และการตรวจสอบข้อมูลกับชุดคำสั่งที่เราได้เตรียมไว้ โดยมีภาพชุดคำสั่งสำหรับการรวบรวมชุดคำสั่งที่ได้กล่าวข้างต้นดังแสดงในภาพที่ 3.22 ในส่วนของ Callback เราใช้ ModelCheckpointCallback ในรูปแบบ csv\_logger ใช้เพื่อบันทึกน้ำหนัก (แบบจำลอง) ในโฟลเดอร์บันทึกหลังการฝึกตามจำนวนของ epoch และบันทึกความแม่นยำและอัตราการสูญเสียในแต่ละ epoch ด้วยวิธีนี้

```
# split data in train and test sets.
train_img, train_mask, test_img, test_mask = split_data()
model.compile(optimizer=Adam(lr=0.0001), loss=[loss_gt(1e-8)], metrics=[dice_coefficient, 'accuracy'])

# fit data to model
history = model.fit(train_img, train_mask,
                    batch_size=1,
                    epochs=200,
                    verbose=1,
                    shuffle=True,
                    validation_split=0.10,
                    callbacks=[model_checkpoint, csv_logger])
```

### รูปที่ 3.22 อัลกอริทึมการเตรียมฝึกการเรียนรู้

เมื่อโครงข่ายประสาทเทียมยูเน็ตสามมิติได้รับการฝึกการเรียนรู้แล้ว โดยใช้ตัวประเมินการสูญเสียและประเมินความถูกต้อง ซึ่งข้อมูลเราได้บันทึกเป็นค่าน้ำหนัก (Weight) ของโมเดลที่ผ่านการฝึกการเรียนรู้แล้วโดยการทำให้เป็นลำดับในแต่ละ epoch ที่เราใช้ในการฝึก และนำค่าที่ใช้สำหรับการประเมินการเรียนรู้ทั้งค่าสัมประสิทธิ์ Dice coefficient score ที่เป็นตัววัดประสิทธิภาพ และฟังก์ชันการสูญเสียเป็น Dice coefficient loss มาบันทึกเป็นกราฟการเรียนรู้ว่าระหว่างการฝึกอบรมมีข้อบ่งชี้ว่าโมเดลทำงานได้ดีทั้งในชุดข้อมูลการฝึกอบรมและชุดการตรวจสอบความถูกต้องหรือไม่ ต้องปรับปรุงอย่างไร เกิด overfitting หรือไม่เราสามารถสังเกตได้จากกราฟการเรียนรู้นี้ โดยจะแสดงภาพในบทความสรุปผลการทดลอง ซึ่งบ่งชี้ได้ว่าโมเดลมีภาพรวมที่ดี ไม่มีการปรับมากเกินไป

### 3.12 การทำนายด้วยสถาปัตยกรรมยูเน็ต U-Net

หลังจากบันทึกโมเดลที่ผ่านการฝึกอบรมและบันทึกไว้ แล้วเราจะประเมินประสิทธิภาพของโมเดลในชุดข้อมูลสำหรับทดสอบด้วยค่าเวลาที่บันทึกไว้ในรูปแบบไฟล์ .h5 โดยเราจะใช้ฟังก์ชัน `model.load_weights` ในการเรียกค่าน้ำหนักที่ได้จากการเรียนรู้มาประเมินด้วยชุดทดสอบที่โมเดลไม่เคยเห็นมาก่อน และใช้ในการทำนายผล การประเมินเป็นกระบวนการระหว่างการประเมินประสิทธิภาพของโมเดลเพื่อตรวจสอบว่าโมเดลนั้นเหมาะสมที่สุดสำหรับการทดสอบที่กำหนดและข้อมูลที่เกี่ยวข้องหรือไม่ ด้วยฟังก์ชันของ Keras จะใช้สำหรับการประเมินด้วยฟังก์ชัน `model.evaluate` ทำการประเมินโมเดล โดยใช้ค่าในการวัดประสิทธิภาพเช่นเดียวกับการเรียนรู้โมเดลด้วยชุดคำสั่งดังภาพที่ 3.23 หลังจากการประเมินเราได้กำหนดฟังก์ชันเพื่อแสดงภาพการทำนายโดยโมเดลการทำนายเป็นขั้นตอนสุดท้ายและผลลัพธ์ที่เราคาดหวังจากการสร้างโมเดล Keras ในวิธีการทำนายเพื่อรับการทำนายของแบบจำลองที่ผ่านการฝึกการเรียนรู้ ด้วยฟังก์ชันการทำนาย `model.predict` ฟังก์ชันต้องอินพุตเป็นอาร์เรย์และเอาต์พุตอาร์เรย์และการ

คาดคะเน จากนั้นเราได้กำหนดขนาดตัวเลขและลงจุดทั้งสาม นั่นคือ รูปภาพ มาสก์ และมาสก์ที่คาดการณ์ไว้ โดยฟังก์ชันที่ใช้ในการประเมินและการแสดงผลการทำนายดังแสดงในภาพ 3.24

หลังจากที่ได้ข้อมูลการทำนายของโมเดลจากชุดทดสอบ และข้อมูลที่ได้นั้นเป็นลักษณะโวลุ่ม 3 มิติ เราได้นำข้อมูลนี้มาสร้างเป็นรูปทรงของหลอดเลือด โดยที่เราต้องแปลงข้อมูลของอาร์เรย์เป็นข้อมูลภาพ NIFTI ไฟล์เสียก่อนด้วยชุดคำสั่งที่ 3.25 และนำไฟล์ไปสร้างรูปทรงด้วยโปรแกรม Matlab ด้วยแอปพลิเคชัน Volume Viewer ดังรูป 3.26

```
model.load_weights('/content/drive/MyDrive/Colab Notebooks/weightrune
t/3D-unet Segmentation of aaa.h5')

score = model.evaluate(test_img, test_mask, batch_size=1, verbose=1)
print(model.metrics_names)
print(f'Test loss: {score[0]} / Test DICE: {score[1]}')
```

รูปที่ 3.23 อัลกอริทึมชุดทดสอบโมเดลเพื่อประเมินประสิทธิภาพ

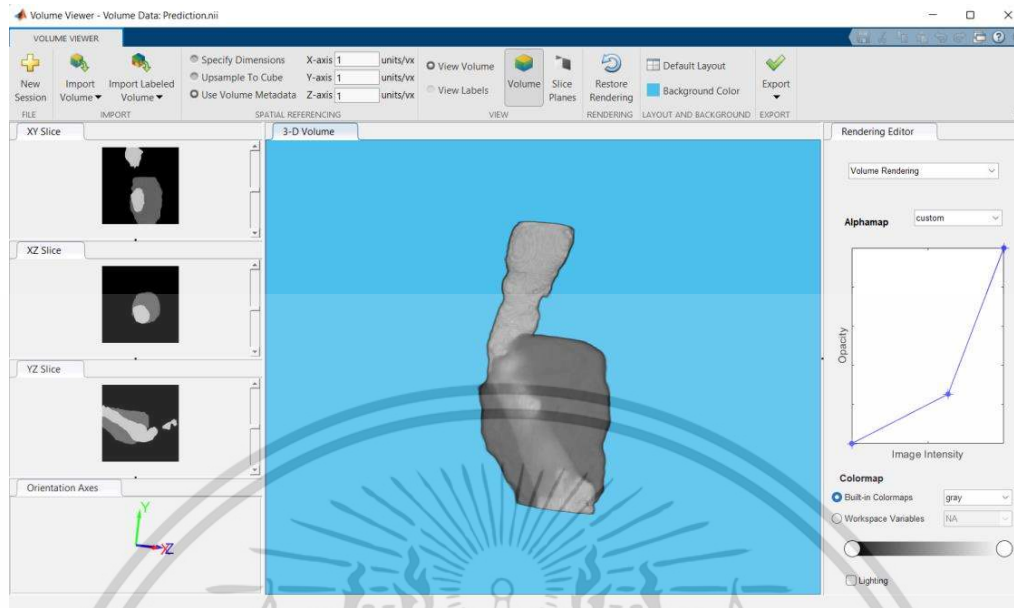
```
test_mask_image = model.predict(test_img, verbose=1, batch_size=1)
np.save('imgs_mask_testunet.npy', test_mask_image)
```

รูปที่ 3.24 อัลกอริทึมชุดทดสอบโมเดลที่ฝึกการเรียนรู้สำเร็จ

```
pred_seg = np.expand_dims(imgs_testd,0)[0]

images = pred_seg[1]
#image = np.squeeze(images, axis=3)
ni_img = nib.Nifti1Image(images, affine=np.eye(4))
nib.save(ni_img, "test_mask.nii")
```

รูปที่ 3.25 อัลกอริทึมแปลงไฟล์อาร์เรย์เป็นไฟล์ภาพ NIFTI



รูปที่ 3.26 การสร้างรูปสามมิติจากไฟล์ภาพที่ได้จากการทำนายของโมเดล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 4

### ผลการทดลอง

จากการสอนโมเดลด้วยการเรียนรู้เชิงลึกด้วยสถาปัตยกรรมยูเน็ต ในการฝึกการเรียนรู้และทำนายภาพในรูปแบบไบนามารัสก์เพื่อนำมาสร้างรูปทรงของหลอดเลือดแดงใหญ่ส่วนท้องที่มีทั้งลักษณะปกติ มีการโป่งพอง และการโป่งพองร่วมกับมีลิ้มเลือดที่ผนังหลอดเลือดทั้งบางและหนา จากภาพเอกซเรย์คอมพิวเตอร์หลอดเลือดแดงใหญ่

ในงานนี้ได้ทำการทดลองการเรียนรู้เชิงลึกโดยใช้ไลบรารีการเรียนรู้จาก Keras และ TensorFlow บนแพลตฟอร์ม NVIDIA-SMI-Persistence-M ด้วยฮาร์ดแวร์เร่งความเร็วด้วย GPU ด้วยภาษาไพทอน บนแพลตฟอร์มที่เป็นการสนับสนุนคุณสมบัติการทำงานร่วมกันของ Google Colaboratory โครงข่ายประสาทเทียมได้รับการฝึกการเรียนรู้และทำนายผลโดยใช้ชุดข้อมูล 85 เปอร์เซ็นต์ สำหรับการฝึกการเรียนรู้ 10 เปอร์เซ็นต์ สำหรับการตรวจสอบความถูกต้องและ 5 เปอร์เซ็นต์ สำหรับการทดสอบ จากชุดข้อมูลที่มีการเพิ่มปริมาณจากการใช้ฟังก์ชันเสริม Data Augmentation จากแพตช์ที่ได้สร้างขึ้นเพื่อจัดเก็บชุดข้อมูล

#### 4.1 ผลการทดลองการใช้สถาปัตยกรรม 3D U-Net

ผลการทดลองประเมินประสิทธิภาพด้วยตัวชี้วัดค่าสัมประสิทธิ์ Dice coefficient score ที่เป็นตัววัดประสิทธิภาพ และฟังก์ชันการสูญเสีย Dice coefficient loss ซึ่งเป็นตัวประเมินการแบ่งส่วนที่ใช้บ่อยที่สุดด้วยสถาปัตยกรรมยูเน็ตสามมิติ ผลการทดลองแสดงไว้ในตารางที่ 4.1 เราได้ผลลัพธ์จากการฝึกการเรียนรู้ด้วยชุดการฝึกและชุดตรวจสอบความถูกต้องและรูปแบบการประเมินด้วยชุดข้อมูลการทดสอบที่ไม่เคยเห็นมาก่อนเพื่อวัดประสิทธิภาพการทำนายของโมเดล จากภาพเอกซเรย์คอมพิวเตอร์ที่จำแนกเป็นหลอดเลือดปกติ ปริมาณหลอดเลือดที่โป่งพอง และหลอดเลือดโป่งพองที่มีลิ้มเลือดที่ผนังหลอดเลือด ผลการประเมินของการแบ่งส่วนจะถูกเปรียบเทียบกับมารัสก์ที่ใช้สร้างขึ้นด้วยการทำด้วยมือเพื่อใช้สำหรับการเรียนรู้

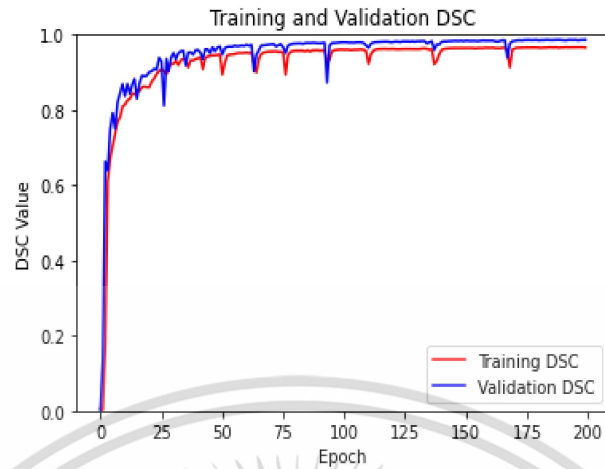
ตารางที่ 4.1 ผลการทดลองจากสถาปัตยกรรมยูเน็ตสามมิติ

Model Architecture	Time for Training	Model Performance		
		Dice Coefficient Score		
		Training Set	Validation Set	Test Set
3D U-Net	1 h 30 min	0.9669	0.9866	0.9868

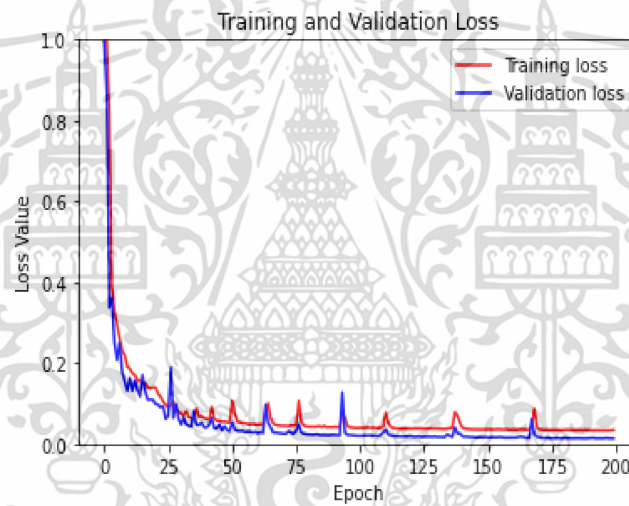
จากตารางที่ 4.1 แสดงถึงผลการทดลองจากการประเมินความถูกต้องของโมเดลซึ่งในที่นี้จะกล่าวถึงค่าสัมประสิทธิ์ Dice Coefficient Score โดยฝึกการเรียนรู้ของโมเดลจำนวน 200 ครั้ง ใช้เวลาในการฝึก 1 ชั่วโมง 30 นาที ความถูกต้องที่ได้จากชุดฝึกการเรียนรู้โดยเป็นค่าที่ได้สูงที่สุดคือ 0.9669 และในชุดตรวจสอบความถูกต้องได้ค่าสูงที่สุดคือ 0.9866 ค่าความสูญเสียคือ 0.0331 และ 0.0134 ตามลำดับ ซึ่งมีค่าใกล้เคียงศูนย์บ่งบอกได้ว่าโมเดลสามารถเรียนรู้ได้ดี และประเมินด้วยชุดทดสอบโมเดลได้ความถูกต้องคือ 0.9868 และค่าความสูญเสียคือ 0.0132

การเรียนรู้ในแต่ละครั้งที่ทำการฝึกโมเดลเมื่อนำมาสร้างเป็นกราฟการเรียนรู้ที่แสดงทั้งประสิทธิภาพและการสูญเสียของโมเดล ของชุดข้อมูลที่ใช้ฝึกการเรียนรู้แทนด้วยเส้นสีแดงและชุดตรวจสอบความถูกต้องแทนด้วยเส้นสีน้ำเงิน ดังแสดงในรูปที่ 4.1 โดยกราฟที่แสดงถึงประสิทธิภาพค่าสัมประสิทธิ์ความคล้าย DSC ในแนวแกนตั้ง โดยมีค่าตั้งแต่ 0 ถึง 1 และจำนวนการฝึกในแนวแกนนอนจำนวน 200 ครั้ง จะเห็นได้ว่าการเรียนรู้ของโมเดลสำหรับการแบ่งภาพหลอดเลือดทั้งสามชนิดนั้นโดยค่าเข้าใกล้ 0 ช่วงของการเรียนรู้เริ่มแรกและเพิ่มขึ้นเรื่อย ๆ เมื่อการฝึกผ่านไป 24 ครั้งประสิทธิภาพเพิ่มที่ได้มากกว่า 0.9 และค่าประสิทธิภาพการเรียนรู้เพิ่มขึ้นตลอดการเรียนรู้จนค่าเข้าใกล้ 1 ดังแสดงในรูปที่ 4.1 (a) ส่วนของกราฟที่แสดงถึงค่าการสูญเสียของโมเดลแทนด้วยแกนตั้ง ค่าการสูญเสียของโมเดลโดยมีค่าตั้งแต่ 0 ถึง 1 ซึ่งค่าพารามิเตอร์อื่นๆ จะเหมือนข้อมูลกราฟข้างต้น ดังแสดงในรูปที่ 4.1 (b) จะเห็นได้ว่าค่าการสูญเสียของโมเดลมีค่าลดลงเข้าใกล้ 0 อย่างรวดเร็วทั้งชุดข้อมูลการเรียนรู้และชุดการตรวจสอบความถูกต้องโมเดล

เมื่อเราได้วัดประสิทธิภาพจากชุดการทดสอบแล้ว ซึ่งมีค่าความถูกต้องมากกว่า 0.9 ค่าใกล้ 1 มาก และได้นำผลภาพที่ได้จากการทำนายของโมเดลที่มีประสิทธิภาพที่สูง เพื่อสังเกตว่าชุดข้อมูลที่ได้จากการทำนายจากชุดทดสอบประสิทธิภาพมีความคล้ายเมื่อนำมาเปรียบเทียบกับมาส์กต้นแบบของชุดทดสอบ ผลการทำนายดังแสดงภาพที่ 4.2 โดยเราได้เลือกผลการทำนายจากชุดทดสอบมาบางส่วนในการแสดงผลในที่นี้ โดยจะเน้นชุดข้อมูลหลอดเลือดแดงที่โป่งที่มีลิ้มเลือดที่ผนังหลอดเลือดร่วมด้วย เพราะเป็นการทดสอบความสามารถของโมเดลอย่างมาก ในการทำนายชุดข้อมูลชนิดนี้ ในส่วนของเส้นเลือดแดงปกติ และเส้นเลือดแดงที่โป่งพองมีประสิทธิภาพการทำนายที่ดีมาก



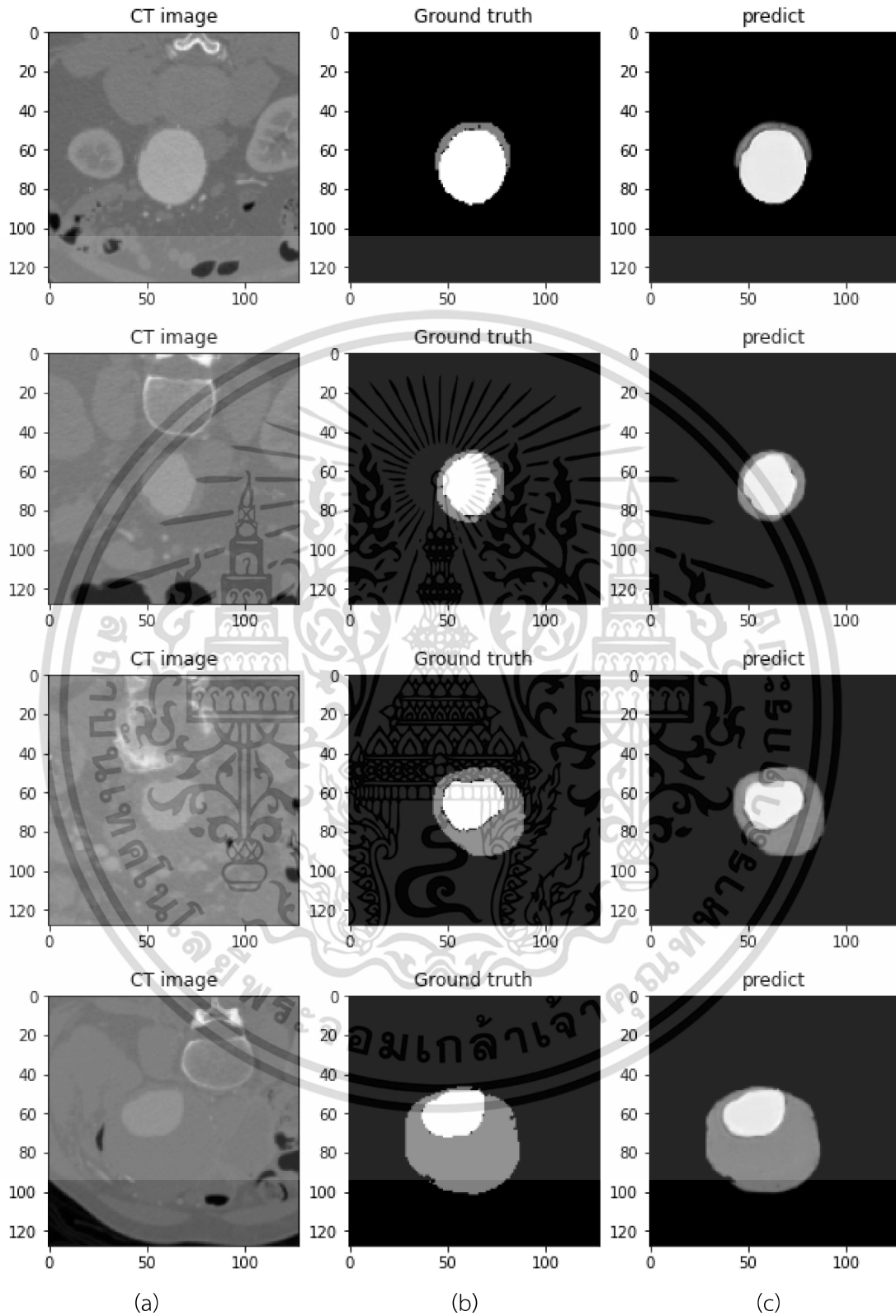
(a)



(b)

รูปที่ 4.1 กราฟแสดงผลจากการเรียนรู้ของโมเดล (a) แสดงประสิทธิภาพของโมเดล (b) การสูญเสียของโมเดล

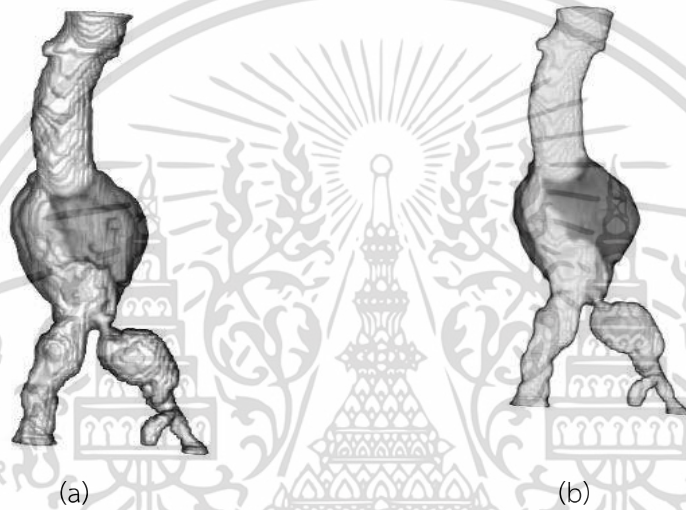
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.2 แสดงภาพสองมิติบางส่วน (a) ภาพ CT (b) ภาพมาสก์ต้นแบบการเรียนรู้ (c) ภาพการทำนายของโมเดล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หลังจากที่ได้ข้อมูลการทำนายของโมเดลจากชุดทดสอบ และข้อมูลที่ได้นั้นเป็นลักษณะโวลุ่ม 3 มิติ เราได้นำข้อมูลนี้มาสร้างเป็นรูปทรงของหลอดเลือด โดยที่เราต้องแปลงข้อมูลอาร์เรย์เป็นข้อมูลภาพ NIFTI ไฟล์เสียก่อน และนำไฟล์ที่ได้ไปสร้างรูปทรงด้วยโปรแกรม Matlab ดังรูป 4.3



รูปที่ 4.3 แสดงภาพรูปทรงสามมิติ (a) สร้างจากมาส์กต้นแบบการเรียนรู้ (b) สร้างจากการทำนายของโมเดล

## บทที่ 5

# สรุปและอภิปรายผลการทดลอง

### 5.1 สรุปผลการทดลอง

จากผลการทดลองเป็นการนำเสนอประสิทธิภาพของโมเดลการเรียนรู้เชิงลึกแบบอัตโนมัติ ด้วยสถาปัตยกรรมยูเน็ตสามมิติ สำหรับการแบ่งส่วนหลอดเลือดแดงใหญ่ในช่องท้อง ซึ่งประกอบด้วยชุดข้อมูลสำหรับการฝึกการเรียนรู้ ชุดข้อมูลสำหรับการตรวจสอบความถูกต้องของโมเดล และชุดข้อมูลสำหรับทดสอบประสิทธิภาพ โดยมีหลอดเลือด 3 ประเภทที่ได้กล่าวมาข้างต้น ได้แก่ หลอดเลือดแดงแบบปกติ หลอดเลือดแดงที่การโป่งพอง และหลอดเลือดแดงที่มีการโป่งพองที่มี ILT จากภาพเอกซเรย์คอมพิวเตอร์หลอดเลือด

จากการทดลองแสดงให้เห็นว่าการใช้ฟังก์ชันเสริมเพื่อเพิ่มชุดข้อมูลสามารถเพิ่มเป็น 300 ชุดข้อมูล จากชุดข้อมูลเดิมมี 60 ชุด โดยประเภทคลาสข้อมูลเป็นแบบไม่สมดุล (Class Imbalance) ซึ่งจะมีชุดข้อมูลประเภทหลอดเลือดโป่งพองที่มี ILT มีจำนวนมากที่สุด ที่ทำให้โมเดลสามารถทำนายภาพมีประสิทธิภาพดีที่สุด โดยมีค่าความถูกต้อง 0.9868 ด้วยการใช้สถาปัตยกรรมยูเน็ตแบบสามมิติจากชุดทดสอบประสิทธิภาพจากการฝึกการเรียนรู้มีความถูกต้อง 0.9669 และความถูกต้องในการตรวจสอบที่ 0.9866 แสดงในตารางที่ 1.

จากรูปภาพการฝึกการเรียนรู้ของโมเดลจะเห็นได้ว่าประสิทธิภาพการฝึกจะมีค่าน้อยเมื่อเริ่มการฝึกช่วงแรก และเพิ่มขึ้นเมื่อการฝึกมากกว่า 24 ครั้ง และเพิ่มขึ้นเรื่อยจนค่าเค้ใกล้เคียง 1 ที่จำนวนการฝึกมากขึ้นสามารถสังเกตจากรูปภาพการเรียนรู้ของกระบวนการฝึก และฟังก์ชันการสูญเสียจากรูปภาพจะเห็นว่ากราฟลดลงเรื่อย ๆ ตั้งแต่เริ่มการฝึก แต่จากรูปภาพทั้งสองจะเห็นได้ว่าชุดตรวจสอบความถูกต้องมีค่าความถูกต้องมากกว่า และมีค่าการสูญเสียที่น้อยกว่าชุดฝึกการเรียนรู้

การทดลองนี้เราได้รับผลลัพธ์จากการทำนายภาพจากโมเดลเป็นรูปแบบ 3 มิติ และสามารถสังเกตการทดลองแต่ละภาพตัดขวางในรูปแบบภาพ 2 มิติ จะเห็นได้ว่าโมเดลมีความสามารถทำนายการแบ่งส่วนภาพของหลอดเลือดที่มีสารทึบรังสีออกมาได้อย่างมีประสิทธิภาพ ทั้งรูปแบบหลอดเลือดปกติและหลอดเลือดที่มีการโป่งพอง ในส่วนของ ILT การแบ่งส่วนมีประสิทธิภาพดีและใกล้เคียงกับมาตรฐานเมื่อเปรียบเทียบกัน ทั้ง ILT ที่มีขนาดบางและหนา ในส่วนของการแสดงผลข้อมูลแบบรูปทรง 3 มิติ เห็นได้ชัดว่าการแบ่งส่วนนั้นมีความสามารถทำนาย ILT ที่บางและหนาได้อย่างมีประสิทธิภาพ ด้วยกระบวนการแบ่งส่วนแบบอัตโนมัติด้วยการใช้สถาปัตยกรรมยูเน็ต

### 5.2 ข้อจำกัดในการศึกษา

ข้อจำกัดในการทำงานของเราคือการลดขนาดภาพเอกซเรย์คอมพิวเตอร์ให้เหลือส่วนที่เราสนใจ แต่ละชุดข้อมูลแบบทำด้วยมือ เพื่อเพิ่มความเร็วในการฝึกการเรียนรู้ด้วยสถาปัตยกรรมโครงข่ายประสาทเทียม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับการแบ่งภาพ AAA แบบอัตโนมัติ และช่วยลดหน่วยความจำของระบบ เพราะหน่วยความจำมีข้อจำกัดของแพลตฟอร์ม และคาดหวังว่างานในอนาคตจะสามารถทำการฝึกการเรียนรู้และทดสอบโดยไม่มีข้อจำกัดขอบเขตชุดข้อมูล โดยจะใช้ซอฟต์แวร์ด้วยความเร็วแบบ GPU ที่ความเร็วสูงกว่างานปัจจุบัน ข้อจำกัดอีกประการหนึ่งคือเกี่ยวกับค่าใช้จ่ายในการใช้ความเร็ว GPU ของแพลตฟอร์ม Google Colab ซึ่งมีค่าใช้จ่ายสูง

### 5.3 บทสรุปข้อเสนอแนะ

โรคหลอดเลือดแดงใหญ่ในช่องท้องโป่งพอง (AAA) เป็นโรคที่มีความอันตรายถึงชีวิต ซึ่งเกิดจากการขยายตัวของส่วนล่างของหลอดเลือดใหญ่ที่ส่งเลือดไปเลี้ยงร่างกาย การคัดกรอง AAA เบื้องต้นจะใช้การทำอัลตราซาวด์ (US) เอกซเรย์คอมพิวเตอร์ (CT) หรือ เครื่องสร้างภาพด้วยสนามแม่เหล็กไฟฟ้า (MRI) เพื่อประเมินระดับของ AAA การศึกษานี้เรานำเสนอการแบ่งส่วน AAA ด้วยวิธีการอัตโนมัติ จากภาพเอกซเรย์คอมพิวเตอร์หลอดเลือด โดยใช้โครงข่ายประสาทเทียมรูปแบบสถาปัตยกรรมยูเน็ตสามมิติ ซึ่งสามารถใช้เป็นเครื่องมือช่วยประเมินขนาด รูปทรงของ AAA ได้อย่างแม่นยำ และเพิ่มประสิทธิภาพด้วยการใช้ฟังก์ชันการเสริมข้อมูล ได้ค่าความถูกต้อง DSC 0.9868 ในชุดทดสอบข้อมูลและผลงานของเรามีประสิทธิภาพที่ดีได้ผลลัพธ์ที่ชัดเจนที่จากการทำนายภาพหลังการฝึกการเรียนรู้ โมเดลสามารถปรับปรุงความถูกต้องได้ด้วย การขยายเวลาสำหรับโมเดลการฝึก เพิ่มการใช้ Dropouts เพื่อช่วยให้โครงข่ายลดเวลาในการฝึกได้ และการใช้ไฮเปอร์พารามิเตอร์อื่นๆ ร่วมด้วย โมเดลจากสถาปัตยกรรมยูเน็ตมีความถูกต้องสูงกว่า 90% เนื่องจากมาส์กต้นแบบสำหรับการเรียนรู้มีคุณภาพสูงเพราะจัดทำโดยนักรังสีการแพทย์ผู้ชำนาญการ และจำนวนข้อมูลเพียงพอสำหรับการเรียนรู้และใช้ฟังก์ชันการเสริมข้อมูลเพื่อเพิ่มปริมาณของข้อมูลเพื่อช่วยลดการเกิด overfitting ของโมเดล แต่ก็มีความเสี่ยงที่จะเกิด overfitting จากการเสริมข้อมูลได้เหมือนกัน เช่นการทดลองของเราจะเห็นว่าประสิทธิภาพชุดการตรวจสอบข้อมูลนั้นมีค่าความถูกต้องมากกว่าชุดฝึกการเรียนรู้ อาจเกิดการเหตุผลนี้ได้เช่นกัน แต่สามารถแก้ไขได้ด้วยการใช้ Dropouts หรือใช้ฟังก์ชันการเสริมข้อมูลแบบสุ่ม และการใช้ early stopping ซึ่งเป็นเทคนิคที่ใช้กันอย่างแพร่หลายเพื่อลดปัญหานี้

## บรรณานุกรม

- [1] หลอดเลือดแดงใหญ่ในช่องท้องโป่งพอง (Abdominal Aorta Aneurysm) [ออนไลน์]. 2563 ,มกราคม 13; Available from: <http://www.lph.go.th/inforlphng/index.php/news/or/mcvt/24-evar>
- [2] โครงข่ายประสาทเทียม (Artificial Neuron Network: ANN) [ออนไลน์]. 2555 ,กรกฎาคม 4; Available from: <https://www.gotoknow.org/posts/163433>
- [3] โครงข่ายประสาทเทียม (Artificial Neuron Network: ANN) [ออนไลน์]. 2565 ,พฤศจิกายน 15 Available from : <https://csit.nu.ac.th/kraisak/ds/ds/chapter07/Chapter07.pdf>
- [4] ทฤษฎีโครงข่ายประสาทเทียมและการประยุกต์ใช้ [ออนไลน์]. 2565 ,พฤศจิกายน 18 Available from : [https://kb.psu.ac.th/psukb/bitstream/2553/2132/8/284819\\_ch3.pdf](https://kb.psu.ac.th/psukb/bitstream/2553/2132/8/284819_ch3.pdf)
- [5] Liu, Xiangbin, Liping Song, Shuai Liu, and Yudong Zhang. 2021. “A Review of Deep-Learning-Based Medical Image Segmentation Methods.” *Sustainability* 13, no. 3: 1224.
- [6] ทฤษฎีที่เกี่ยวข้องงานวิจัยที่เกี่ยวข้อง [ออนไลน์]. 2565 ,ธันวาคม 10 Available from : <http://digital.csmsu.net:8080/library/bitstream/123456789/137/7/07.%E0%B8%9A%E0%B8%97%E0%B8%97%E0%B8%B5%E0%B9%882%E0%B8%97%E0%B8%A4%E0%B8%A9%E0%B8%8E%E0%B8%B5%E0%B9%81%E0%B8%A5%E0%B8%B0%E0%B8%87%E0%B8%B2%E0%B8%99%E0%B8%A7%E0%B8%B4%E0%B8%88%E0%B8%B1%E0%B8%A2%E0%B8%97%E0%B8%B5%E0%B9%88%E0%B9%80%E0%B8%81%E0%B8%B5%E0%B9%88%E0%B8%A2%E0%B8%A7%E0%B8%82%E0%B9%89%E0%B8>
- [7] Convolutional Neural Network [ออนไลน์]. 2563 ,มกราคม 28; Available from: <https://medium.com/@hadee2531earvesdrop/convolutional-neural-network-%E0%B8%84%E0%B8%B7%E0%B8%AD%E0%B8%AD%E0%B8%B0%E0%B9%84%E0%B8%A3-42c45f7ec16b>
- [8] Convolutional Neural Network [ออนไลน์]. 2563 ,มกราคม; Available from: <https://guopai.github.io/ml-blog19.html>
- [9] ชั้นการเชื่อมโยงเต็มรูปแบบ (Fully Connected Layer) [ออนไลน์]. 2559; Available from: [https://dric.nrct.go.th/direct\\_download.php?bibid=311229&doc\\_id=784079&name\\_dw=6\\_Chap\\_2.pdf](https://dric.nrct.go.th/direct_download.php?bibid=311229&doc_id=784079&name_dw=6_Chap_2.pdf)

- [10] Convolutional neural networks for segmentation [ออนไลน์]. 2562 ,มิถุนายน 6; Available from: <https://divamgupta.com/image-segmentation/2019/06/06/deep-learning-semantic-segmentation-keras.html>
- [11] Image Segmentation Using Deep Learning: A Survey [ออนไลน์]. 2563 ,พฤษภาคม 11; Available from: <https://medium.com/swlh/image-segmentation-using-deep-learning-a-survey-e37e0f0a1489>
- [12] CT Scan [ออนไลน์]. 2565 ,ธันวาคม 20; Available from: <https://www.pobpad.com/ct-scan>
- [13] การตรวจด้วยเครื่องสร้างภาพด้วยสนามแม่เหล็กไฟฟ้า (MRI) [ออนไลน์]. 2565 ,ธันวาคม 20; Available from: <https://www.rama.mahidol.ac.th/aimc/th/content/09122017-1414-th>
- [14] เอ็มอาร์ไอ (Magnetic Resonance Imaging : MRI) [ออนไลน์]. 2563 กรกฎาคม 20 ; Available from: <https://www.synphaet.co.th/%E0%B9%80%E0%B8%AD%E0%B9%87%E0%B8%A1%E0%B8%AD%E0%B8%B2%E0%B8%A3%E0%B9%8C%E0%B9%84%E0%B8%AD-magnetic-resonance-imaging-mri/>
- [15] Digital Radiography Systems [ออนไลน์]. 2565 ,ธันวาคม 20; Available from: <https://sites.google.com/site/medicaldigitalimage/home/digital-radiography-systems>
- [16] Medical Image Analysis with Deep Learning [ออนไลน์]. 2566 ,มกราคม 2; Available from: <https://www.kdnuggets.com/2017/07/medical-image-analysis-deep-learning-part-4.html>
- [17] Semantic Segmentation: U-Net [ออนไลน์]. 2564 พฤศจิกายน 5; Available from: <https://medium.com/red-buffer/semantic-segmentation-u-net-1e5c0f4516a5>
- [18] S. Jadon. 2020. "A survey of loss functions for semantic segmentation." IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), 1-7
- [19] Binary cross-entropy [ออนไลน์]. 2564 สิงหาคม 29; Available from: <https://androidkt.com/choose-cross-entropy-loss-function-in-keras/>
- [20] The dice loss [ออนไลน์]. 2564 มิถุนายน 27; Available from: <https://pycad.co/the-difference-between-dice-and-dice-loss/>
- [21] นาย พิชญุทธ บุญตน บทที่ 2 การขยายข้อมูล (Data Augmentation). หน้า 15

[22] Google Colab [ออนไลน์]. 2561 กันยายน 23; Available from:

<https://medium.com/@gnothaigamero/%E0%B8%AA%E0%B8%B2%E0%B8%A2-ml-%E0%B9%84%E0%B8%A1%E0%B9%88%E0%B8%84%E0%B8%A7%E0%B8%A3%E0%B8%9E%E0%B8%A5%E0%B8%B2%E0%B8%94%E0%B8%81%E0%B8%B1%E0%B8%9A-google-colab-a9f826cbf156>

[23] 3D Slicer [ออนไลน์]. 2566 กุมภาพันธ์ 22; Available from:

[https://en.wikipedia.org/wiki/3D\\_Slicer](https://en.wikipedia.org/wiki/3D_Slicer)

[24] Volume Viewer [ออนไลน์]. 2566 กุมภาพันธ์ 22; Available from:

<https://www.mathworks.com/help/images/ref/volumeviewer-app.html>



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้