

การสร้างภาพสามมิติของกระดูกต้นขาที่แตกหักจากภาพถ่ายรังสีสองมุมมองด้วยวิธีการเรียนรู้เชิง  
ลึก



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต

สาขาวิชาวิศวกรรมชีวเวช

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2565

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

3D IMAGE RECONSTRUCTION OF PROXIMAL FEMORAL FRACTURE FROM TWO-  
VIEW X-RAY IMAGES USING DEEP LEARNING



A Dissertation Submitted in Partial Fulfillment of the Requirements  
for the Degree of Doctor of Philosophy in Biomedical Engineering

FACULTY OF ENGINEERING

Chulalongkorn University

Academic Year 2022

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การสร้างภาพสามมิติของกระดูกต้นขาที่แตกหักจาก ภาพถ่ายรังสีสองมุมมองด้วยวิธีการเรียนรู้เชิงลึก
โดย	นายอนุพงษ์ บุตรทองคำ
สาขาวิชา	วิศวกรรมชีวเวช
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	รองศาสตราจารย์ ดร.ไพรัช ตั้งพรประเสริฐ
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม	รองศาสตราจารย์ ดร.ธนรัตน์ ชลิดาพงศ์ รองศาสตราจารย์ นายแพทย์วัชร  วิไลรัตน์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้รับวิทยานิพนธ์ฉบับนี้เป็นส่วน  
หนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต

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

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ  
(รองศาสตราจารย์นายแพทย์พรภวิษญ์ ศรีภิรมย์)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก  
(รองศาสตราจารย์ ดร.ไพรัช ตั้งพรประเสริฐ)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม  
(รองศาสตราจารย์ ดร.ธนรัตน์ ชลิดาพงศ์)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม  
(รองศาสตราจารย์ นายแพทย์วัชร วิไลรัตน์)

..... กรรมการ  
(ดร. นายแพทย์ศรัณย์ ตันตีทวิสุทธิ)

..... กรรมการ  
(อาจารย์ ดร.เอกพล ช่างสูวณิช)

..... กรรมการ  
(รองศาสตราจารย์ ดร.ดวงดาว วิชาดากุล)

ดุษฎีบัณฑิตกิตติมศักดิ์ : การสร้างภาพสามมิติของกระดูกต้นขาที่แตกหักจากภาพถ่ายรังสีสองมุมมองด้วยวิธีการเรียนรู้เชิงลึก. ( 3D IMAGE RECONSTRUCTION OF PROXIMAL FEMORAL FRACTURE FROM TWO-VIEW X-RAY IMAGES USING DEEP LEARNING) อ.ที่ปรึกษาหลัก : รศ. ดร.ไพรัช ตั้งพรประเสริฐ, อ.ที่ปรึกษาร่วม : รศ. ดร.ธนารัตน์ ชลิตาพงศ์, รศ. นพ.วัชร วิไลรัตน์

ภาวะกระดูกต้นขาส່วนปลายหักนั้นเป็นภาวะบาดเจ็บที่ค่อนข้างรุนแรง มีสาเหตุจากการบาดเจ็บทั้งภาวะบาดเจ็บจากอุบัติเหตุและการหักผ่านรอยโรค การวินิจฉัยและการวางแผนการผ่าตัดเป็นขบวนการที่สำคัญในการรักษา ซึ่งต้องพึ่งพาการถ่ายภาพรังสีเอกซเรย์และเอกซเรย์คอมพิวเตอร์ อย่างไรก็ตามการถ่ายภาพรังสีเอกซเรย์คอมพิวเตอร์นั้นมีความใช้จ่ยมาก ปริมาณรังสีสูง และใช้ระยะเวลาในการถ่ายภาพยาวนานกว่าการถ่ายภาพเอกซเรย์ ดังนั้นการสร้างภาพสามมิติของกระดูกต้นขาที่แตกหักจากภาพถ่ายรังสีเอกซเรย์จึงเป็นที่ต้องการและเป็นโจทย์ปัญหาที่ยังไม่ได้รับการแก้ไข อีกทั้งยังขาดแคลนข้อมูลในการวิจัย งานวิจัยนี้จึงนำเสนอโมเดลสร้างภาพสามมิติของกระดูกต้นขาส່วนปลายที่แตกหักจากภาพถ่ายรังสีสองมุมมอง เพื่อเพิ่มในภาพการมองเห็นแบบสามมิติใช้สำหรับการวินิจฉัยและการวางแผนการผ่าตัด โมเดลการสร้างภาพสามมิติได้ประยุกต์ใช้การเรียนรู้เชิงลึกซึ่งเป็นโครงข่ายประสาทเทียมคอนโวลูชันแบบพีระมิด คลาสช่วยฝึกสอนเป็นตัวแทนของบริเวณกระดูกที่แตกหักช่วยสนับสนุนการเรียนรู้รายละเอียดการรอยแตกหักของโมเดล เนื่องจากตัวอย่างภาพสำหรับฝึกสอนมีไม่เพียงพอ เทคนิคตัวอย่างฝึกสอนที่แตกหักจึงถูกพัฒนาขึ้นเพื่อขยายขนาดตัวอย่างสำหรับฝึกสอนและเพื่อเพิ่มความแม่นยำของผลลัพธ์ นอกจากนี้โมเดลสำหรับผสมคุณสมบัติตามแนวแกนจึงถูกพัฒนาขึ้นภายในโมเดลเพื่อเอาชนะปัญหาความคลาดเคลื่อนเชิงมุมระหว่างภาพถ่ายรังสี โมเดลที่นำเสนอสามารถทำงานกับชุดข้อมูลที่มีคลาดเคลื่อนเชิงมุมระหว่างภาพถ่ายรังสีได้ถึง 10 องศา โดยมีค่า mIoU ที่  $0.827 \pm 0.083SD$ . และ mASD มีค่า  $1.043 \pm 0.481SD$ . โมเดลที่นำเสนอให้รูปทรงกระดูกต้นขาและรายละเอียดของรอยแตกหักที่แม่นยำใกล้เคียงกับความเป็นจริง

สาขาวิชา	วิศวกรรมชีวเวช	ลายมือชื่อนิติ
		.....
ปี	2565	ลายมือชื่อ อ.ที่ปรึกษาหลัก
การศึกษา		.....
		ลายมือชื่อ อ.ที่ปรึกษาร่วม

# # 6271036721 : MAJOR BIOMEDICAL ENGINEERING

KEYWORD: Deep learning, Computer vision, Proximal femoral fracture, Auxiliary class, Fractural augmentation, Axial-fusion module

Danupong Buttongkum : 3D IMAGE RECONSTRUCTION OF PROXIMAL FEMORAL FRACTURE FROM TWO-VIEW X-RAY IMAGES USING DEEP LEARNING. Advisor: Assoc. Prof. PAIRAT TANGPORNPRASERT, Ph.D. Co-advisor: Assoc. Prof. THANARAT CHALIDABHONGSE, Ph.D., Assoc. Prof. VAJARA WILAIRATANA, MD.

A proximal femoral fracture is a severe injury occurring in traumatic and pathologic causes. Diagnosis and Preoperative planning are necessary procedures relying on preoperative radiographs such as X-ray and CT images. However, CT imaging has a higher radiation dose, cost, and longer acquisition time than X-ray imaging. Therefore, 3D reconstruction of the fracture from X-ray images had been requisite and remains a challenging problem, as well as a lack of dataset. This research proposes a 3D proximal femoral fracture reconstruction from biplanar radiographs to improve the 3D visualization of bone fragments for diagnosis and preoperative planning. A novel 3D Fracture Reconstruction Network is proposed which applies a deep learning-based, Fully Convolutional Network with Feature Pyramid Network architecture. The Auxiliary class is proposed, which refers to fracture representation. It encourages network learning to reconstruct the fracture. Since the samples are scarce to acquire, the Fracture augmentation is invented to enlarge the fracture training samples and improve reconstruction accuracy. Moreover, Axial-fusion module is also invented and built into the model to overcome misalignment problem. The model can operate with misaligned data up to 10 degrees of rotational error, achieving mIoU of  $0.827 \pm 0.083SD$ . and mASSD of  $1.043 \pm 0.481SD$ . The results show the precise shape and fracture detail similar to the real femoral fracture.

Field of Study: Biomedical Engineering Student's Signature

.....

Academic 2022

Advisor's Signature

## กิตติกรรมประกาศ

ขอขอบคุณ รศ.ดร.ไพรัช ตั้งพรประเสริฐ รศ.ดร.ธนารัตน์ ชลิตาพงศ์ รศ.นพ.วัชระ วิไลรัตน์ และ รศ.ดร.ชัญญาพันธ์ วิรุฬห์ศรี อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ได้ให้คำปรึกษาและการสนับสนุนในการทำงานวิจัยนี้จนสำเร็จ และได้อบรมสั่งสอนข้าพเจ้าทั้งในด้านวิชาความรู้และจริยธรรมซึ่งเป็นประโยชน์อย่างยิ่งในการทำงานและการดำเนินชีวิตในอนาคต ขอขอบคุณคณาจารย์ ภาควิชาวิศวกรรมชีวเวช และวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ทุกๆ ท่านที่ได้ให้ความรู้ทางวิชาการแก่ข้าพเจ้า ขอขอบคุณคณาจารย์แผนกรังสีและออร์โธปิดิกส์ คณะแพทยศาสตร์ จุฬาลงกรณ์มหาวิทยาลัยและโรงพยาบาลศิริราชที่ได้ให้ข้อมูลสำหรับทำวิจัยและความรู้ทางการแพทย์แก่ข้าพเจ้า ขอขอบคุณรุ่นพี่ เพื่อนๆ และน้องทุกๆ คนที่ได้มีส่วนร่วม และได้ให้ความช่วยเหลือในงานวิจัยนี้ ขอขอบคุณคณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ที่ได้ให้การสนับสนุนการศึกษา และเป็นสถานที่ทำวิจัยในครั้งนี้จนสำเร็จไปได้ด้วยดี สุดท้ายขอขอบทูน Second Century (C2F) ที่ให้ทุนการศึกษาและทุนวิจัยในครั้งนี้ ข้าพเจ้าหวังว่างานวิจัยนี้ จะเป็นประโยชน์ต่อทางมหาวิทยาลัย และเป็นส่วนสำคัญในการพัฒนาประเทศไทย

دنۇفۇنۇ نۇتۇرۇتۇنۇ

จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY

## สารบัญ

	หน้า
บทคัดย่อภาษาไทย .....	ค
บทคัดย่อภาษาอังกฤษ.....	ง
กิตติกรรมประกาศ.....	จ
สารบัญ.....	ฉ
สารบัญตาราง.....	ณ
สารบัญรูป .....	ฏ
บทที่ 1 บทนำ .....	1
1.1 ความสำคัญและที่มาของปัญหา (Rationale) .....	1
1.2 วัตถุประสงค์ของการวิจัย (Objectives) .....	2
1.3 ขอบเขตงานวิจัย (Scope of Research) .....	2
1.4 ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย (Expected or Anticipated Benefit Gain).....	3
บทที่ 2 ทบทวนวรรณกรรมงานวิจัยที่เกี่ยวข้อง (Review Literature).....	4
2.1 การบาดเจ็บของรยางค์ส่วนล่าง .....	4
2.2 ภาพถ่ายรังสีทางการแพทย์ (Medical imaging) .....	6
2.2.1 ทฤษฎีภาพถ่ายรังสีเอกซเรย์ (X-ray radiography) .....	6
2.2.2 ภาพถ่ายรังสีเอกซเรย์คอมพิวเตอร์ (Computed Tomography, CT).....	8
2.3.3 ความสำคัญของภาพสามมิติกับการวางแผนการรักษา.....	9
2.2.4 ความเสี่ยงจากการถ่ายภาพรังสีทางการแพทย์ (Risk of X-ray imaging).....	11
2.3 สร้างโมเดลสามมิติด้วยแบบจำลองรูปร่างเชิงสถิติ (3D Reconstruction using Statistic shape modeling).....	13

2.4 การสร้างโมเดลสามมิติจากการเรียนรู้ของเครื่อง (Machine learning) .....	15
2.4.1 ความรู้พื้นฐานเกี่ยวกับการเรียนรู้ของเครื่อง (Basic Knowlegde of Machine Learning) .....	15
2.4.2 งานวิจัยที่เกี่ยวข้อง (Related work).....	24
บทที่ 3 การออกแบบโมเดลสร้างรูปทรงสามมิติเบื้องต้น (Preliminary Design 3D Reconstruction Model) .....	31
3.1 การเก็บข้อมูล (Data collection).....	31
3.2 การจัดเตรียมข้อมูลสำหรับการสอน (Data Preprocessing).....	32
3.2.1 การแบ่งส่วนกระดูกต้นขา (Fractural Femur Segmentation) .....	32
3.2.2 การสร้างภาพ Scout film .....	35
3.3 โมเดลสร้างรูปทรงสามมิติ (3D Reconstruction Network) .....	38
3.4 การฝึกสอนโมเดล (Model training) .....	43
3.5 การวัดความแม่นยำเชิงปริมาณ (Quantitative evaluation).....	45
3.6 ผลลัพธ์ (Result).....	46
บทที่ 4 วิธีการที่นำเสนอ (Proposed Method).....	49
4.1 เทคนิคการเรียนรู้ด้วยคลาสช่วยฝึกสอน (Auxiliary class) .....	49
4.2 เทคนิคการเพิ่มขนาดกลุ่มตัวอย่าง (Data Augmentation) .....	51
4.3 ผลลัพธ์จากการใช้ Proposed method .....	54
บทที่ 5 การปรับปรุงประสิทธิภาพของโมเดลสร้างรูปทรงสามมิติ .....	59
5.1 การสร้างภาพถ่ายรังสีแบบดิจิทัลแบบมวมคลาดเคลื่อน (Misaligned Digitally Reconstructed Radiograph).....	59
5.2 การปรับปรุงโมเดลการสร้างภาพสามมิติที่สามารถประมวลผลภาพถ่ายรังสีสองมุมมองที่มีความคลาดเคลื่อนเชิงมุม .....	60



5.3 ผลลัพธ์จากการปรับปรุงโมเดลสร้างรูปทรงสามมิติด้วย Axial-fusion module .....	64
บทที่ 6 การประเมินเชิงคุณภาพของโมเดลสร้างรูปทรงสามมิติ .....	68
6.1 วิธีการประเมินเชิงคุณภาพ .....	68
6.2 ผลลัพธ์จากการประเมินเชิงคุณภาพ .....	70
บทที่ 7 สรุปผลการวิจัยและข้อเสนอแนะงานวิจัยในอนาคต .....	73
7.1 สรุปผลการวิจัย .....	73
7.2 ข้อเสนอแนะงานวิจัยในอนาคต .....	74
ภาคผนวก A ผลลัพธ์การสร้างรูปทรงสามมิติ วัดผลด้วย IoU metric .....	76
ภาคผนวก B ผลลัพธ์การสร้างรูปทรงสามมิติ วัดผลด้วย ASSD metric .....	78
ภาคผนวก C ผลลัพธ์จากการสร้างรูปทรงสามมิติแบบ 3D Volume และ Surface distance error .....	80
ภาคผนวก D เอกสารจริยธรรมงานวิจัยในมนุษย์ .....	83
บรรณานุกรม .....	86
บรรณานุกรม .....	91
ประวัติผู้เขียน .....	93

## สารบัญตาราง

	หน้า
ตารางที่ 1 ผลลัพธ์จากการผ่าตัดเปรียบเทียบระหว่างศัลยแพทย์กลุ่ม A และ B [18].....	11
ตารางที่ 2 ปริมาณรังสีที่ได้รับจากการวินิจฉัยโรคตามอวัยวะบริเวณต่างๆ กัน [9] .....	12
ตารางที่ 3 DICOM Attribute ของค่า Rescale Intercept และ Slope .....	36
ตารางที่ 4 ความแม่นยำของโมเดลสร้างรูปทรงสามมิติของกระดูกชนิดต่างๆ .....	46
ตารางที่ 5 เปรียบเทียบค่า mIoU และ mASSD ด้วยโมเดลสร้างภาพสามมิติต่างๆ .....	54
ตารางที่ 6 การทดสอบแบบจับคู่ (Paired T-test) เปรียบเทียบความแตกต่างระหว่างผลลัพธ์ที่ได้จากเทคนิคพิเศษต่าง .....	55
ตารางที่ 7 เปรียบเทียบจำนวน learnable parameters, vRAM ที่ใช้, เวลาในการฝึกสอนและ อนุมานผลลัพธ์.....	55
ตารางที่ 8 เปรียบเทียบความแตกต่างระหว่างผลลัพธ์ที่ได้จากโมเดล FracReconNet-Aligned, FracReconNet-Misaligned และ FracReconNetAxial-Misaligned ด้วย IoU .....	66
ตารางที่ 9 เปรียบเทียบความแตกต่างระหว่างผลลัพธ์ที่ได้จากโมเดล FracReconNet-Aligned, FracReconNet-Misaligned และ FracReconNetAxial-Misaligned ด้วย ASSD .....	66
ตารางที่ 10 เปรียบเทียบจำนวนพารามิเตอร์ (Learnable parameters), การใช้ vRAM, ระยะเวลา ฝึกสอน (Training time) และระยะเวลาอนุมานผลลัพธ์ (Inference time) ระหว่างโมเดล FracReconNet ที่ใช้ Fusion module แบบเดียวกับแบบ Axial-fusion .....	67
ตารางที่ 11 แสดงผลลัพธ์การสร้างรูปทรงสามมิติด้วยโมเดล FracReconNet ซึ่งถูกฝึกสอนด้วย ภาพถ่ายรังสีสองมุมมองที่ตั้งฉากกัน (aligned data) .....	76
ตารางที่ 12 แสดงผลลัพธ์การสร้างรูปทรงสามมิติด้วยโมเดล FracReconNet ซึ่งถูกฝึกสอนด้วย ภาพถ่ายรังสีสองมุมมองที่มีความคลาดเคลื่อนเชิงมุม (misaligned data) .....	76
ตารางที่ 13 แสดงผลลัพธ์การสร้างรูปทรงสามมิติด้วยโมเดล FracReconNet-Axial ซึ่งถูกฝึกสอน ด้วยภาพถ่ายรังสีสองมุมมองที่มีความคลาดเคลื่อนเชิงมุม (misaligned data) .....	77

ตารางที่ 14 แสดงผลลัพธ์การสร้างรูปทรงสามมิติด้วยโมเดล FracReconNet ซึ่งถูกฝึกสอนด้วย  
 ภาพถ่ายรังสีสองมุมมองที่ตั้งฉากกัน (aligned data) ..... 78

ตารางที่ 15 แสดงผลลัพธ์การสร้างรูปทรงสามมิติด้วยโมเดล FracReconNet ซึ่งถูกฝึกสอนด้วย  
 ภาพถ่ายรังสีสองมุมมองที่มีความคลาดเคลื่อนเชิงมุม (misaligned data) ..... 78

ตารางที่ 16 แสดงผลลัพธ์การสร้างรูปทรงสามมิติด้วยโมเดล FracReconNet-Axial ซึ่งถูกฝึกสอน  
 ด้วยภาพถ่ายรังสีสองมุมมองที่มีความคลาดเคลื่อนเชิงมุม (misaligned data) ..... 79



## สารบัญรูป

	หน้า
รูปที่ 1 การจำแนกการแตกหักของคอกระดูกต้นขาหักของ Garden [14].....	5
รูปที่ 2 การจำแนกการแตกหักของกระดูกบริเวณ Intertrochanteric ของ Evan [14].....	5
รูปที่ 3 การจำแนกการแตกหักของกระดูกต้นขาบริเวณ Subtrochanteric ของ Russell-Taylor [14].....	6
รูปที่ 4 ปรากฏการณ์ระหว่างรังสีเอกซ์เรย์กับสสารที่รังสีเคลื่อนผ่าน .....	7
รูปที่ 5 อธิบายการดูดกลืนรังสีตามกฎของ Beer-Lambert .....	7
รูปที่ 6 ค่า mass attenuation ของ (a) กระดูก ( $\rho = 1.920 \text{ g/cm}^3$ ) (b) กล้ามเนื้อ ( $\rho = 1.044 \text{ g/cm}^3$ ) และ (c) น้ำ ( $\rho = 1.000 \text{ g/cm}^3$ ) เมื่อรังสีเอกซ์เรย์ส่องผ่านที่ค่าระดับพลังงานต่างๆ [17]... 8	8
รูปที่ 7 แผนภาพเครื่องถ่ายภาพรังสีเอกซ์เรย์คอมพิวเตอร์ [13].....	9
รูปที่ 8 ตัวอย่างภาพถ่ายรังสีเอกซ์เรย์คอมพิวเตอร์ [13].....	9
รูปที่ 9 ข้อมูลภาพถ่ายรังสีที่สัลยแพทย์ใช้สำหรับการวางแผนการผ่าตัด [18].....	10
รูปที่ 10 Learning curve ความเร็วในการผ่าตัดระหว่างสัลยแพทย์กลุ่ม A และ B [18] .....	11
รูปที่ 11 การสร้างรูปร่างเฉลี่ยของกระดูกต้นขา (Shape model construction) [21] .....	14
รูปที่ 12 ขบวนการซ้อนทับภาพระหว่างโมเดลสามมิติกับภาพเอกซ์เรย์ [21].....	15
รูปที่ 13 เพอร์เซปตรอนหนึ่งหน่วย (Perceptron unit) [22].....	17
รูปที่ 14 ฟังก์ชันกระตุ้น (Activation function) [22] .....	17
รูปที่ 15 โครงข่ายการเรียนรู้เชิงลึก (Deep learning) [22].....	18
รูปที่ 16 โครงข่ายประสาทแบบคอนโวลูชัน (Convolutional neural network ; ConvNet) [22] 21	21
รูปที่ 17 การเปรียบเทียบโครงสร้าง (ก) fully-connected layer กับ (ข) Convolutional layer [22] .....	21
รูปที่ 18 ระดับชั้นรวม(Pooling layer) [22].....	22

รูปที่ 19	เปรียบเทียบ Deconvolution และ Unpooling กับ Convolution และ Pooling [23]	22
รูปที่ 20	โครงข่าย CNN ที่มีรูปแบบการเชื่อมต่อแบบพีระมิด [24]	22
รูปที่ 21	โครงข่าย Single-image Tomography [28]	24
รูปที่ 22	ภาพรวมของโมเดล X2CT-GAN [29]	25
รูปที่ 23	Generator ของ X2CT-GAN [29]	26
รูปที่ 24	TL-Network architecture [30]	27
รูปที่ 25	Statistical spline model [31]	28
รูปที่ 26	ภาพรวมโมเดลการสร้างรูปทรงสามมิติแบบ CNN SSM fitting [31]	29
รูปที่ 27	โครงข่าย CNN สำหรับ spline landmark detection [31]	30
รูปที่ 28	ผลลัพธ์จากการสร้างรูปทรงสามมิติของ spline [31]	30
รูปที่ 29	การแบ่งกลุ่มของกลุ่มตัวอย่างสำหรับการเรียนรู้ของเครื่อง	32
รูปที่ 30	ขั้นตอนการแบ่งส่วนปริมาตรของกระดูกต้นขาที่แตกหัก	32
รูปที่ 31	การแบ่งส่วนกระดูกด้วยวิธีระดับความเข้ม (Thresholding segmentation)	33
รูปที่ 32	การแยกชิ้นส่วนกระดูกแต่ละชิ้น	34
รูปที่ 33	การเติมเต็มช่องว่างในกระดูก	34
รูปที่ 34	กระดูกต้นขาที่ถูกแบ่งส่วนเรียบร้อยแล้ว (Complete Femur Mask)	34
รูปที่ 35	แผนภาพการสังเคราะห์ Scout film [20]	35
รูปที่ 36	ผลลัพธ์จากการคำนวณ LAC	36
รูปที่ 37	ผลลัพธ์การสังเคราะห์ Scout film หรือ ภาพรังสีแบบดิจิทัล	37
รูปที่ 38	ตัวอย่างสำหรับฝึกสอนโมเดลสร้างรูปทรงสามมิติของกระดูกต้นขา	38
รูปที่ 39	แผนภาพโมเดลสร้างรูปทรงสามมิติของกระดูกต้นขา (3DReconNet)	39
รูปที่ 40	การรวมคุณสมบัติ (Fusion) $X_v^{(i)}$ ของแต่ละมุมมองเข้าด้วยกัน	43
รูปที่ 41	กราฟค่าความสูญเสีย Focal loss ต่อค่าความน่าจะเป็น [33]	44

รูปที่ 42 ผลลัพธ์การสร้างรูปทรงสามมิติของกระดูกต้นขาปกติ .....	47
รูปที่ 43 ผลลัพธ์การสร้างรูปทรงสามมิติของกระดูกต้นขาหักแบบไม่เคลื่อนที่ .....	47
รูปที่ 44 ผลลัพธ์การสร้างรูปทรงสามมิติของกระดูกต้นขาหักแบบเคลื่อนที่ .....	48
รูปที่ 45 ขั้นตอนการสังเคราะห์คลาสช่วยฝึกสอน (Auxiliary class) จากโมเดลกระดูกที่แตกหัก	50
รูปที่ 46 กราฟ Power spectrum density .....	52
รูปที่ 47 กราฟ Uniformly random distributed frequency .....	52
รูปที่ 48 Fractal surface สำหรับ Data Augmentation .....	53
รูปที่ 49 การเพิ่มขนาดตัวอย่างกระดูกที่แตกหัก (Fracture augmentation) .....	53
รูปที่ 50 เปรียบเทียบผลลัพธ์การสร้างรูปทรงสามมิติของกระดูกต้นขาปกติ .....	56
รูปที่ 51 เปรียบเทียบผลลัพธ์การสร้างรูปทรงสามมิติของกระดูกต้นขาหักแบบไม่เคลื่อนที่ .....	57
รูปที่ 52 เปรียบเทียบผลลัพธ์การสร้างรูปทรงสามมิติของกระดูกต้นขาหักแบบเคลื่อนที่ .....	57
รูปที่ 53 การสังเคราะห์ Scout film ที่มีความคลาดเคลื่อนเชิงมุมที่ $90 \pm 10$ องศา .....	59
รูปที่ 54 ตัวอย่างภาพคุณสมบัติ $D_1^{(i)}$ กับ $P(D_2^{(i)})$ ที่ซ้อนทับกันของภาพรังสีสองมุมมองที่มีความ คลาดเคลื่อนเชิงมุม 10 องศา กรอบสีน้ำเงินแสดงถึงตำแหน่ง key-point บน anatomical structure เดียวกันของแต่ละมุมมอง .....	61
รูปที่ 55 การผสมคุณสมบัติของ Native fusion module ในหัวข้อที่ 3.3 หมายถึงกรอบสีน้ำเงิน ทั้งสองแสดงตำแหน่งการผสมกันในอุดมคติ ส่วนกรอบสีแดงแสดงตำแหน่งการผสมกันที่เกิดขึ้น ขณะใช้ Conv3D .....	62
รูปที่ 56 การผสมคุณสมบัติด้วย 2D Convolution แบบ Axial-fusion .....	63
รูปที่ 57 โครงสร้าง Axial-fusion module ใช้สำหรับเพิ่มประสิทธิภาพการรวมคุณสมบัติที่มี มุมมองคลาดเคลื่อนกัน .....	64
รูปที่ 58 กราฟแสดงค่า Average IoU ของผลลัพธ์การสร้างรูปทรงสามมิติ ที่ความคลาดเคลื่อน เชิงมุมต่าง (2.5 ถึง 10 องศา) .....	65

รูปที่ 59 กราฟแสดงค่า Average ASSD ของผลลัพธ์การสร้างรูปทรงสามมิติ ที่ความคลาดเคลื่อนเชิงมุมต่าง (2.5 ถึง 10 องศา).....	66
รูปที่ 60 ตัวอย่างทดสอบ ซึ่งประกอบไปด้วย ภาพถ่ายรังสี (Input) และรูปทรงสามมิติที่ถูกสร้างขึ้น (Output) เพื่อใช้สำหรับประเมินเชิงคุณภาพ .....	68
รูปที่ 61 การจำแนกการแตกหักของกระดูกต้นขาแบบต่างๆ .....	69
รูปที่ 62 ความถูกต้องของการวินิจฉัยเพื่อจำแนกประเภทการแตกหักของกระดูกต้นขา .....	70
รูปที่ 63 ผลลัพธ์จากการประเมินเชิงคุณภาพของโมเดลสร้างรูปทรงสามมิติ.....	71
รูปที่ 64 แสดงผลลัพธ์แบบ 3D volume จากการสร้างรูปทรงสามมิติของโมเดลต่างๆ จากภาพถ่ายรังสีสองมุมมองที่ตั้งฉากกัน .....	80
รูปที่ 65 แสดงผลลัพธ์แบบ 3D volume จากการสร้างรูปทรงสามมิติของโมเดลต่างๆ จากภาพถ่ายรังสีสองมุมมองที่มีความคลาดเคลื่อนเชิงมุม 5 องศา.....	80
รูปที่ 66 แสดงผลลัพธ์แบบ 3D volume จากการสร้างรูปทรงสามมิติของโมเดลต่างๆ จากภาพถ่ายรังสีสองมุมมองที่มีความคลาดเคลื่อนเชิงมุม 10 องศา.....	81
รูปที่ 67 แสดงผลลัพธ์แบบ Surface distance error จากการสร้างรูปทรงสามมิติของโมเดลต่างๆ จากภาพถ่ายรังสีสองมุมมองที่ตั้งฉากกัน .....	81
รูปที่ 68 แสดงผลลัพธ์แบบ Surface distance error จากการสร้างรูปทรงสามมิติของโมเดลต่างๆ จากภาพถ่ายรังสีสองมุมมองที่มีความคลาดเคลื่อนเชิงมุม 5 องศา.....	82
รูปที่ 69 แสดงผลลัพธ์แบบ Surface distance error จากการสร้างรูปทรงสามมิติของโมเดลต่างๆ จากภาพถ่ายรังสีสองมุมมองที่มีความคลาดเคลื่อนเชิงมุม 10 องศา.....	82

## บทที่ 1 บทนำ

### 1.1 ความสำคัญและที่มาของปัญหา (Rationale)

การบาดเจ็บกระดูกหักนั้นเป็นภาวะการบาดเจ็บที่ค่อนข้างรุนแรง สามารถเกิดขึ้นได้กับคนทุกเพศทุกวัย โดยเฉพาะกับผู้สูงอายุ และเพศหญิงมีโอกาสกระดูกหักมากกว่าเพศชาย สาเหตุการหักของกระดูกนั้นเกิดได้ทั้งจากภาวะบาดเจ็บ (Traumatic fracture) เช่น การประสบอุบัติเหตุต่างๆ และการหักผ่านรอยโรค (Pathologic fracture) จากความผิดปกติของกระดูกของผู้ป่วยรายนั้นๆ เช่น มีเนื้องอกหรือมีการติดเชื้ออยู่ก่อนแล้วหรือมีภาวะกระดูกพรุน ทำให้กระดูกเปราะแตกหักง่าย [1] บริเวณที่เกิดการแตกหักที่พบได้บ่อยสุด คือ กระดูกต้นขาส่วนต้น (Proximal femur) [2, 3] เนื่องจากกระดูกบริเวณนั้นได้รับความเค้นสูง (Mechanical stress concentration) จากลักษณะรูปทรงของกระดูกเองและภาระน้ำหนักที่กระทำ[4] ประกอบกับการเกิดโรคทางพยาธิวิทยาในบริเวณนั้น

สำหรับการรักษากระดูกต้นขาส่วนต้นหักจะเริ่มจากการวินิจฉัยลักษณะการแตกหักของกระดูกโดยใช้ภาพถ่ายรังสีเอกซเรย์ (X-ray image) ทำการวางแผนการผ่าตัด (Preoperative planning) เพื่อเลือกวิธีการรักษาและอุปกรณ์ที่ต้องใช้ และทำการผ่าตัดโดยเร็วที่สุดเพื่อให้หลังจากเกิดการแตกหักของกระดูก[5] การผ่าตัดที่นิยมใช้รักษากระดูกต้นขาหัก คือ การผ่าตัดใส่โลหะยึดตรึงกระดูก ซึ่งเป็นการผ่าตัดจัดกระดูกแบบแผลปิด (Closed reduction) เนื่องจากเป็นวิธีการรักษาที่รุกรานน้อยที่สุด (Minimally invasive) ใช้เวลาผ่าตัดน้อย ให้ผลลัพธ์จากการรักษาที่ดี และผู้ป่วยสามารถลุกขึ้นเดินได้อย่างรวดเร็ว[6, 7] วิธีการผ่าตัดแบบนี้เหมาะกับศัลยแพทย์ผู้มีประสบการณ์สูงและมีความสามารถในการสร้างนโนภาพรูปทรงสามมิติของกระดูกที่หักแต่ละชิ้นส่วนจากภาพถ่ายรังสีเอกซเรย์ทั้งก่อนและในระหว่างผ่าตัด เพื่อที่จะทำการจัดกระดูกกลับเข้าที่ตำแหน่งเดิมได้อย่างถูกต้องก่อนจะยึดตรึงกระดูก[8] หากไม่สามารถจัดตำแหน่งของกระดูกให้เข้าที่เดิมได้อย่างแม่นยำหรือเลือกใช้อุปกรณ์ยึดตรึงกระดูกที่ไม่เหมาะสม อาจจะทำให้เกิดภาวะแทรกซ้อนที่ตามมา เช่น กระดูกที่หักแต่ละชิ้นไม่เชื่อมติดกันหรือเชื่อมติดกันช้า (Non- or Delay-union) กระดูกเชื่อมกันแบบผิดรูป (Malunion) อาจเกิดการแตกหักซ้ำ ซึ่งทำการผ่าตัดใหม่ซึ่งจะยากลำบากกว่าการผ่าตัดรอบแรกมาก [4, 5] และเนื่องจากต้องมีการถ่ายภาพรังสีซ้ำๆ หลายครั้งในระหว่างผ่าตัด เพื่อเล็งตำแหน่งของกระดูก ส่งผลให้ทั้งศัลยแพทย์และผู้ป่วยได้รับปริมาณรังสีที่สูงตามด้วย อาจจะทำให้เกิดผลกระทบบจากรังสีในระยะยาวอีกด้วย [9, 10]



การถ่ายภาพรังสีเอกซเรย์คอมพิวเตอร์ (Computerized Tomography; CT) สำหรับใช้ในการวางแผนการรักษาจึงเอามามีส่วนสำคัญในการ เนื่องจากความสามารถในการถ่ายภาพกายวิภาคตรงบริเวณที่สนใจของผู้ป่วยเฉพาะราย (Individual target anatomy) ในรูปแบบสามมิติ และสามารถนำไปประมวลผลต่อเพื่อให้ได้รูปทรงกระดูกแบบสามมิติ (Reconstruction image) ทำให้แพทย์สามารถมองเห็นรูปทรงกระดูกที่แตกหักก่อน จึงสามารถวางแผนการและเลือกวิธีการผ่าตัดได้อย่างรวดเร็ว แม่นยำ ลดจำนวนครั้งในการถ่ายภาพรังสีในระหว่างตัด และลดปริมาณเลือดที่สูงเสียไปในระหว่างผ่าตัด เมื่อเทียบกับการวางแผนก่อนผ่าตัดโดยใช้ภาพถ่ายรังสีเอกซเรย์แบบปกติ [11, 12] แต่ก็มีข้อเสียที่การถ่ายภาพ CT นั้นใช้ปริมาณรังสีที่สูงมากๆ ใช้ระยะเวลาในการประมวลผลนาน และมีค่าใช้จ่ายที่สูงกว่า เมื่อเทียบกับภาพถ่ายรังสีเอกซเรย์แบบปกติ [9, 13] การพัฒนาเทคนิคการถ่ายภาพทางการแพทย์แบบสามมิติที่ใช้ปริมาณรังสีต่ำจึงเป็นสิ่งสำคัญที่ต่อพัฒนาต่อไป

## 1.2 วัตถุประสงค์ของการวิจัย (Objectives)

งานวิจัยนี้จึงสนใจที่จะพัฒนาอัลกอริทึมสำหรับการสร้างรูปทรงกระดูกต้นขาสามมิติ ทั้งกระดูกปกติ (Intact) และกระดูกหักจากภาวะบาดเจ็บ (Traumatic fracture) ที่มีความแม่นยำ ด้วยภาพถ่ายรังสีสองมุมมอง เพื่อเพิ่มมโนภาพการมองเห็นรูปทรงและลักษณะการแตกหักของกระดูกแบบสามมิติให้กับบุคลากรทางการแพทย์ สำหรับวินิจฉัยและการวางแผนการผ่าตัด (Preoperative planning) ของผู้ป่วยกระดูกต้นขาแตกหัก

## 1.3 ขอบเขตงานวิจัย (Scope of Research)

- 1) เนื่องจากระยะเวลาทำวิจัยและงบประมาณที่มีจำกัด งานวิจัยนี้เป็นการวิจัยศึกษาจากข้อมูลย้อนหลัง (Retrospective study) เท่านั้น ใช้ข้อมูลภาพถ่ายรังสีย้อนหลังหรือข้อมูลที่มีอยู่แล้วในฐานะข้อมูลของโรงพยาบาลจุฬาลงกรณ์ เพื่อนำไปวิเคราะห์และพัฒนาโปรแกรมคอมพิวเตอร์ ไม่มีการรับสมัครผู้ป่วยรายใหม่มาเพื่อถ่ายภาพรังสี
- 2) ใช้ Scout film ที่ได้มาจากภาพถ่าย CT สำหรับฝึกสอนแทนการใช้ภาพถ่ายรังสีที่ถ่ายจากผู้ป่วยจริงๆ เนื่องจากเป็นการศึกษาแบบย้อนหลังจึงไม่สามารถเก็บข้อมูลภาพถ่ายรังสีที่คู่กับภาพถ่ายรังสีเอกซเรย์คอมพิวเตอร์ของผู้ป่วยเฉพาะรายได้
- 3) ความแม่นยำ นิยามโดยค่า Average IoU ที่มากกว่า 0.7 และการประเมินความถูกต้องโดยแพทย์ผู้ชำนาญ

- 4) การทดสอบอัลกอริทึมจะใช้ข้อมูล 10% ของข้อมูลทั้งหมด ซึ่งไม่ได้ถูกใช้ในการฝึกสอนหรือทดสอบมาก่อน ข้อมูลป้อนเข้าจะเป็นภาพ Scout film หรือ ภาพถ่ายรังสี (ในกรณีที่สามารถค้นหาได้จากฐานข้อมูลได้)

#### 1.4 ประโยชน์ที่คาดว่าจะได้รับการวิจัย (Expected or Anticipated Benefit Gain)

ประโยชน์ทางด้านการแพทย์ คือ ระบบนี้ช่วยในศัลยแพทย์สามารถเห็นภาพสามมิติของกระดูกที่แตกหักได้โดยใช้เพียงภาพถ่ายรังสีเอกซเรย์ เป็นประโยชน์ต่อการวินิจฉัยและการวางแผนการผ่าตัด ทดแทนการถ่ายภาพเอกซเรย์คอมพิวเตอร์ที่มีปริมาณรังสีที่สูงกว่า ประโยชน์ทางสังคม คือ ช่วยลดค่าใช้จ่ายในการรักษาให้กับผู้ป่วยที่มีความจำเป็นที่จะต้องเข้ารับการรักษาโดยใช้ภาพถ่ายเอกซเรย์คอมพิวเตอร์ สามารถเข้าถึงเทคนิคสร้างภาพสามมิติด้วยวิธีการใหม่ที่ผู้วิจัยกำลังนำเสนอได้ ประโยชน์ทางเศรษฐกิจ คือ หากงานวิจัยนี้ได้ผลสำเร็จตามที่นำเสนอไป และสามารถกระจายการใช้งานไปสู่โรงพยาบาลทั่วประเทศได้ ช่วยประหยัดค่าใช้จ่ายในการรักษาได้ในวงกว้าง

## บทที่ 2

### บททวนวรรณกรรมงานวิจัยที่เกี่ยวข้อง (Review Literature)

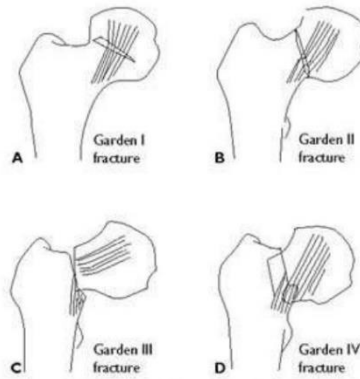
#### 2.1 การบาดเจ็บของรยางค์ส่วนล่าง

การบาดเจ็บของรยางค์ส่วนล่าง คือ การบาดเจ็บตั้งแต่กระดูกเชิงกรานไปจนถึงบริเวณปลายนิ้วเท้า เป็นการบาดเจ็บที่มีความสำคัญและพบได้บ่อย เกิดได้จากหลากหลายสาเหตุ เช่น อุบัติเหตุทางจราจร การหกล้ม และการเสื่อมสภาพตามอายุที่มากขึ้น เป็นต้น ในงานวิจัยนี้มีขอบเขตการวิจัยเฉพาะการแตกหักของกระดูกต้นขาเท่านั้น ซึ่งสามารถแบ่งได้ตามบริเวณที่แตกหัก โดยมีรายละเอียดดังต่อไปนี้

##### 1) คอกระดูกต้นขาหัก (Femur neck fracture)

เป็นภาวะที่พบได้บ่อยในผู้ป่วยสูงอายุ มีโอกาสเกิดในสตรี 80% และความสัมพันธ์กับโรคกระดูกพรุนโดยตรง อีกทั้งยังเกิดได้ในผู้ป่วยที่มีอายุน้อยจากกรณีเกิดอุบัติเหตุที่รุนแรง เช่น ตกจากที่สูง อุบัติเหตุทางจราจร เป็นต้น การจำแนกประเภทของการแตกหักบริเวณคอกระดูกต้นขาตามนิยามของ Garden classification อธิบายได้ดังนี้ (รูปที่ 1)

- |            |                                                                                                                                                                                              |
|------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Garden I   | หักแบบไม่สมบูรณ์ (incomplete fracture) และหัวกระดูกต้นขาอัดฝัง เข้ากับกระดูกคอสะโพก (valgus impact) จะเห็นรอยหักของ cortex ส่วนบนเท่านั้น และอาจเห็นเงาของกระดูกที่ซ้อนกันเป็นเส้นที่บดสีขาว |
| Garden II  | หักแบบสมบูรณ์ (complete fracture) แต่ไม่มีการเคลื่อนของกระดูก ทั้งในท่าตรง และ lateral cross table จะเห็นรอยหักผ่านทั้งสอง cortex                                                            |
| Garden III | หักแบบสมบูรณ์และมีการเคลื่อนที่เกยกันของกระดูกบางส่วน (partial displaced) จะสังเกตเห็นเส้น trabecular ของหัวกระดูกทำมุมกับเส้น trabecular ที่คอกระดูก                                        |
| Garden IV  | หักแบบสมบูรณ์และมีการเคลื่อนหลุดระหว่างกระดูกสองหัวและคอ (total displaced) จะสังเกตเห็นเส้น trabecular ของหัวกระดูก ขนานกับเส้น trabecular ที่คอกระดูก[14]                                   |



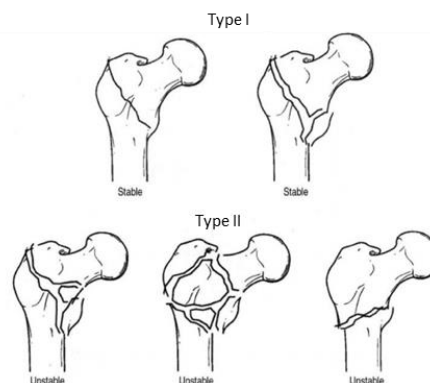
รูปที่ 1 การจำแนกการแตกหักของคอกระดูกต้นขาหักของ Garden [14]

## 2) Intertrochanteric Fracture

คือภาวะกระดูกหักบริเวณระหว่าง greater trochanter กับ lesser trochanter เป็นกรณีที่พบได้บ่อยที่สุดของภาวะกระดูกต้นขาส่วนบนหัก ซึ่งมีความสัมพันธ์โดยตรงกับภาวะกระดูกพรุนสาเหตุส่วนมากนั้นมาจากอุบัติเหตุในกลุ่มผู้ป่วยที่มีภาวะกระดูกพรุนร่วมด้วย การหักของกระดูกต้นขาบริเวณนี้สามารถจำแนกตาม Evans classification ได้เป็น 2 ประเภท ตามเสถียรภาพของกระดูก แสดงดังรูปที่ 2 คือ

Evan type I คือการหักแบบมั่นคง (stable fracture) กระดูกบริเวณ posteromedial ของบริเวณ trochanter ยังไม่หักหรืออาจหักเป็นชิ้นเล็ก แต่กระดูกยังคงจัดเรียงตัวกันได้อย่างมั่นคง มีเสถียรภาพในการคงรูปขณะรับน้ำหนักเบาๆ

Evan type II คือการหักแบบไม่มั่นคง (unstable fracture) มีการแตกของกระดูกบริเวณ posteromedial เป็นชิ้นใหญ่ หรือแตกเป็นหลายชิ้น มีการแตกจาก lesser trochanter ลงไปยังบริเวณกระดูกส่วน subtrochanteric ลักษณะของการแตกหักแบบนี้ กระดูกจะไม่สามารถคงรูปได้เมื่อบริเวณนี้รับน้ำหนัก [14, 15]

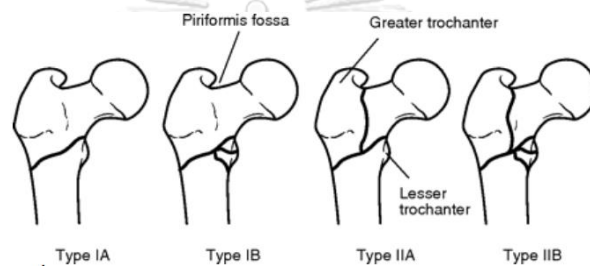


รูปที่ 2 การจำแนกการแตกหักของกระดูกบริเวณ Intertrochanteric ของ Evan [14]

### 3) Subtrochanteric Fracture

คือการหักของกระดูกต้นขาตั้งแต่บริเวณ lesser trochanter ลงมาไม่เกิน 5 cm. ซึ่งเป็นบริเวณที่มีเลือดมาเลี้ยง cortical bone น้อย ประกอบกับการเป็นบริเวณที่รับความเค้นสูง จึงทำให้การเชื่อมติดกันของกระดูกบริเวณนี้เกิดขึ้นช้า และพบภาวะแทรกซ้อนจากการผ่าตัดได้สูงกว่ากระดูกต้นขาบริเวณอื่น สาเหตุการหักมาจากการประคบอุบัติเหตุ หรือผู้ป่วยสูงอายุที่มีภาวะกระดูกพรุน ตามคำนิยามของ Russell-Taylor ได้แบ่งชนิดของการหักได้เป็น 2 ประเภท ตามรูปที่ 3 คือ

- Type I คือการหักที่ไม่มีการแตกของ piriformis fossa ร่วมด้วย  
 Type II คือการหักที่มีการแตกร้าวเข้าไปที่บริเวณ piriformis fossa [14]

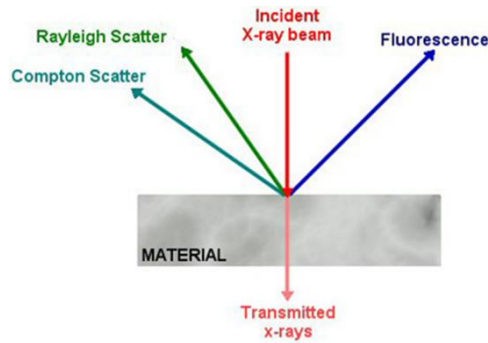


รูปที่ 3 การจำแนกการแตกหักของกระดูกต้นขาบริเวณ Subtrochanteric ของ Russell-Taylor [14]

## 2.2 ภาพถ่ายรังสีทางการแพทย์ (Medical imaging)

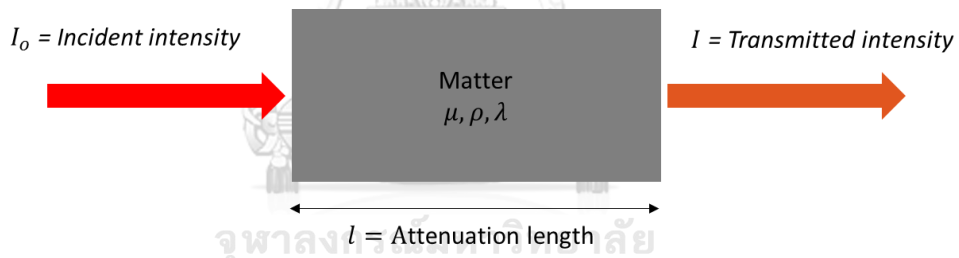
### 2.2.1 ทฤษฎีภาพถ่ายรังสีเอกซเรย์ (X-ray radiography)

รังสีเอกซเรย์ (X-ray) เป็นการแผ่รังสีแม่เหล็กไฟฟ้า (Electromagnetic radiation) พลังงานสูงรูปแบบหนึ่ง ที่มีความยาวคลื่นอยู่ในช่วง 0.03 - 0.3 nm , ความถี่ในช่วง  $30 \times 10^{15} - 30 \times 10^{18}$  Hz และพลังงานในช่วง 100 eV – 200 KeV มีความสามารถในการทะลุผ่านสสารต่างๆ และตกกระทบบนฟิล์มหรือตัวรับภาพทำให้เกิดเป็นภาพถ่ายรังสีเอกซเรย์ขึ้นที่สามารถมองเห็นรูปร่างของสสารที่แตกต่างกัน ในกรณีนำไปใช้กับร่างกายมนุษย์จะสามารถเห็นรูปร่างและแยกแยะชนิดของอวัยวะต่างๆ ภายในร่างกายที่รังสีส่องผ่านไปได้ จากการที่อวัยวะต่างๆ ในร่างกายมีความสามารถในการดูดกลืนรังสีที่แตกต่างกัน ทำให้ระดับความเข้มของภาพเอกซเรย์ในแต่ละตำแหน่งมีค่าแตกต่างกันด้วย [13] โดยปกติรังสีเอกซเรย์นั้นจะมีปฏิสัมพันธ์กับสสารเกิดเป็นปรากฏการณ์สำคัญที่ส่งผลต่อภาพถ่ายรังสีแสดงดังรูปที่ 4



รูปที่ 4 ปฏิกิริยาการกระเจิงรังสีเอกซ์เรย์กับสสารที่รังสีเคลื่อนผ่าน

การดูดกลืนรังสีของสสารเป็นตัวแปรที่สำคัญต่อการประมวลผลภาพถ่ายรังสี เนื่องจากความเข้มข้นของรังสีที่ตกกระทบกับตัวรับภาพแปรผกผันกับอัตราการดูดกลืนของสสารที่รังสีเคลื่อนผ่าน ความสัมพันธ์ระหว่างความเข้มข้นของรังสีกับความสามารถในการดูดกลืนของสสารนั้น สามารถอธิบายได้ด้วยกฎของ Beer-Lambert กล่าวไว้ว่า อัตราการลดทอนลงหรือการดูดกลืนของรังสีเมื่อเคลื่อนที่ผ่านสสารใดๆ นั้น มีความสัมพันธ์เชิงเส้นกับค่าสัมประสิทธิ์การลดทอนของสสาร (attenuation coefficient) และความหนาของสสาร[16] ตามรูปที่ 5 และสมการที่ 1



รูปที่ 5 อธิบายการดูดกลืนรังสีตามกฎของ Beer-Lambert

$$A = \log \left( \frac{I_0}{I} \right) = \mu l = \left( \frac{\mu}{\rho} \right) \lambda \quad \text{สมการที่ 1}$$

โดยที่  $A$  คือ ค่าการดูดกลืนรังสี (Absortivity) ของสสาร (เท่า)

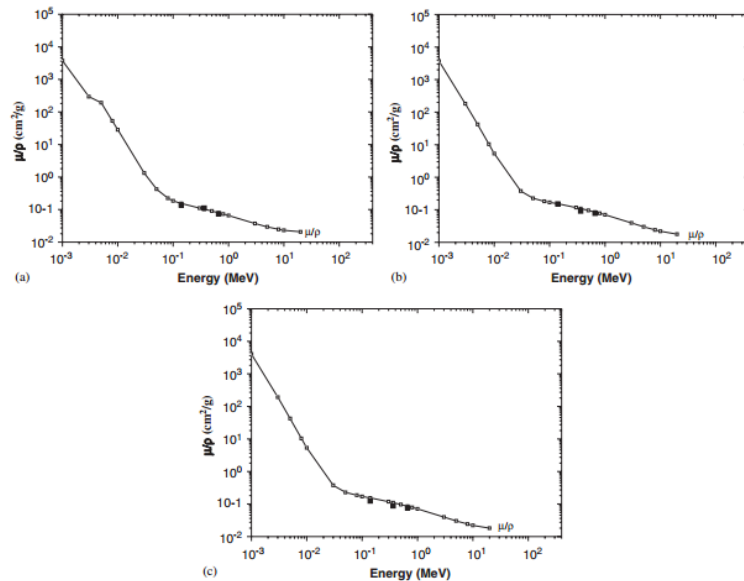
$I_0$  และ  $I$  คือค่าความเข้มข้นของรังสีตกกระทบและส่งผ่าน ตามลำดับ

$\mu$  คือ ค่าสัมประสิทธิ์การลดทอนเชิงเส้น (linear attenuation) ( $\text{cm}^{-1}$ ) จากการทดลอง ทั่วๆไป จะไม่สามารถหาค่า linear attenuation ได้โดยตรง แต่จะได้ค่า mass attenuation ( $\frac{\mu}{\rho}$ ) มาแทนแสดงดังรูปที่ 6 [17]

$l$  คือ ความหนาหรือความลึกของสสาร ณ ตำแหน่งที่รังสีเคลื่อนที่ผ่าน (cm)

$\rho$  คือ ความหนาแน่นของสสาร ( $\text{g/cm}^3$ )

$\lambda$  คือ ความหนาแน่นเชิงพื้นที่ (Area density) ของสสาร ( $\text{g}/\text{cm}^2$ ) [16]



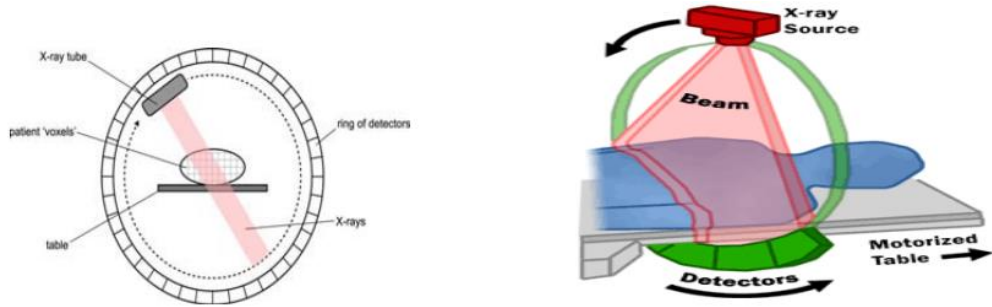
รูปที่ 6 ค่า mass attenuation ของ (a) กระจุก ( $\rho = 1.920 \text{ g}/\text{cm}^3$ ) (b) ก๊าซเนื้อ ( $\rho = 1.044 \text{ g}/\text{cm}^3$ ) และ (c) น้ำ ( $\rho = 1.000 \text{ g}/\text{cm}^3$ ) เมื่อรังสีเอกซเรย์ส่องผ่านที่ค่าระดับพลังงานต่างๆ [17]

ปัจจุบัน ภาพถ่ายรังสีเอกซเรย์เป็นที่นิยมใช้กันอย่างแพร่หลาย เนื่องจากเป็นการวินิจฉัยที่ง่าย รวดเร็ว และผู้ป่วยได้รับปริมาณรังสีที่ต่ำ จึงมีประโยชน์อย่างมากต่อการวินิจฉัยโรคต่างๆ เช่น โรคทางกระดูก ปอด ช่องท้อง ฟัน ลำคอ และอื่นๆ อีกมากมาย และยังมีการนำภาพถ่ายรังสีดังกล่าวไปประยุกต์ใช้ในการนำทางระหว่างผ่าตัด (Intraoperative surgical navigation) เช่น การผ่าตัดใส่โลหะยึดตรึงกระดูกแบบแผลปิด เป็นต้น

## 2.2.2 ภาพถ่ายรังสีเอกซเรย์คอมพิวเตอร์ (Computed Tomography, CT)

เอกซเรย์คอมพิวเตอร์เป็นเทคโนโลยีที่ต่อยอดมาจากการถ่ายภาพรังสีเอกซเรย์ เป็นการรวมเครื่องถ่ายภาพรังสีเอกซเรย์เข้ากับกลไกหมุนถ่ายภาพแบบวงกลม 360 องศา (rotating frame) เพื่อเก็บภาพตัดขวางแบบรอบด้านแสดงดังรูปที่ 7 ทุกครั้งที่กลไกหมุนเพื่อฉายรังสีครบหนึ่งรอบจะได้ภาพตัดขวางหนึ่งภาพ ตัวเครื่องสามารถเลื่อนเก็บภาพตัดขวางตามแนวยาวของผู้ป่วยทำให้ได้ภาพตัดขวางจำนวนหลายๆ ภาพ สามารถนำมาเรียงต่อกันและประมวลผลเพื่อสร้างโมเดลสามมิติ (3D reconstruction) ด้วย อัลกอริทึมบนระบบคอมพิวเตอร์ [13] ให้ความแม่นยำและรายละเอียดของกระดูกที่สูงที่สุดในการถ่ายภาพทางการแพทย์ทั้งหมด มีประโยชน์อย่างยิ่งต่อการวินิจฉัยโรค การวางแผนทำการรักษาหรือผ่าตัด และเป็นส่วนสำคัญที่นำไปใช้ต่อ

ยอดในระบบคอมพิวเตอร์ช่วยการผ่าตัด ตัวอย่างภาพถ่ายรังสีเอกซเรย์คอมพิวเตอร์แสดงดังรูปที่



(a) แบบรังสีขนาน (Parallel beam) (b) แบบรังสีรูปพัด (CT fan beam)

รูปที่ 7 แผนภาพเครื่องถ่ายภาพรังสีเอกซเรย์คอมพิวเตอร์ [13]



รูปที่ 8 ตัวอย่างภาพถ่ายรังสีเอกซเรย์คอมพิวเตอร์ [13]

### 2.3.3 ความสำคัญของภาพสามมิติกับการวางแผนการรักษา

ภาพถ่ายรังสีทางการแพทย์ที่อยู่ในรูปแบบสามมิติ เช่น ภาพถ่าย CT หรือ MRI นั้นมีประโยชน์อย่างมากต่อการวางแผนก่อนการรักษ (Preoperative planning) ช่วยให้บุคลากรทางการแพทย์สามารถวินิจฉัยลักษณะการแตกหักได้อย่างถูกต้องแม่นยำกว่าเมื่อเทียบกับภาพถ่ายรังสีที่อยู่ในรูปแบบสองมิติ เช่น X-ray หรือ Fluoroscopic เนื่องจากรูปสามมิติให้ข้อมูลได้มากกว่าภาพสองมิติ แต่ยังมีข้อเสียที่การถ่ายภาพสามมิตินั้นใช้เวลาานานกว่า และผู้ป่วยจะได้รับปริมาณรังสีที่มากกว่าซึ่งจะกล่าวถึงในหัวข้อถัดไป

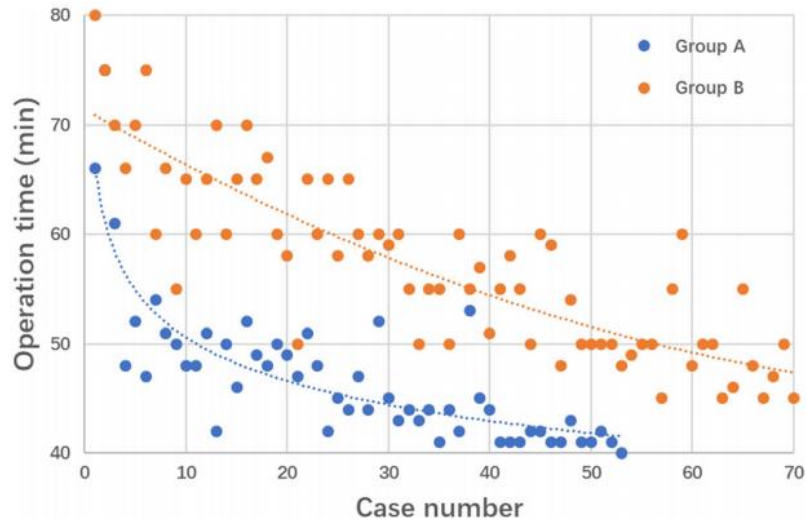


ในหัวข้อนี้เราจะอธิบายถึงความสำคัญของภาพสามมิติที่มีผลต่อการรักษา โดยมีงานวิจัยที่ศึกษาเกี่ยวกับการเทียบเปรียบเทียบผลการรักษาจากการผ่าตัดจัดกระดูกแบบแผลปิดของกระดูกต้นขา ระหว่างศัลยแพทย์ฝึกหัดสองกลุ่มซึ่งยังไม่มีประสบการณ์การผ่าตัดจัดกระดูกมาก่อน โดยศัลยแพทย์กลุ่ม A ได้รับภาพถ่ายสามมิติ (ภาพ CT) (n = 53) ในการวางแผนการผ่าตัดดังรูปที่ 9(ก) ส่วนศัลยแพทย์กลุ่ม B ได้รับภาพถ่ายสองมิติ (ภาพ X-ray) (n = 72) ในการวางแผนการผ่าตัดเท่านั้นดังรูปที่ 9(ข) งานวิจัยนี้เป็นการศึกษาแบบเก็บข้อมูลย้อนหลัง (Retrospective study) โดยข้อมูลที่จะนำมาเปรียบเทียบ คือ เวลาที่ใช้ในการผ่าตัด (operating time), ปริมาณเลือดที่เสียไประหว่างผ่าตัด (blood loss), จำนวนครั้งที่ต้องถ่ายภาพรังสี fluoroscopic ในระหว่างผ่าตัด ระยะเวลาเฝ้าติดตามอาการหลังจากการผ่าตัด (follow-up time) และระยะเวลาที่เชื่อมติดกันของกระดูกที่หัก (fracture healing time)



(ก) กลุ่มแรก: ภาพ reconstructed จากภาพถ่าย CT (ข) กลุ่มสอง: ภาพถ่ายรังสี X-ray  
รูปที่ 9 ข้อมูลภาพถ่ายรังสีที่ศัลยแพทย์ใช้สำหรับการวางแผนการผ่าตัด [18]

จากการศึกษาพบว่าศัลยแพทย์กลุ่ม A ใช้เวลาในการผ่าตัดเร็วกว่าศัลยแพทย์กลุ่ม B ระยะเวลาที่ใช้ในการผ่าตัดของศัลยแพทย์กลุ่ม A จะลดลงอย่างรวดเร็วเมื่อได้ทำการผ่าตัดบ่อยๆ หรือมีประสบการณ์ผ่าตัดมากขึ้น เมื่อเทียบกับศัลยแพทย์กลุ่ม B ที่ระยะเวลาในการผ่าตัดจะค่อยๆ ลดลงเป็นเส้นตรงมีประสบการณ์มากขึ้นแสดงดังรูปที่ 10 และพบว่ากรณีที่ศัลยแพทย์กลุ่ม A ได้เห็นรูปทรงสามมิติของกระดูกต้นขาที่แตกหักเพื่อวางแผนการผ่าตัดก่อน ยังช่วยลดการเสียเลือดของผู้ป่วย และจำนวนครั้งในการถ่ายภาพรังสีในระหว่างผ่าตัดน้อยลงเมื่อเทียบกับศัลยแพทย์กลุ่ม B ดังตารางที่ 1 เนื่องจากศัลยแพทย์กลุ่ม A ได้เห็นและเข้าใจลักษณะรอยแตกหักในสามมิติ (เทียบกับศัลยแพทย์กลุ่ม B ได้เห็นเพียงรอยแตกหักในสองมิติเท่านั้น) ทำให้สามารถการผ่าตัดได้อย่างรวดเร็ว แม่นยำขึ้น โดยที่พึงพาภาพถ่ายในระหว่างผ่าตัดน้อยลง



รูปที่ 10 Learning curve ความเร็วในการผ่าตัดระหว่างศัลยแพทย์กลุ่ม A และ B [18]

ตารางที่ 1 ผลลัพธ์จากการผ่าตัดเปรียบเทียบระหว่างศัลยแพทย์กลุ่ม A และ B [18]

	Group A (n = 53)	Group B (n = 72)	P value
Operation time (min)	45.00(42.00, 50.00)	55.00 (50.00, 60.00)	< 0.01*
Blood loss (mL)	160.00 (140.00, 170.00)	250.00 (195.00, 279.50)	< 0.01*
Radiation frequency (times)	13.02 ± 2.32	20.92 ± 3.27	< 0.01*
Postoperative complications			
Wound infection	1	0	
Urinary-tract infection	0	3	
Cerebral infraction	1	0	
Death	2	4	
Complication rate (%)	7.55% (4/53)	9.72% (7/72)	0.671

Data are median (25% quartile, 75% quartile) or mean ± SD




\*significant difference (P < 0.01)




### 2.2.4 ความเสี่ยงจากการถ่ายภาพรังสีทางการแพทย์ (Risk of X-ray imaging)

เป็นที่ทราบกันอยู่แล้วว่าภาพถ่ายรังสีทางการแพทย์มีประโยชน์มากมายต่อการวินิจฉัยโรคทางการแพทย์ แต่ก็ยังมีข้อเสียอยู่เช่นกัน คือ รังสีเอกซเรย์นั้นเป็นการแผ่รังสีแบบคลื่นแม่เหล็กไฟฟ้าที่มีพลังงานสูง เมื่อนำมาฉายผ่านผู้ป่วยจะมีรังสีส่วนหนึ่งสามารถทะลุผ่านร่างกายไปได้ และมีอีกส่วนหนึ่งคือปริมาณรังสีที่ได้รับ (radiation dose) ซึ่งรังสีส่วนนี้สามารถก่อให้เกิดผลกระทบทางชีววิทยาต่อผู้ที่ได้รับ แบ่งได้เป็นผลแบบที่ปรากฏชัดเจน (Deterministic effect) เช่น เป็นผื่นแดง, ผม่วรง, คลื่นไส้, ท้องร่วง, เป็นหมัน, การฟ่อของอวัยวะ, ฟังผิดในเนื้อเยื่อและอาจเสียชีวิตได้ เป็นต้น ซึ่งเกิดจากการได้รับรังสี ในปริมาณมากๆ และผลกระทบแบบสุ่ม (Stochastic effect) เช่น เกิดโรคมะเร็งในระยะยาว และเกิดการเปลี่ยนแปลงทางพันธุกรรม ซึ่งเกิดจากการได้รับรังสีปริมาณต่ำในช่วงเวลานาน [19] เมื่อฉายรังสี อวัยวะแต่ละส่วนจะได้รับปริมาณรังสีที่แตกต่างกันเนื่องจากมีความไวต่อรังสีที่ไม่เท่ากัน ค่าปริมาณรังสีที่ได้รับนี้จะเรียกว่า ค่าปริมาณ

รังสียังผล (Effective dose) เป็นผลรวมของของปริมาณรังสีสมมูลหลังปรับเทียบสภาพความไวต่อรังสีของแต่ละเนื้อเยื่อทั่วร่างกาย มีหน่วยเป็น MilliSievert (mSv) ข้อมูลจาก International Commission on Radiological Protection (ICRP) ได้กำหนดค่าปริมาณรังสีสูงสุดที่รับได้ (maximum permissible dose) สำหรับเจ้าหน้าที่ผู้ปฏิบัติงานด้านรังสีเท่ากับ 20 mSv ต่อปี โดยเฉลี่ยติดต่อกัน 5 ปี (แต่ละปีต้องรับรังสีได้ไม่เกิน 50 mSv) ซึ่งเป็นปริมาณรังสีที่ไม่ก่อให้เกิด deterministic effect ใดๆ และเสี่ยงต่อการเกิดมะเร็งหรือการผ่าเหล่าในอัตราที่ไม่แตกต่างกับคนปกติทั่วไป ตารางที่ 2 สรุปค่าปริมาณรังสียังผลจากการวินิจฉัยโรค ณ อวัยวะที่แตกต่างกัน และช่วงเวลาเชิงเปรียบเทียบกับปริมาณรังสีพื้นหลังที่คนทั่วไปได้รับตามธรรมชาติ (natural background radiation) โดยเฉลี่ยแล้วคนเราจะได้รับรังสีพื้นหลังประมาณ 3-4.5 mSv ต่อปี (จะ ได้รับปริมาณรังสีมากหรือน้อยขึ้นอยู่กับบริเวณที่อยู่อาศัย) [9]

ตารางที่ 2 ปริมาณรังสีที่ได้รับจากการวินิจฉัยโรคตามอวัยวะบริเวณต่างๆ กัน [9]

	Procedure	Approximate effective radiation dose	Comparable to natural background radiation for:
<b>ABDOMINAL REGION</b> 	Computed Tomography (CT)–Abdomen and Pelvis	10 mSv	3 years
	Computed Tomography (CT)–Abdomen and Pelvis, repeated with and without contrast material	20 mSv	7 years
	Computed Tomography (CT)–Colonography	6 mSv	2 years
	Intravenous Pyelogram (IVP)	3 mSv	1 year
	Barium Enema (Lower GI X-ray)	8 mSv	3 years
	Upper GI Study with Barium	6 mSv	2 years
<b>BONE</b> 	Spine X-ray	1.5 mSv	6 months
	Extremity (hand, foot, etc.) X-ray	0.001 mSv	3 hours
<b>CENTRAL NERVOUS SYSTEM</b> 	Computed Tomography (CT)–Head	2 mSv	8 months
	Computed Tomography (CT)–Head, repeated with and without contrast material	4 mSv	16 months
	Computed Tomography (CT)–Spine	6 mSv	2 years

<b>CHEST</b>	<b>Procedure</b>	<b>Approximate effective radiation dose</b>	<b>Comparable to natural background radiation for:</b>
	Computed Tomography (CT)–Chest	7 mSv	2 years
	Computed Tomography (CT)–Lung Cancer Screening	1.5 mSv	6 months
	Chest X-ray	0.1 mSv	10 days
<b>DENTAL</b>	<b>Procedure</b>	<b>Approximate effective radiation dose</b>	<b>Comparable to natural background radiation for:</b>
	Dental X-ray	0.005 mSv	1 day
<b>HEART</b>	<b>Procedure</b>	<b>Approximate effective radiation dose</b>	<b>Comparable to natural background radiation for:</b>
	Coronary Computed Tomography Angiography (CTA)	12 mSv	4 years
	Cardiac CT for Calcium Scoring	3 mSv	1 year

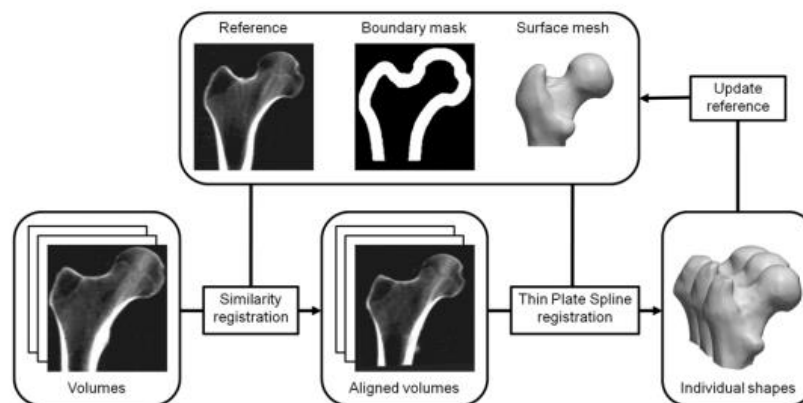
จากตารางที่ 2 หากเปรียบเทียบระหว่างการถ่ายภาพรังสีเอกซเรย์ กับ เอกซเรย์คอมพิวเตอร์ สำหรับวินิจฉัยโรคทางกระดูก จะพบว่า ปริมาณรังสียังผลที่ผู้ป่วยได้รับจากการถ่ายภาพรังสีเอกซเรย์มีค่าประมาณ 0.001 mSv (เทียบปริมาณรังสีพื้นหลังตามธรรมชาติ 3 เดือน) แต่สำหรับการถ่ายภาพรังสีเอกซเรย์คอมพิวเตอร์นั้นผู้ป่วยจะได้รับปริมาณรังสีสูงถึง 10-20 mSv (เทียบปริมาณรังสีพื้นหลังตามธรรมชาติ 3-7 ปี) [9] ซึ่งผู้ป่วยจะได้รับปริมาณรังสีที่สูงกว่าการถ่ายภาพรังสีเอกซเรย์อย่างมาก จึงเพิ่มโอกาสเกิดมะเร็งมากกว่า มีค่าใช้จ่ายที่สูงกว่า และเครื่อง CT scan ส่วนมากไม่สามารถถ่ายภาพในระหว่างผ่าตัดได้

### 2.3 สร้างโมเดลสามมิติด้วยแบบจำลองรูปร่างเชิงสถิติ (3D Reconstruction using Statistic shape modeling)

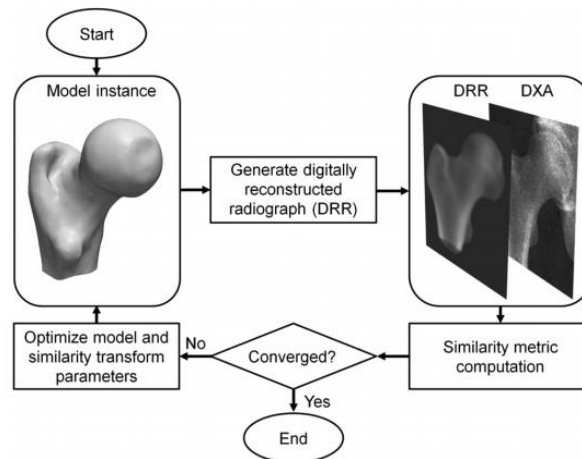
ในหลายทศวรรษที่ผ่านมา การสร้างโมเดลสามมิติของกระดูกต้นขาจากภาพเอกซเรย์ในงานวิจัยส่วนใหญ่เป็นการสร้างจากแบบจำลองรูปร่างเชิงสถิติโดยใช้รูปทรงเป็นหลัก (Statistic shape model หรือ SSM) โดยการเรียนรู้รูปร่างของวัตถุที่สนใจจากลักษณะรูปร่าง (Shape constraint variable) วิธีนี้จะเก็บข้อมูลโมเดลสามมิติที่สนใจหลายๆ โมเดลมาทำการหาค่าเฉลี่ยของตัวแปรที่กำหนดโหมด (mode) [20] ตัวอย่างงานวิจัยที่ประสบผลสำเร็จ เช่น งานวิจัยของ Tristan Whitmarsh, et. al. วิจัยเกี่ยวกับการสร้างโมเดลสามมิติของกระดูกต้นขาส่วนบนพร้อมทั้งการกระจายตัวของความหนาแน่นกระดูก [21] มีขบวนการสร้างโมเดลสามมิติหลักๆ สองขั้นตอนคือ

1) การสร้างรูปทรงทางสถิติของกระดูกต้นขา (Statistic Shape model construction) ดังรูปที่ 11 รูปทรงกระดูกต้นขาจะอยู่ในรูปแบบของ Polygon surface หรือ Surface mesh โดยมีจุดที่กำหนดตำแหน่งบน surface เป็นตำแหน่ง landmark ของโมเดลกระดูกต้นขา การสร้างโมเดล SSM จะเป็นการนำตำแหน่ง landmark ของโมเดลสามมิติของกระดูกแต่ละชิ้นในกลุ่มตัวอย่างมาหาตำแหน่งเฉลี่ยของแต่ละ landmark โดยใช้ Principal component analysis (PCA) ซึ่งจะได้เซตของค่าเฉลี่ยและความเบี่ยงเบนมาตรฐานของ landmark แต่ละตำแหน่ง SSM จะมีรูปทรงเบื้องต้นตามค่าเฉลี่ยของ landmark ที่คำนวณได้ และ SSM สามารถเปลี่ยนรูปทรงไปตาม landmark ที่เปลี่ยนแปลงไปค่าตามความเบี่ยงเบนมาตรฐาน

2) การซ้อนทับโมเดลสามมิติเข้ากับภาพเอกซเรย์ (2D/3D registration) ซึ่งเป็นการหาความเหมือน (Similarity measure) ของระหว่างภาพเอกซเรย์และภาพเอกซเรย์สังเคราะห์ของ Statistic model หรือเรียกว่า Digitally reconstructed radiograph (DRR) ดังรูปที่ 12 ในส่วนนี้จะเป็นการทำการซ้อนทับแบบวนซ้ำ (Iteration) แต่ละรอบของการทำงานเริ่มจากการนำ SSM มาสร้างภาพ DRR เพื่อนำไปเปรียบเทียบและวัดค่าความเหมือน (Similarity measure) กับภาพ X-ray (Input) ถ้าสองภาพนั้นยังไม่เหมือนกัน จะทำการหาค่าตำแหน่งของ landmark ของ SSM ที่เหมาะสมที่สุด (Optimization) เพื่อให้ค่าความเหมือนมีค่าน้อยที่สุด ทำเช่นนี้ซ้ำไปเรื่อยๆ จนกระทั่งเกิดการลู่เข้า และโมเดล SSM ในการทำซ้ำครั้งสุดท้ายจะเป็นโมเดลสามมิติผลลัพธ์



รูปที่ 11 การสร้างรูปร่างเฉลี่ยของกระดูกต้นขา (Shape model construction) [21]



รูปที่ 12 ขบวนการซ้อนทับภาพระหว่างโมเดลสามมิติกับภาพเอกซเรย์ [21]

วิธี Statistic shape model มีข้อดีตรงที่สามารถสร้างรูปร่างที่มีความแม่นยำได้สูง และใช้จำนวนข้อมูลหรือตัวอย่างสอนสำหรับสร้างโมเดลที่น้อยกว่าวิธีอื่นๆ แต่จะมีข้อเสีย คือ วิธีนี้สามารถใช้สร้างกระดูกที่มีความสมบูรณ์ได้เพียงอย่างเดียว ไม่สามารถนำไปใช้สร้างโมเดลกระดูกที่ไม่สมบูรณ์ เช่น มีการแตกหัก หรือ บิดงอผิดรูปได้

## 2.4 การสร้างโมเดลสามมิติจากการเรียนรู้ของเครื่อง (Machine learning)

### 2.4.1 ความรู้พื้นฐานเกี่ยวกับการเรียนรู้ของเครื่อง (Basic Knowledge of Machine Learning)

การเรียนรู้ของเครื่องเป็นเทคนิคการแก้ปัญหาทางคอมพิวเตอร์ระดับสูง ใช้ในการเรียนรู้ความสัมพันธ์บางอย่างที่สำคัญในกลุ่มตัวอย่างเพื่อทำหน้าที่ทำนายผลลัพธ์ (Prediction) ที่ซับซ้อนเกินกว่าจะหาผลลัพธ์ที่แม่นยำได้ ลักษณะการสอนให้เครื่องสามารถเรียนรู้ได้จะเป็นการใส่ข้อมูลสอน (Training sample) ซึ่งประกอบได้ด้วยข้อมูลขาเข้า (Input)  $x$  และข้อมูลขาออก (Output)  $y$  ที่สอดคล้องกันเป็นจำนวนหลายๆ ตัว เพื่อให้อัลกอริทึมการเรียนรู้สามารถคำนวณเพื่อจับความสัมพันธ์ของข้อมูลสอนออกมาเป็นฟังก์ชันสมมติฐาน (Hypothesis)  $h$  ซึ่งจะแตกต่างจากการพัฒนาโปรแกรมแบบดั้งเดิม (Traditional programming) ตรงที่ผู้พัฒนาโปรแกรมไม่จำเป็นต้องนิยามฟังก์ชันสมมติฐานเพื่อให้คำนวณข้อมูลขาออกเองโดยตรง แต่จะเป็นให้อัลกอริทึมเรียนรู้ความสัมพันธ์ระหว่างข้อมูลขาเข้าและข้อมูลขาออกเอง การเรียนรู้ของเครื่องนั้นสามารถแบ่งออกเป็น 3 กลุ่ม ตามลักษณะของข้อมูลสอน (Training sample) คือ Supervised learning เช่น สร้างฟังก์ชันถดถอย (Regression) การแยกแยะข้อมูล (Classification), Unsupervised

learning เช่น การจัดกลุ่มข้อมูล (Clustering) และ Reinforcement learning การเรียนรู้แบบเสริมกำลัง โดยในงานวิจัยนี้จะสนใจเฉพาะ Supervised learning เท่านั้น

### โครงข่ายประสาทเทียม (Artificial neural network หรือ Perceptron neural network)

โครงข่ายประสาทเทียมเป็นการจำลองการทำงานหรือการประมวลผลของสมองมนุษย์ ซึ่งมีเพอร์เซปตรอน (Perceptron) เป็นแกนหลักแสดงดังรูปที่ 13 โดยจะรับข้อมูลขาเข้า  $x_1, x_2, x_3, \dots, x_n$  ข้อมูลขาเข้า  $x_i$  นั้นแสดงถึงคุณสมบัติเฉพาะของข้อมูลขาเข้า (feature หรือ attribute) จากนั้นข้อมูลขาเข้าแต่ละตัวจะถูกคูณด้วยค่าน้ำหนัก (weight) เฉพาะแต่ละตัวของมัน ซึ่งมีค่าเป็น  $W_0, W_1, W_2, W_3, \dots, W_n$  ซึ่งมี  $W_0$  เป็นค่าน้ำหนักเบื้องต้น (bias weight) จากนั้นทำการบวกรวมกัน ของค่าข้อมูลขาเข้า  $x_i$  คูณกับ  $W_i$  ตามสมการต่อไปนี้ โดยที่ตัวแปร  $i$  แทนจำนวนของคุณสมบัติเฉพาะของข้อมูลขาเข้า

$$\sum_{i=0}^n W_i x_i = W_0 + W_1 x_1 + W_2 x_2 + \dots + W_n x_n$$

จากนั้นผลรวมทั้งหมดจะถูกป้อนเข้าไปในฟังก์ชันกระตุ้น (Activation function)  $g$  เพื่อคำนวณข้อมูลขาออก  $y$  หรือฟังก์ชันสมมติฐาน  $h$  ตามสมการต่อไปนี้

$$y = h = g(W_0 + W_1 x_1 + W_2 x_2 + \dots + W_n x_n)$$

ฟังก์ชันกระตุ้น  $g$  จะทำหน้าที่ในการแบ่งขอบเขตของข้อมูลขาออก  $y$  ซึ่งมีให้เลือกหลากหลายประเภทตามลักษณะของข้อมูลขาออกตามที่ต้องการ เช่น Sigmoid function ให้ค่า 0 หรือ 1, Rectified linear unit ให้ค่า 0 หรือ  $x$  แสดงดังรูปที่ 14 ตัวอย่างของฟังก์ชันกระตุ้นมีดังต่อไปนี้

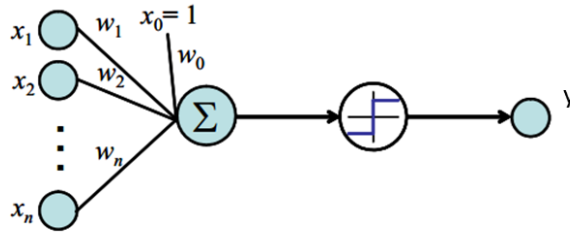
1) Sigmoid function ให้ค่าข้อมูลขาออกเป็น 0 เมื่อข้อมูลขาเข้าน้อยกว่า 0 และให้ค่าข้อมูลขาออกเป็น 1 เมื่อข้อมูลขาเข้ามีค่ามากกว่า 0 สามารถคำนวณค่าได้ตามสมการด้านล่าง

$$g_{sigmoid}(x) = \sigma = \frac{1}{1 + e^{-x}}$$

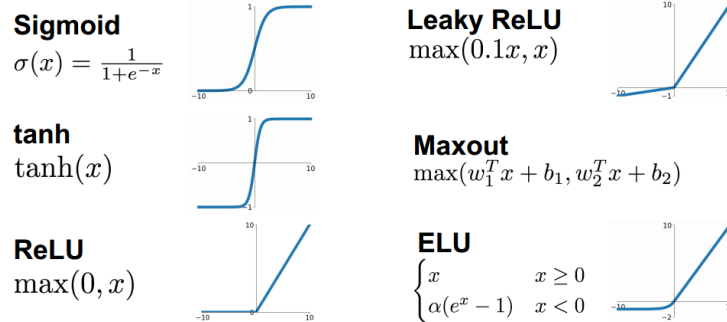
2) Rectified linear unit function ให้ค่าข้อมูลขาออกเป็น 0 เมื่อข้อมูลขาเข้าน้อยกว่า 0 และให้ค่าข้อมูลขาออกเป็น  $x$  เมื่อข้อมูลขาเข้ามีค่ามากกว่า 0 สามารถคำนวณค่าได้ตามสมการด้านล่าง

$$g_{ReLU}(x) = \begin{cases} 0 & ; x < 0 \\ x & ; x \geq 0 \end{cases}$$

ข้อมูลขาออกของ  $y$  ที่คำนวณได้อาจอยู่ในรูปของ ค่าตัวแทนของชนิด (class label) หรือ ค่าทำนาย (predicted value) เป็นต้น เราจะเรียกการคำนวณโดยเริ่มจากการใส่ข้อมูลขาเข้ามา คำนวณในเพอร์เซปตรอนตั้งแต่ระดับชั้นแรกไปเรื่อยๆ จนกระทั่งได้ข้อมูลขาออกนี้ว่า การแพร่กระจายไปข้างหน้า (Forward propagation)



รูปที่ 13 เพอร์เซปตรอนหนึ่งหน่วย (Perceptron unit) [22]



รูปที่ 14 ฟังก์ชันกระตุ้น (Activation function) [22]

ฟังก์ชันความสูญเสีย (Loss function)

ฟังก์ชันของค่าความผิดพลาดที่คำนวณได้จากเพอร์เซปตรอนถูกเรียกว่า ฟังก์ชันความสูญเสีย  $J$  ซึ่งสามารถคำนวณได้หลายวิธีขึ้นอยู่กับ เช่น

สำหรับปัญหาการถดถอยจะคำนวณค่าความสูญเสียโดยหาผลรวมของค่าความผิดพลาดระหว่างค่าข้อมูลขาออกที่ต้องการกับข้อมูลขาออกที่คำนวณได้แล้วยกกำลังสอง จะได้ฟังก์ชันความสูญเสียตามเป็นฟังก์ชันนูน (Convex function) สมการด้านล่าง มีข้อดีตรงที่เป็นฟังก์ชันที่สามารถหาอนุพันธ์ได้ (Differentiable function) และมีจุดต่ำสุดสัมบูรณ์ (Global minimum) เพียงจุดเดียว

$$J = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$$

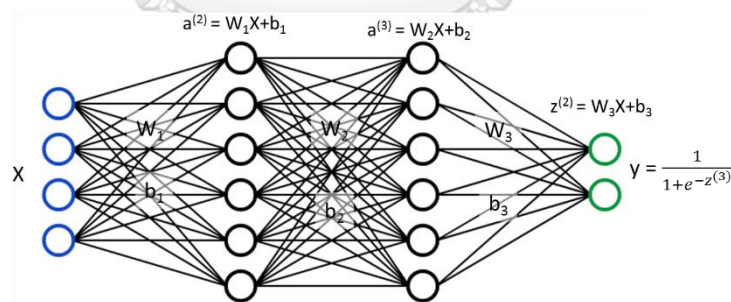


หรือ สำหรับปัญหาการจำแนกประเภท (Classification problem) จะนิยามโดยใช้ฟังก์ชันเอนโทรปี (Entropy) ของตัวแปรสุ่ม ซึ่งนิยามตามสมการต่อไปนี้

$$J = -\frac{1}{m} \left[ \sum_{i=1}^m y^i \log h_{\theta}(x^{(i)}) + (1 - y^i) \log (1 - h_{\theta}(x^{(i)})) \right]$$

ค่าฟังก์ชันสูญเสียที่ยกตัวอย่างไว้ข้างต้นจะถูกนำไปคำนวณเพื่อปรับค่าน้ำหนักของเพอร์เซปตรอน ซึ่งจะถูกอธิบายอีกครั้งในย่อหน้าถัดไป

ในทางปฏิบัติ สำหรับการพัฒนาระบบการเรียนรู้ของเครื่องจะไม่ใช้เพอร์เซปตรอนเพียงแค่ตัวเดียว แต่จะนำเพอร์เซปตรอนหลายตัวมาต่อกันเป็นโครงข่ายของประสาทเทียมที่มีความลึกหลายๆ ระดับชั้น (Layer) แต่ละชั้นมีจำนวนเพอร์เซปตรอนมากกว่าหนึ่งตัว สามารถแก้ปัญหาที่ซับซ้อนได้มากยิ่งขึ้น เรียกว่า โครงข่ายของประสาทเทียมที่มีความลึกหลายๆ ระดับชั้นเช่นนี้ว่าเป็นการเรียนรู้เชิงลึก (Deep learning) ดังรูปที่ 15 จะเห็นว่าในระดับชั้นที่  $l$  ใดๆ จะมีการเชื่อมต่อข้อมูลขาเข้า  $x$  หรือ  $a^l$  แต่ละตัวกับกับฟังก์ชันกระตุ้นแต่ละตัวในระดับชั้นถัดไป  $a^{l+1}$  ด้วยตัวแปรค่าน้ำหนัก  $W^l$  โครงสร้างโครงข่ายแบบนี้จะมีชื่อเรียกอีกอย่างหนึ่งว่าเป็น การเชื่อมต่อแบบสมบูรณ์ (Fully-connected layer) เนื่องจากโครงข่ายประสาทเทียมมีความซับซ้อนมากเราจึงขอนิยามตัวแปรได้ดังต่อไปนี้



รูปที่ 15 โครงข่ายการเรียนรู้เชิงลึก (Deep learning) [22]

- กำหนดให้  $L$  = จำนวนระดับชั้นของโครงข่ายประสาทเทียมทั้งหมด
- $s_l$  = จำนวนเพอร์เซปตรอนของระดับชั้นที่  $l$  ใดๆ
- $K$  = จำนวนชนิดของข้อมูลขาออก

นิยามให้

$W^{(j)}$  = เมทริกซ์ของค่าน้ำหนักเพอร์เซปตรอนตัวที่  $j$  ในระดับชั้นที่  $j$

$a_i^{(j)}$  = เมทริกซ์ของฟังก์ชันกระตุ้นของเพอร์เซปตรอนตัวที่  $i$  ในระดับชั้นที่  $j$

$y$  = ข้อมูลขาออก

$g$  = ฟังก์ชันกระตุ้นใดๆ เช่น ReLU, Sigmoid เป็นต้น

ตัวอย่างของเมทริกซ์ของฟังก์ชันกระตุ้นระดับชั้นที่ 2 และระดับชั้นที่  $j$  แสดงตามสมการด้านล่าง

$$a^{(j)} = \begin{bmatrix} a_1^{(j)} \\ a_2^{(j)} \\ \vdots \\ a_{s_l}^{(j)} \end{bmatrix} = \begin{bmatrix} g(W_{10}^{(j-1)} + W_{11}^{(j-1)} a_1^{(j-1)} + W_{12}^{(j-1)} a_2^{(j-1)} + \dots + W_{1s_{l-1}}^{(j-1)} a_{s_{l-1}}^{(j-1)}) \\ g(W_{20}^{(j-1)} + W_{21}^{(j-1)} a_1^{(j-1)} + W_{22}^{(j-1)} a_2^{(j-1)} + \dots + W_{2s_{l-1}}^{(j-1)} a_{s_{l-1}}^{(j-1)}) \\ \vdots \\ g(W_{s_l 0}^{(j-1)} + W_{s_l 1}^{(j-1)} a_1^{(j-1)} + W_{s_l 2}^{(j-1)} a_2^{(j-1)} + \dots + W_{s_l s_{l-1}}^{(j-1)} a_{s_{l-1}}^{(j-1)}) \end{bmatrix}$$

การแพร่กระจายย้อนกลับ(Back propagation)

อัลกอริทึมการแพร่กระจายย้อนกลับ ใช้ในการหาค่าน้ำหนักของเพอร์เซปตรอน ( $w_1, w_2, w_3, \dots, w_n$ ) ที่เหมาะสมที่สุดที่ทำให้ค่าความผิดพลาดของข้อมูลขาออกมีค่าต่ำที่สุด โดยใช้การคำนวณแบบทำซ้ำ (Iteration) จนมีการลู่ออกของผลลัพธ์ การแพร่กระจายย้อนกลับจะเป็นการคำนวณเพื่อปรับค่าน้ำหนัก โดยใช้วิธีการนำค่าอนุพันธ์อันดับที่หนึ่ง (First-order derivative) หรือ gradient ของฟังก์ชันที่ใช้แพร่กระจายไปข้างหน้าเทียบกับตัวแปรค่าน้ำหนักนั้นๆ เพื่อนำมาเป็นค่าอัตราการเปลี่ยนแปลงของค่าความสูญเสีย เมื่อเทียบกับตัวแปรค่าน้ำหนักใดๆ ในแต่ละระดับชั้นของเพอร์เซปตรอน อาจจะเป็นการหาอนุพันธ์อันดับที่หนึ่งสามารถทำได้จากการหาค่าอนุพันธ์โดยตรงตามหลักการแคลคูลัส หรือใช้ระเบียบวิธีเชิงตัวเลข (Numerical gradient) ตามสมการด้านล่างก็ได้ และมีการใช้กฎอนุพันธ์ลูกโซ่ (Chain rule) ร่วมด้วย

$$f(x) = \lim_{h \rightarrow \infty} \frac{f(x + 2h) - f(x - 2h)}{2h}$$

จากเพอร์เซปตรอนที่ให้มาตามรูปที่ 15 ซึ่งใช้ฟังก์ชันคำนวณแบบเชิงเส้นและใช้ sigmoid เป็นฟังก์ชันกระตุ้น เราสามารถทำการแพร่กระจายไปข้างหน้าได้ตามสมการต่อไปนี้

$$a^{(2)} = W_1 X + b_1$$

$$a^{(3)} = W_2 a^{(2)} + b_2$$

$$z^{(3)} = W_3 a^{(3)} + b_3$$

$$y = a^{(4)} = g(z^{(3)}) = \frac{1}{1 + e^{-z^{(3)}}}$$

และจะทำการแพร่กระจายย้อนกลับ โดยเริ่มจากการหาค่าอนุพันธ์อันดับที่หนึ่งของฟังก์ชันความสูญเสียเทียบกับตัวแปรค่าน้ำหนักของเพอร์เซปตรอนแต่ละชั้น เมื่อเราทำ back propagation จนครบทุกระดับชั้นของเพอร์เซปตรอนแล้ว เราจะได้ค่าอนุพันธ์อันดับที่หนึ่งของฟังก์ชันสูญเสียเมื่อเทียบกับค่าตัวแปรค่าน้ำหนักของแต่ละระดับชั้น คือ

$$\frac{\partial L}{\partial w_1} \quad \frac{\partial L}{\partial b_1} \quad \frac{\partial L}{\partial w_2} \quad \frac{\partial L}{\partial b_2} \quad \frac{\partial L}{\partial w_3} \quad \text{และ} \quad \frac{\partial L}{\partial b_3}$$

ซึ่งจะถูกนำไปใช้ในการปรับค่าน้ำหนักด้วยวิธี Stochastic Gradient descent ดังต่อไปนี้

$$\begin{aligned} W_1 &= W_1 + \alpha \frac{\partial L}{\partial W_1} & b_1 &= b_1 + \alpha \frac{\partial L}{\partial b_1} \\ W_2 &= W_2 + \alpha \frac{\partial L}{\partial W_2} & b_2 &= b_2 + \alpha \frac{\partial L}{\partial b_1} \\ W_3 &= W_3 + \alpha \frac{\partial L}{\partial W_3} & b_3 &= b_3 + \alpha \frac{\partial L}{\partial b_3} \end{aligned}$$

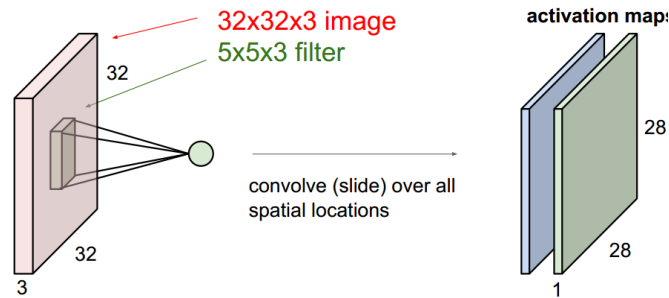
ซึ่ง  $\alpha$  คือ ค่าอัตราการเรียนรู้ ค่ามีน้อยๆ อยู่ในช่วง  $1 \times 10^{-6} - 1 \times 10^{-3}$  (ขึ้นอยู่กับลักษณะการเรียนรู้) เมื่อทำการคำนวณซ้ำโดยใช้ back propagation ไปเรื่อยๆ จะทำให้ค่าน้ำหนักแต่ละตัวมีค่าเปลี่ยนแปลงไปในทางที่ทำให้ค่าฟังก์ชันสูญเสีย (Error) มีค่าน้อยที่สุด ซึ่งจะทำให้การใช้เพอร์เซปตรอนคำนวณข้อมูลขาออกมีค่าแม่นยำมากที่สุด

### โครงข่ายประสาทแบบคอนโวลูชัน (Convolutional neural network)

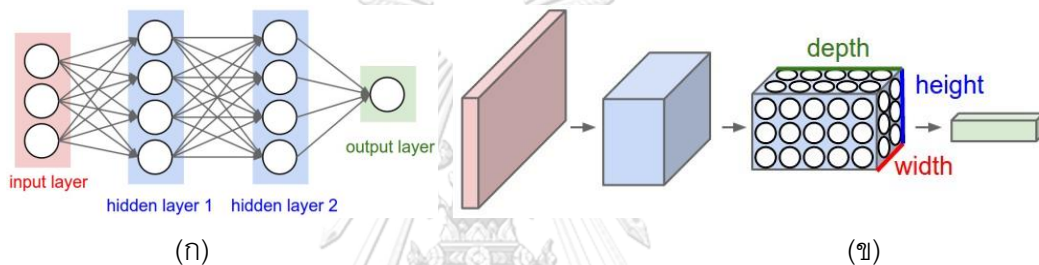
โครงข่ายประสาทแบบคอนโวลูชัน หรือเรียกสั้นๆ ว่า ConvNet นิยมใช้ในการประมวลผลข้อมูลขาเข้าที่อยู่ในรูปแบบของข้อมูลเชิงพื้นที่ (Spatial data) เช่น ภาพถ่ายสี ภาพเอกซเรย์ วิดีโอ และโมเดลสามมิติ เป็นต้น ประกอบไปด้วยค่าน้ำหนักต่างๆ ซึ่งใช้ในการคูณเชิงสเกลาร์ (dot product) กับคุณสมบัติของข้อมูลขาเข้า โดยค่าน้ำหนักของ ConvNet จะอยู่ในรูปแบบตัวกรอง (Filter) ขนาด  $F \times F$  คงที่ในแต่ละระดับชั้น (ในแต่ละระดับชั้นมีตัวกรองได้มากกว่า 1 ตัว จำนวนตัวกรองในแต่ละระดับชั้นเท่ากับ  $K$  ตัว) ข้อมูลขาเข้าจะถูกนำไปดำเนินการคอนโวลูชัน (Convolution operation) ซึ่งเป็นการนำตัวกรองที่มีขนาด  $F \times F$  ทำการคูณเชิงสเกลาร์กับข้อมูลขาเข้าที่ตำแหน่ง  $(x, y)$  ใดๆ และทำการเลื่อนตำแหน่งตัวกรองไปที่ตำแหน่งใกล้เคียงด้วยขนาดการเลื่อน (stride)  $S$  จะได้ข้อมูลขาออกที่อยู่ในรูปสองมิติหลายๆ ชุดซ้อนกันเท่ากับจำนวนตัวกรองที่ใช้ แสดงดังรูปที่ 16 การดำเนินการคอนโวลูชันจะเป็นการสกัดคุณสมบัติที่สำคัญของข้อมูลขาเข้า (feature extraction) เพื่อนำคุณสมบัตินี้ไปใช้ในการเรียนรู้ต่อไป

ข้อดีของ ConvNet คือ จะมีการแชร์ตัวแปรค่าน้ำหนักให้กับข้อมูลขาเข้าในแต่ละระดับชั้น มีจำนวนตัวแปรน้ำหนักเท่ากับขนาดและจำนวนของตัวกรอง คือ  $(F \times F) \times K + K$  ซึ่งจะใช้จำนวน

ตัวแปรค่าน้ำหนักจะมีจำนวนที่คงที่ตามที่ผู้ออกแบบกำหนดไว้โดยไม่ขึ้นอยู่กับขนาดของข้อมูลขาเข้า ทำให้มีจำนวนตัวแปรค่าน้ำหนักมีจำนวนที่น้อยกว่าโครงข่ายแบบเพอร์เซปตรอนอย่างมาก การเรียนรู้ข้อมูลข้อมูลประเภทภาพสองหรือสามมิติจึงสามารถทำได้โดยรวดเร็วมากยิ่งขึ้น [22]

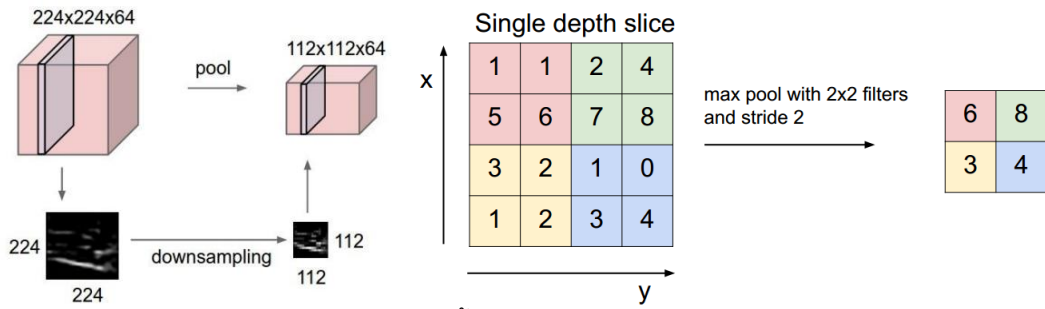


รูปที่ 16 โครงข่ายประสาทแบบคอนโวลูชัน (Convolutional neural network ; ConvNet) [22]

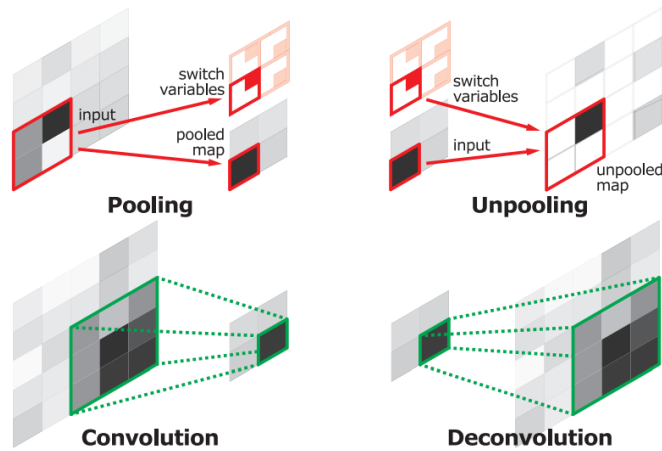


รูปที่ 17 การเปรียบเทียบโครงสร้าง (ก) fully-connected layer กับ (ข) Convolutional layer [22]

โดยทั่วไป ConvNet จะไม่ได้ใช้ในการประมวลผลเพียงอย่างเดียวในระบบโครงข่าย แต่จะทำงานร่วมกับระดับชั้นรวม (Pooling layer) ซึ่งจะทำหน้าที่รวบรวมข้อมูลเชิงพื้นที่ที่อยู่ใกล้ๆ กันภายในระยะตัวกรอง  $F \times F$  ด้วยการหาค่าเฉลี่ย (Average pooling) หรือ การค่ามากที่สุด (Max pooling) ของกลุ่มข้อมูลที่อยู่ภายในระยะ  $F \times F$  จะให้ข้อมูลขาออกที่มีขนาดลดลงเป็นจำนวนเท่าของค่า  $F$  ที่ใช้ในการทำ pooling ตามรูปที่ 18 ทั้ง ConvNet และ Pooling ที่ได้กล่าวไปข้างต้นจะเป็นการทำประมวลผลข้อมูลขาเข้าที่ทำให้ข้อมูลขาออกมีขนาดที่เล็กลง จะเรียกว่าเป็นการลดขนาดของตัวอย่าง (Down sampling) และจะมีโครงข่ายอีกชนิดสองชนิด คือ DeConvNet และ Unpooling จะทำงานแบบตรงข้ามกับ ConvNet และ Pooling จะเป็นการประมวลผลแบบเพื่อเพิ่มขนาดของข้อมูลตัวอย่าง (Up sampling) โดยจะประมวลผลและทำให้ข้อมูลขาออกมีขนาดที่ใหญ่ขึ้น แสดงดังรูปที่ 19

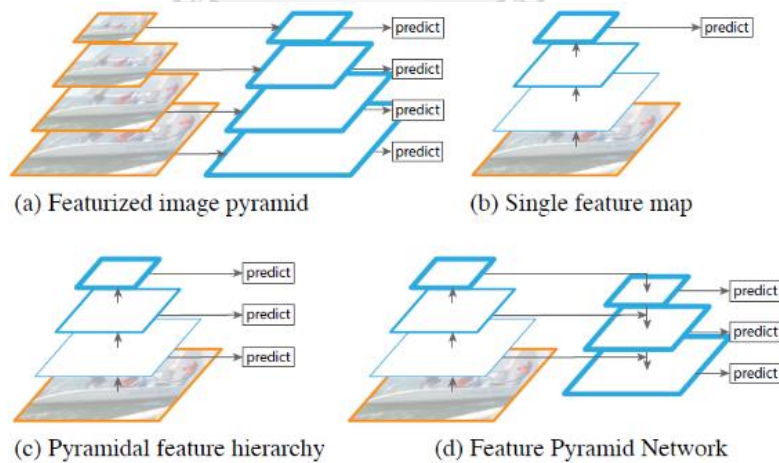


รูปที่ 18 ระดับชั้นรวม(Pooling layer) [22]



รูปที่ 19 เปรียบเทียบ Deconvolution และ Unpooling กับ Convolution และ Pooling [23]

รูปแบบโครงสร้างของ CNN (CNN Architecture type)



รูปที่ 20 โครงข่าย CNN ที่มีรูปแบบการเชื่อมต่อแบบพีระมิด [24]

ในปัจจุบันโครงข่าย CNN เป็นที่นิยมใช้ในการประมวลผลภาพและคอมพิวเตอร์วิทัศน์อย่างมาก จึงมีการวิจัยและคิดค้นโครงสร้างขึ้นมาหลากหลายรูปแบบรูปที่ 20 แต่ละรูปแบบมีลักษณะการเชื่อมต่อเพื่อประมวลผลและมีข้อดี-ข้อเสียที่แตกต่างกันไป มีรายละเอียดดังต่อไปนี้

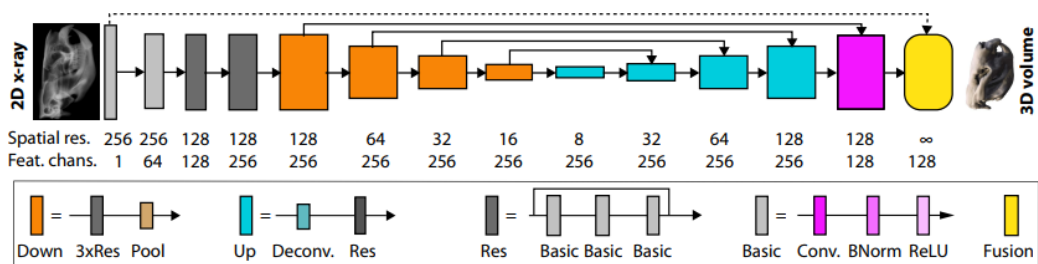
- 1) Featurized image pyramid (รูปที่ 20-a) แนวคิดนี้ใช้วิธีลดขนาดของภาพตั้งต้น (input image) ให้มีขนาดเล็กลงเป็นสัดส่วน 1, 2, 4, 8, ... เท่าจากภาพตั้งต้น และใช้ Convolutional operation ที่ใช้ hand-engineered-filter ขนาดคงที่ในการสกัดคุณสมบัติของวัตถุออกจากภาพแต่ละขนาด เพื่อให้คุณสมบัติที่สกัดออกมาสามารถใช้แยกแยะหรือจดจำวัตถุที่มีขนาดเล็ก/ใหญ่แตกต่างกันได้ หรือ เรียกว่า Scale-invariance นิยมใช้ในงาน classification แต่ข้อเสียคือทำงานช้าเนื่องจากต้องประมวลผลรูปภาพที่ถูกลดสัดส่วนลงจำนวนหลายรูป และยังมีความแม่นยำในการจดจำวัตถุยังไม่ดีพอใช้ระดับที่จะนำไปใช้งานได้จริง
- 2) Single feature image (รูปที่ 20-b) แนวคิดเป็นพื้นที่ของ CNN ในปัจจุบัน จะใช้ Convolutional operation ที่มี filter ที่สามารถเรียนรู้ค่าน้ำหนักภายในได้ (learnable weight of filter) มาต่อกันเป็นระดับชั้น Convolutional operation พวกนี้จะทำการสกัดคุณสมบัติออกจากภาพ และที่มีการ down-sampling คุณสมบัติเพื่อให้มีขนาดเล็กลงครึ่งหนึ่งจำนวนหลายๆระดับความลึก นิยมใช้ในงาน Classification จากคุณสมบัติในระดับที่อยู่ลึกที่สุดเท่านั้น ตัวอย่างเช่น VGG16 [25] ข้อดีคือมีความแม่นยำในการจดจำวัตถุสูงขึ้นตามระดับความลึกในการประมวลผลของโครงสร้าง แต่จะมีข้อเสียคือหากใช้โครงสร้างที่มีความลึกมากๆ จะทำให้เกิด Overfitting ของโมเดล และมีโอกาสเกิด gradient vanishing สูง อีกทั้งการจดจำวัตถุขนาดเล็กทำได้ยากเพราะเทคนิคนี้ใช้คุณสมบัติระดับที่ลึกที่สุดในการทำ Classification เท่านั้น
- 3) Pyramid feature hierachy (รูปที่ 20-c) โครงสร้างแบบนี้มีลักษณะคล้ายกับ Single feature image ในข้อที่แล้วมาก เพียงแต่จะใช้คุณสมบัติที่ถูกสกัดได้จาก Convolutional operation ทุกระดับความลึกมารวมกันเพื่อแยกแยะหรือจดจำวัตถุที่มีขนาดเล็ก/ใหญ่ต่างกัน นิยมใช้ในงาน Detection ตัวอย่างเช่น Single Shot Detector (SSD) [26] ข้อดีคือโครงสร้างแบบนี้มีคุณสมบัติ Scale-invariance สามารถจดจำวัตถุที่มีขนาดแตกต่างกันได้ และมีความแม่นยำสูง ข้อเสียคือหากปรับโครงสร้างให้มีความลึกมากๆ ยังคงประสบปัญหา Overfitting แล Gradient vanishing ในระหว่างฝึกสอนได้
- 4) Feature pyramid network (รูปที่ 20-d) โครงสร้างนี้ปรับปรุงมาจาก Pyramid feature hierachy อีกชั้น โดยจะมีการใช้ down-sampling ในช่วงครึ่งแรกและ up-sampling ในช่วงครึ่งหลังของ Convolutional operation มีลักษณะเป็นตัวเข้ารหัส-ถอดรหัส

(Encoder-Decoder) และเพิ่มการเชื่อมต่อระหว่างคุณสมบัติในช่วงแรกๆ(ระดับขั้นต้นๆ) และช่วงครึ่งหลัง(ระดับขั้นลึกๆ) เข้าด้วยกัน (Concatenation) เพื่อให้ในระหว่างการแพร่กระจายย้อนกลับ (back propagation) ค่า gradient สามารถวิ่งผ่านเชื่อมต่อของคุณสมบัติต้น-ลึกได้อย่างง่าย สามารถลดโอกาสเกิด Overfitting และ Gradient vanishing ได้อย่างมีประสิทธิภาพ นิยมใช้ในงาน Segmentation ตัวอย่างเช่น UNet [27]

## 2.4.2 งานวิจัยที่เกี่ยวข้อง (Related work)

### 1) Single-image Tomography: 3D Volumes from 2D Cranial X-Rays

Single-image Tomography [28] เป็นงานวิจัยชิ้นแรกๆที่นำเสนอเกี่ยวกับการสร้างรูปทรงสามมิติจากภาพ X-ray โดยใช้โครงข่ายแบบ CNN ที่มีโครงสร้างแบบ Feature Pyramid Network (FPN) มีลักษณะเป็นตัวเข้ารหัสและถอดรหัสดังรูปที่ 21 ตัวเข้ารหัสจะรับภาพ X-ray ขนาด 256x256 เข้ามาประมวลผลด้วย DenseNet (สีเทา) และมีการลดขนาดหรือ down-sampling ด้วย Pooling (สีส้ม) จะได้คุณสมบัติที่ถูกเข้ารหัส (Encoded features) ตัวถอดรหัส (สีฟ้า) รับ encoded features มาเรียนรู้เพื่อถอดรหัส โดยจะเพิ่มขนาดหรือ up-sampling ให้มีขนาดใหญ่ขึ้น ในระดับขั้นจะมีการเชื่อมต่อ encoded features กับคุณสมบัติที่ถูกถอดรหัสในขั้นก่อนที่ถูกเพิ่มขนาดแล้ว ในขั้นตอนสุดท้ายของการทำ Conv2d จะกำหนดขนาด channel สุดท้ายให้มีขนาด 128 ซึ่งเท่ากับความกว้างและสูง 128x128 หรือเป็นปริมาตรขนาด 128x128x128 (สีม่วง) และฟิวชั่น (Fusion) (สีเหลือง) นำคุณสมบัติที่ได้จากขั้นตอนที่แล้วมารวมกันให้ได้ปริมาตรของรูปทรงที่ต้องการ



**Figure 4:** Input to our architecture is a 2D x-ray gray image (left). The network converts this image into an internal representation with decreased spatial resolution (here seen as a block's height) and increasing depth (depicted as a block's width). Each type of block (encoded as colors) is defined as a combination of other blocks. Solid lines are learned, dotted lines are non-learned. For details, please see the text.

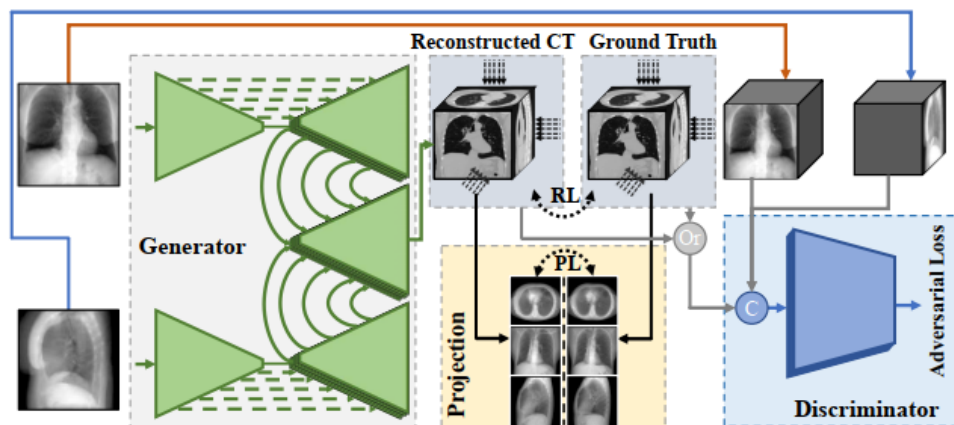
### รูปที่ 21 โครงข่าย Single-image Tomography [28]

ในการฝึกสอนโมเดลนี้จะใช้ density error ซึ่งดัดแปลงมาจากกฎของ Beer-lambert (สมการที่ 1) เป็นฟังก์ชันสูญเสีย และทำการฝึกสอนด้วย Mammalia UTCT database ที่เป็นรูปทรงสามมิติของกะโหลกสัตว์เลี้ยงลูกด้วยนมที่ได้จาก 3D scan เป็นข้อมูลแบบวีอกเซล

(Voxel) ใช้เป็นข้อมูลขาออกที่แท้จริงของการฝึกสอน และใช้สำหรับสังเคราะห์ภาพ X-ray แบบดิจิทัล (Digitally reconstructed radiograph, DRR) เนื่องจากไม่มีภาพ X-ray ที่มีมุมมองและตำแหน่งที่ตรงกันกับรูปทรงสามมิติพอดี ภาพ DRR นี้จะถูกใช้เป็นข้อมูลขาเข้าของการฝึกสอน

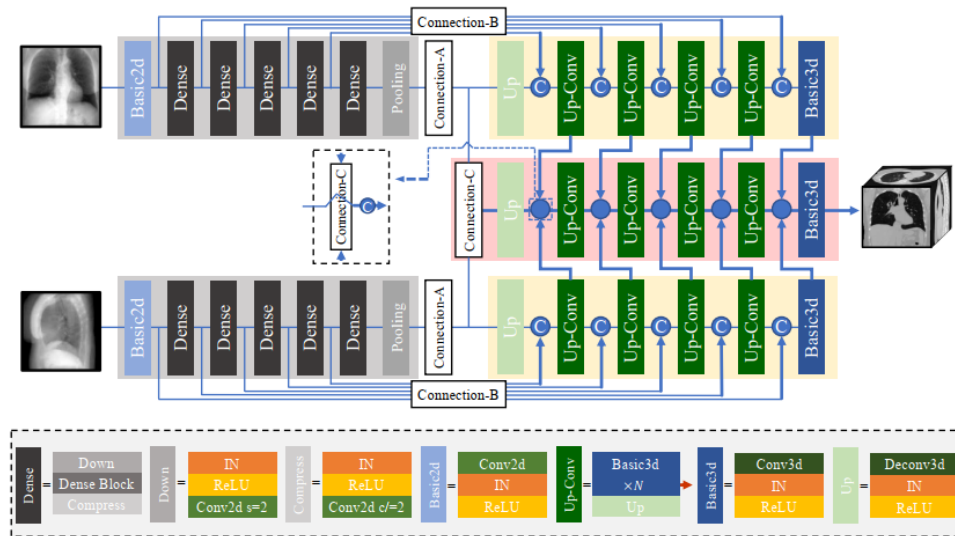
ข้อดีของโมเดลนี้ คือ สามารถสร้างรูปทรงกะโหลกสัตว์ได้อย่างแม่นยำโดยใช้ภาพ X-ray สังเคราะห์เพียงภาพเดียว ให้อารยะเอียดที่จำเพาะของตัวอย่างนั้นๆ ได้ดี ซึ่งต่างจากการสร้างรูปทรงสามมิติด้วยวิธีแบบจำลองรูปร่างเชิงสถิติ แต่ก็มีข้อเสียหรือข้อสังเกต คือ ในความเป็นจริงอวัยวะของคนหรือสัตว์ไม่สามารถถูกจัดวางอยู่ในสมมาตรก่อนทำการ CT-scan เพื่อหารูปทรงสามมิติ ดังนั้นหากรูปทรงสามมิติที่ถูกลำเอียงมาใช้ฝึกสอนหรือทดสอบนั้น ไม่ได้ถูกจัดวางอยู่ในตำแหน่งที่ได้สมมาตร หรือ ไม่ได้อยู่ตรงกลาง Voxel พอดี หรือ วางเลื้อนไปในทางลึกของ Voxel (ซึ่งไม่สามารถสังเกตเห็นได้จากภาพ DRR ที่เป็นข้อมูลขาเข้า) อาจทำให้รูปทรงสามมิติที่โมเดลสร้างขึ้นนั้นมีการเลื่อนตำแหน่งหรือไม่ซ้อนทับกับตำแหน่งของรูปทรงที่แท้จริงได้ เป็นผลให้ความแม่นยำลดลงอย่างมาก

## 2) X2CT-GAN: Reconstructing CT from Biplanar X-Rays with Generative Adversarial Networks



รูปที่ 22 ภาพรวมของโมเดล X2CT-GAN [29]





รูปที่ 23 Generator ของ X2CT-GAN [29]

โมเดล X2CT-GAN [29] ถูกพัฒนาขึ้นเพื่อแก้ปัญหาการเลือนของตำแหน่งตามแนวลึกที่เกิดขึ้นในโมเดล Single-Image Tomography [28] โดยการใช้ข้อมูลขาเข้าเป็นภาพ DRR สองมุมมอง ได้แก่ Anterior-Posterior และ Lateral view ซึ่งสังเคราะห์มาจากภาพ CT บริเวณปอด จาก The lung image database consortium (LIDC) and image database resource initiative (IDRI) จุดสำคัญของโมเดลนี้คือการใช้การฝึกสอนแบบขัดแย้ง (Generative and Adversarial) ระหว่าง Generator และ Discriminator ช่วยให้ผลลัพธ์มีความสมจริงมากขึ้น แต่ละส่วนประกอบมีรายละเอียดต่อไปนี้

- Generator ประกอบไปด้วย 3 ส่วนคือ Encoder, Decoder และ Fusion ตัว Encoder (Conv2D) สกัดและเข้ารหัสคุณสมบัติจากภาพ DRR ให้ตัว Decoder (Conv3D) ถอดรหัส โดย Encoder และ Decoder มีรูปแบบการเชื่อมต่อของโครงสร้างกันแบบ FPN และจะมีสองชุดทำงานขนานกันสำหรับประมวลผลภาพ DRR แต่ละมุมมอง จากนั้นส่งคุณสมบัติที่ถอดรหัสไปที่ Fusion เพื่อรวมคุณสมบัติและสร้างรูปทรงสามมิติด้วย Conv3D และ ConvTranpose3D

- Discriminator เป็น Classification network ทำหน้าที่แยกว่ารูปทรงสามมิติที่รับเข้ามา นั้นเป็นรูปทรงสามมิติของจริงหรือถูกสังเคราะห์ขึ้นมา ใช้ Conv3D และ Fully-connected layer ในการแยกประเภท

การฝึกสอน X2CT-GAN จะใช้ Mean square loss, Projection loss เพื่อคำนวณความสูญเสียของ Generator ในการสร้างรูปทรงสามมิติ และ Adversarial loss ใช้เพื่อหาความผิดพลาดจากการทำนายรูปทรงสามมิติ เริ่มทำการฝึกสอน Generator ให้สร้างรูปทรงก่อน

แล้วค่อยส่งรูปทรงสามมิติที่ Generator สร้างขึ้น หรือ Ground Truth ให้กับ Discriminator ทำนายว่าอันไหนเป็นของจริงหรือไม่ จนกระทั่ง Discriminator ไม่สามารถแยกรูปทรงสามมิติระหว่างกันได้

ข้อดีของโมเดล X2CT-GAN คือสามารถสร้าง CT-image สมจริงจากการฝึกสอนแบบ Adversarial attack ซึ่งโมเดลอื่นไม่สามารถทำได้ ข้อเสียคือการฝึกสอนแบบ Adversarial attack นั้นมีการลู่อู่มากๆ

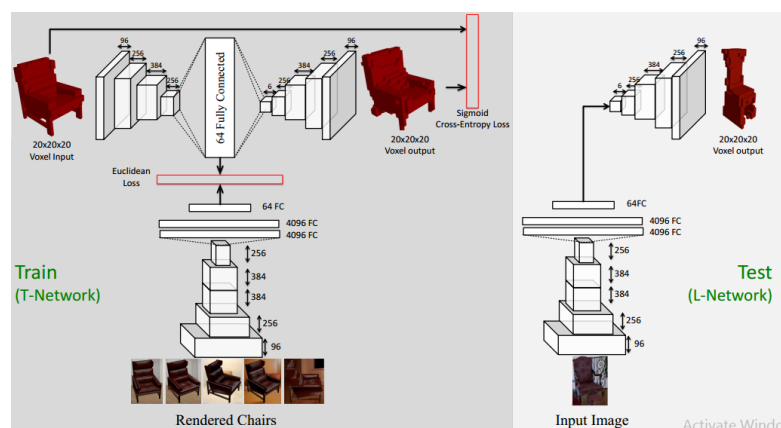
### 3) Learning a Predictable and Generative Vector Representation for Objects

ในงานวิจัยนี้ เป็นการสร้าง representation ของโมเดลสามมิติของวัตถุให้อยู่ในรูปแบบของ vector ขนาด 64x1 เรียกโมเดลนี้ว่า TL-Network [30] มีส่วนประกอบหลักสองส่วนคือ

- Autoencoder part ทำหน้าที่เข้ารหัสโมเดลสามมิติของวัตถุให้อยู่ในรูปของ vector ขนาด 64x1 มี Encoder และ Decoder ทำหน้าที่เข้ารหัสและถอดรหัส โดยใช้ Conv3D สกัดคุณสมบัติจากโมเดลสามมิติของวัตถุ จากนั้นทำแบนเป็น vector (Flatten) ให้อยู่ในรูปของ 64-vector ซึ่งเป็น Object Representation และถอดรหัส vector ด้วย Conv3D และ ConvTranspose3D ให้กลับมาเป็นโมเดลสามมิติขนาดเท่าเดิมอีกครั้งนี้

- Predictor part รับภาพสองมิติของวัตถุนั้นๆมาเพื่อใช้ทำนายโมเดลสามมิติ โดยประมวลผลด้วย Conv2D และ Pooling ในมีขนาดเล็กลง และใช้ Fully-connected layer เพื่อทำนาย 64-vector ให้เหมือนกับ 64-vector จาก Autoencoder ซึ่งเป็น Object representation ที่ต้องการ

ในการอนุมานผลลัพธ์ (Inference) ทำการตัด Encoder ออก เหลือเพียงแค่ Predictor (ทำนาย 64-vector) และ Decoder (สร้างโมเดลสามมิติจาก 64-vector ที่ Predictor ทำนายไว้)

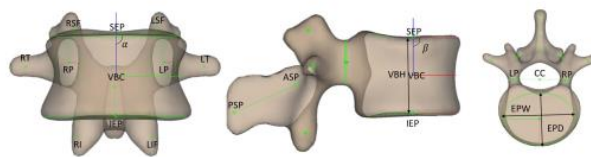


รูปที่ 24 TL-Network architecture [30]

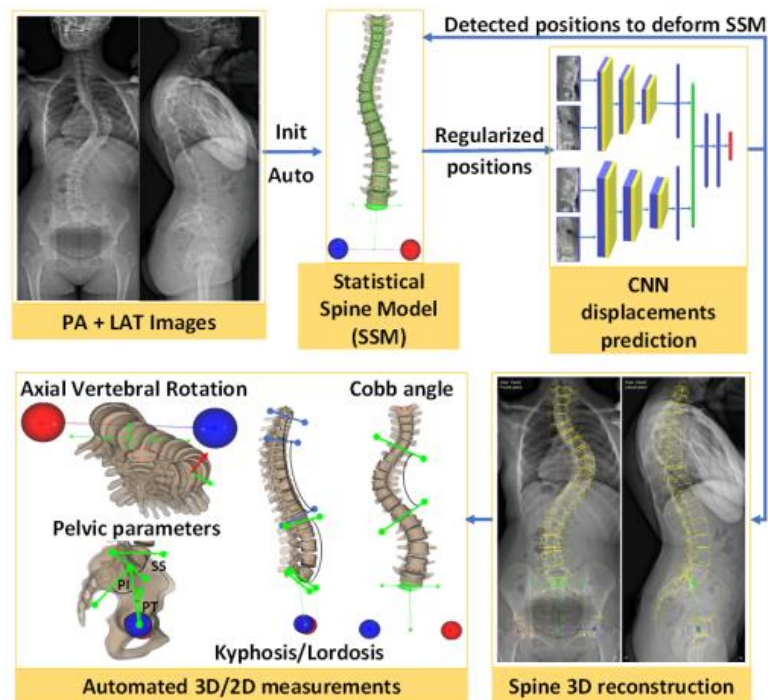
ข้อดีของ TL-Network คือ ให้โมเดลสามมิติที่มีความแม่นยำสูงจากการประมวลข้อมูลเข้า  
เข้าเพียงภาพเดียว และ Vector representation ของวัตถุ สามารถนำไปใช้วิเคราะห์รูปทรง  
ทางคณิตศาสตร์ (Arithmetic geometry) ต่อได้ง่าย ข้อเสียคือโครงสร้างนี้ใช้ Fully-  
connected layer ซึ่งใช้จำนวนพารามิเตอร์ค่อนข้างมากตามขนาดของ Vector  
representation ที่เพิ่มขึ้น หากต้องการโมเดลสามมิติที่มีความละเอียดสูงก็ความจำเป็นที่  
จะต้องใช้ Vector Representation ที่มีขนาดมากๆ จำนวนพารามิเตอร์และการอนุมาน  
ผลลัพธ์ก็จะใช้เวลามากตามไปด้วย ถึงแม้ว่าโครงสร้าง TL-network ให้ความแม่นยำสูงจาก  
ข้อมูลป้อนเข้าเพียงภาพเดียว แต่ก็เหมาะกับงานสร้างรูปทรงสามมิติที่มีขนาดเล็กๆ เท่านั้น

#### 4) 3D Spine Reconstruction from Biplanar Radiographs Using CNN for Statistical Spine Model Fitting

ในงานวิจัยนี้เป็นการสร้างรูปทรงสามมิติจากการทำงานร่วมกันของโมเดล Machine  
learning กับ Statistical shape ของอวัยวะที่สนใจ (ในงานวิจัยนี้คือ statistical spline  
model (รูปที่ 25)) โดยโมเดล Machine learning ใช้ตรวจจับตำแหน่งสำคัญทางกายวิภาค  
(anatomical landmark) เพื่อใช้สำหรับการซ้อนทับ (registration) Statistical spline model  
เข้ากับภาพ X-ray สองมุมมอง (AP-view และ Lateral-view) เพื่อหารูปทรงสามมิติของ  
spline ที่มีความแม่นยำสูง [31] ดังรูปที่ 26 โดยไม่มีการสังเคราะห์ภาพ DRR ของ Statistical  
spline model เพื่อนำไปหา similarity กับภาพ X-ray แต่ลดละมุมมอง จึงช่วยลดเวลาที่เสียไป  
กับการทำซ้ำจากวิธีการสร้างรูปทรงสามมิติแบบ Statistic shape model เดิม ที่ได้อธิบายไป  
แล้วในหัวข้อ 2.3

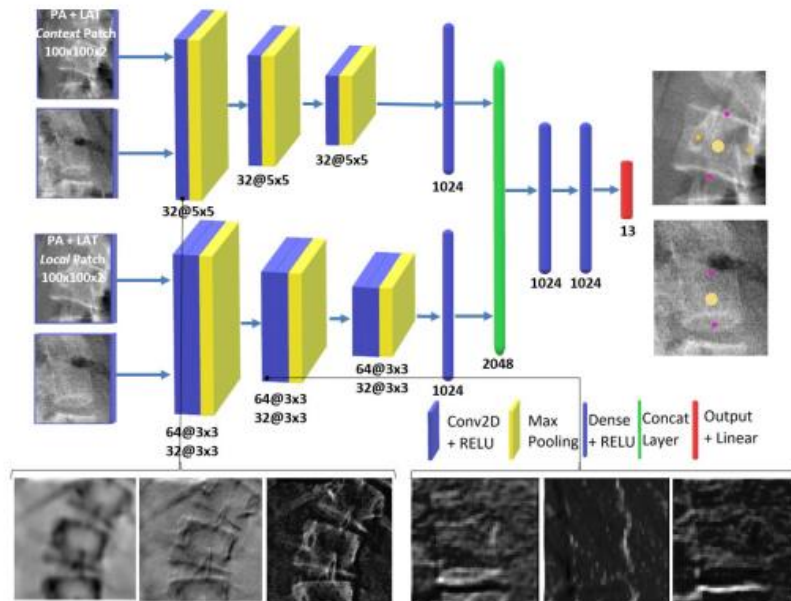


รูปที่ 25 Statistical spline model [31]

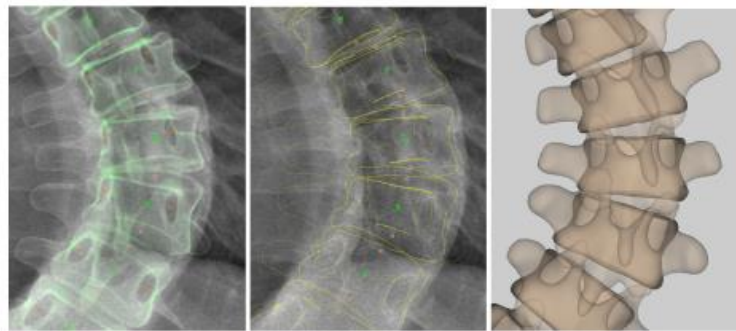


รูปที่ 26 ภาพรวมโมเดลการสร้างรูปทรงสามมิติแบบ CNN SSM fitting [31]

โมเดล Machine learning ในงานวิจัยนี้ จะใช้โครงข่าย CNN สองชุด แต่ละชุดจะรับภาพ X-ray สองมุมมอง CNN ชุดบนจะใช้สกัดคุณสมบัติบริบท (context feature) จากภาพ X-ray โดยใช้วิธีการ down-sampling 1/4 ส่วน CNN ชุดล่างจะใช้สกัดคุณสมบัติเฉพาะที่ (local feature) ด้วยวิธีการ down-sampling 1/3 คุณสมบัติจาก CNN แต่ละชุดจะถูกนำมาต่อกันให้อยู่ในรูปของ vector และใช้ ANN ในการทำ regression เพื่อทำนายตำแหน่งของ landmark ต่างๆ ของ statistical spline model ในสามมิติเทียบกับตำแหน่งของ pixel ในภาพ X-ray สองมุมมองที่รับเข้ามา landmark ต่างๆ ที่ได้จาก CNN นี้จะถูกใช้ในการเปลี่ยนแปลงรูปทรงของ statistical spline model โดยตรงโดยไม่ต้องมีการสังเคราะห์ภาพ DRR ของรูปทรงสามมิติขึ้นมาเพื่อเปรียบเทียบกับภาพ X-ray ที่ป้อนเข้าไป



รูปที่ 27 โครงข่าย CNN สำหรับ spline landmark detection [31]



รูปที่ 28 ผลลัพธ์จากการสร้างรูปทรงสามมิติของ spline [31]

ข้อดีของการสร้างรูปทรงสามมิติในงานวิจัยนี้คือ สามารถสร้างรูปทรงสามมิติที่มีความแม่นยำสูงมาก มีความรวดเร็วกว่าการสร้างรูปทรงสามมิติโดยใช้การทำซ้ำอย่างมาก และสามารถสร้างรูปทรงสามมิติได้หลายชิ้นพร้อมๆ กันแสดงดังรูปที่ 28 ข้อเสียคือ เนื่องจาก Statistical model ยังไม่สามารถใช้สร้างเมช (mesh) ของกระดูกที่แตกหักได้ การสร้างรูปทรงสามมิติของกระดูกที่แตกหักจึงเป็นไปได้ไม่สำหรับวิธีนี้

### บทที่ 3

#### การออกแบบโมเดลสร้างรูปทรงสามมิติเบื้องต้น (Preliminary Design 3D Reconstruction Model)

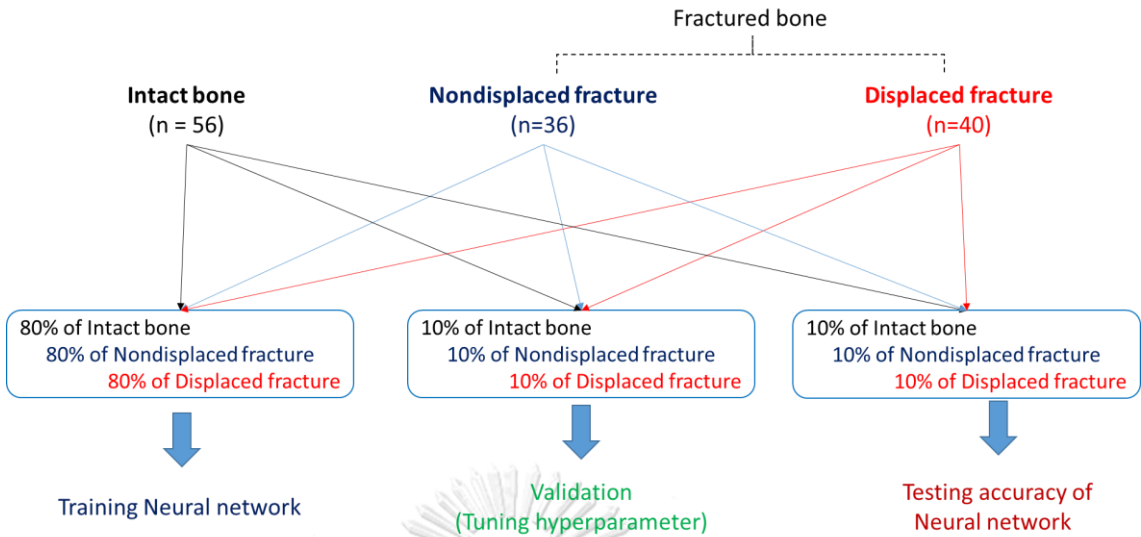
จากการศึกษาความรู้พื้นฐานและงานวิจัยที่เกี่ยวข้อง พบว่าโมเดลการเรียนรู้ของเครื่องแบบ CNN [22] โดยใช้โครงสร้างแบบ Feature-Pyramid Network (FPN) [24] มีความเป็นไปได้สูงในการนำมาประยุกต์ใช้สร้างรูปทรงสามมิติจากภาพสองมิติ เช่น โมเดล X2CT-GAN [29] ในบทนี้จะการสร้างโมเดลเบื้องต้นเพื่อสร้างรูปทรงกระดูกต้นขาที่แตกหักจากภาพถ่ายรังสีสองมิติสองมุมมอง โดยใช้โครงข่ายแบบ CNN โครงสร้างแบบ FPN และทำการฝึกสอนระบบเพื่อนำผลลัพธ์ที่ได้มาวิเคราะห์และคิดค้นเทคนิคใหม่ๆ เพื่อโมเดลสามารถสร้างรูปทรงสามมิติของกระดูกที่แตกหักได้อย่างมีประสิทธิภาพและมีความแม่นยำต่อไป โดยขั้นตอนการดำเนินงานจะประกอบได้ด้วย การเก็บข้อมูล การจัดเตรียมข้อมูลสำหรับการสอน การออกแบบโมเดลการสร้างรูปทรงสามมิติ และผลลัพธ์เบื้องต้น มีรายละเอียดดังต่อไปนี้

#### 3.1 การเก็บข้อมูล (Data collection)

ในงานนี้จะเก็บข้อมูลแบบย้อนหลังที่มีอยู่แล้วในฐานข้อมูลของโรงพยาบาลจุฬาลงกรณ์ จะเก็บภาพถ่ายรังสีเอกซเรย์คอมพิวเตอร์ (CT image) ของผู้ป่วยกระดูกต้นขาหรือกระดูกต้นขาหักข้างซ้ายหรือขวาก็ได้ โดยจะมีกระดูกปกติในข้างที่ไม่ได้รับบาดเจ็บด้วย ก่อนการเก็บข้อมูลผู้วิจัยได้ทำการยื่นเรื่องขออนุญาตเก็บข้อมูลและขอจริยธรรมงานวิจัยกับทางโรงพยาบาลจุฬาลงกรณ์เรียบร้อยแล้ว (IRB No.249/64) จากข้อมูลภาพ CT ที่ได้มาทั้งหมดทั้งกระดูกต้นขาแบบปกติและแบบหักทั้งหมดมีจำนวน 132 ตัวอย่าง เราจะแบ่งตัวอย่างออกเป็น 3 ชนิดตามลักษณะของกระดูก คือ 1) กระดูกต้นขาแบบปกติ(Intact femur) 2) กระดูกต้นขาที่หักแบบไม่มีการเคลื่อนที่(Nondisplaced fracture femur) และ 3) กระดูกต้นขาที่หักแบบมีการเคลื่อนที่(Displaced fracture femur) (ตามรูปที่ 29) ตามหลักการเรียนรู้ของเครื่อง ดังนี้

- 1) 80% ของตัวอย่างสอนทั้งหมด ใช้สำหรับเป็นข้อมูลสอนระบบ (Training sample)
- 2) 10% ของตัวอย่างสอนทั้งหมด ใช้สำหรับปรับแต่งตัวแปรของระบบ (Validation sample)
- 3) 10% ของตัวอย่างสอนทั้งหมด ใช้สำหรับทดสอบความแม่นยำของระบบ (Test sample)

ทั้งนี้สัดส่วนของกลุ่มตัวอย่างทั้งสามอาจมีการเปลี่ยนแปลงได้ เพื่อให้เหมาะสมกับลักษณะงาน

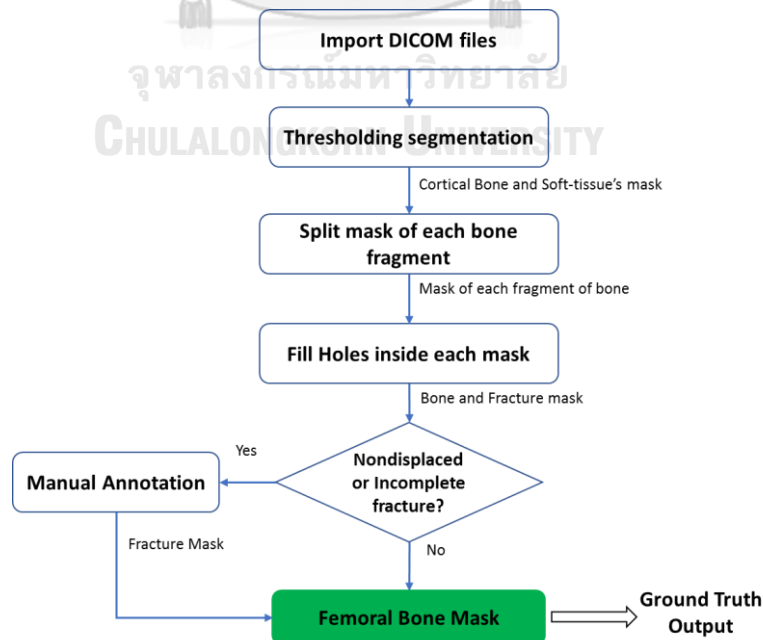


รูปที่ 29 การแบ่งกลุ่มของกลุ่มตัวอย่างสำหรับการเรียนรู้ของเครื่อง

### 3.2 การจัดเตรียมข้อมูลสำหรับการสอน (Data Preprocessing)

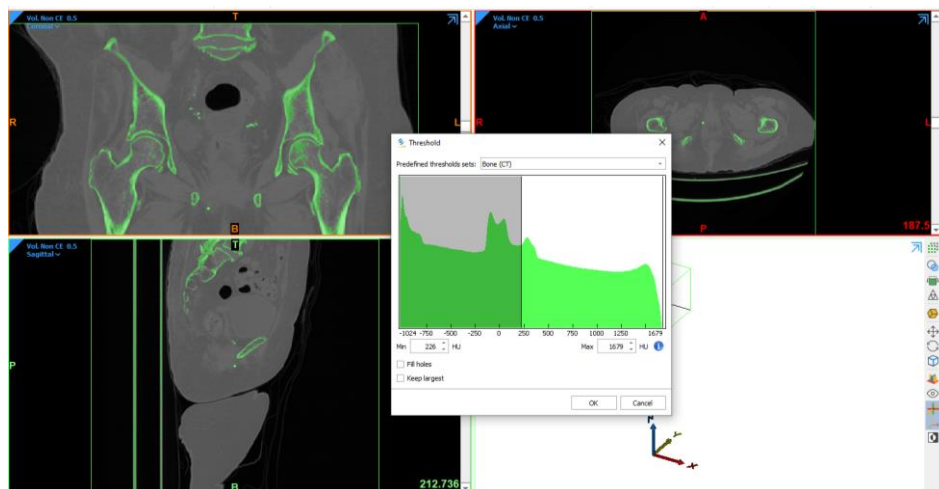
#### 3.2.1 การแบ่งส่วนกระดูกต้นขา (Fractural Femur Segmentation)

ในหัวข้อนี้จะเป็นการแบ่งส่วนปริมาตร (Volume segmentation) ของปริมาตรกระดูกต้นขาที่เราสนใจออกจากภาพ CT ที่ได้เก็บข้อมูลมา ในขั้นตอนนี้เราจะใช้โปรแกรมสำหรับประมวลผลภาพถ่ายรังสีโดยเฉพาะชื่อ Mimics 21.0 (Materialise Innovation Suite) ในการแบ่งส่วนปริมาตรกระดูกส่วนที่เราสนใจ โดยมีขั้นตอนแสดงดังรูปที่ และมีรายละเอียดดังต่อไปนี้



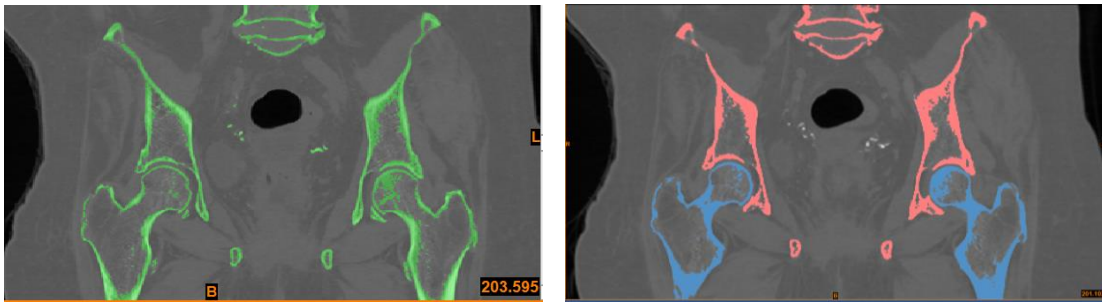
รูปที่ 30 ขั้นตอนการแบ่งส่วนปริมาตรของกระดูกต้นขาที่แตกหัก

- 1) เริ่มต้นเปิดโปรแกรม MIMIC แล้ว import ชุดไฟล์ CT image นามสกุล DICOM ของตัวอย่างที่ต้องการ จากการนั้น ทำการแบ่งส่วนกระดูกทึบ (cortical bone) ด้วยระดับความเข้มของพิกเซล (Thresholding segmentation) โดยเลือกค่าความเข้มในหน่วย Hounsfield ช่วงประมาณ 175-226 จนถึงค่าความเข้มสูงที่สุดของตัวอย่างนั้นๆ จะได้ Mask ของกระดูกทึบออกมา แสดงดังรูปที่ 31
- 2) จากนั้นทำการแยกส่วนกระดูกแต่ละชิ้นออกจากกัน ด้วยคำสั่ง Split Mask โดยระบาย Seed mask ของกระดูกแต่ละชิ้น โปรแกรมจะทำการแบ่งส่วนกระดูกแต่ละชิ้นออกจากกันเป็น mask ใหม่ เป็นกระดูกต้นขาชิ้นที่ 1, ชิ้นที่ 2, ..., และกระดูกเชิงกราน แสดงดังรูปที่ 32
- 3) หลังจากแบ่งส่วนกระดูกแต่ละชิ้นเรียบร้อยแล้ว นำ Mask ของกระดูกแต่ละชิ้นมาเติมเต็มช่องว่างภายใน (Filling boundary) โดยใช้คำสั่ง Smart Fill ทำการเลือกปรับค่า global hole closing distance ที่เหมาะสม จะได้ Mask ที่ถูกเติมเต็มช่องว่างภายในเรียบร้อยแล้ว แสดงดังรูปที่ 33
- 4) ในกรณีที่มีการหักของกระดูกตัวขาแบบไม่เคลื่อนที่หรือไม่สมบูรณ์ (Nondisplaced or Incomplete fracture) ให้ทำการสร้าง Mask ใหม่ และทำเครื่องหมาย (Manual annotation) อาณาบริเวณที่เป็นรอยแตกหักของกระดูก (ตำแหน่งใดๆ ในปริมาณของกระดูกที่มีค่าระดับความเข้ม Hounsfield ที่ต่ำกว่าค่าความเข้มของกระดูก)
- 5) เมื่อทำตามขั้นตอนที่ 1) ถึง 4) จะได้ปริมาตรสามมิติ หรือ Mask จากการแบ่งส่วนของกระดูกต้นขา แสดงดังรูปที่ 34 ซึ่ง Mask นี้จะใช้เป็นข้อมูลขาออกที่แท้จริง (Ground truth output) ของโมเดลสร้างภาพสามมิติ ซึ่งจะถูกอธิบายอีกครั้งในหัวข้อที่ 3.3



รูปที่ 31 การแบ่งส่วนกระดูกด้วยวิธีระดับความเข้ม (Thresholding segmentation)

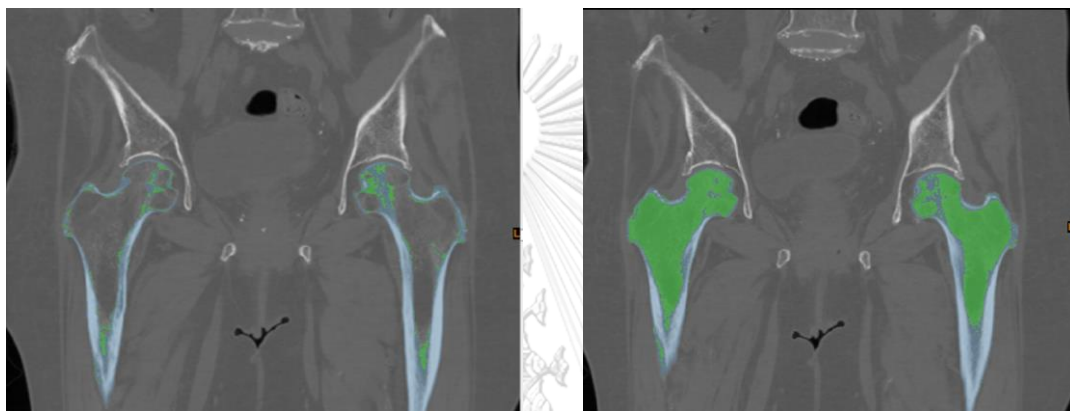




(ก) ก่อนทำการแยกกระดูก

(ข) หลังทำการแยกกระดูก

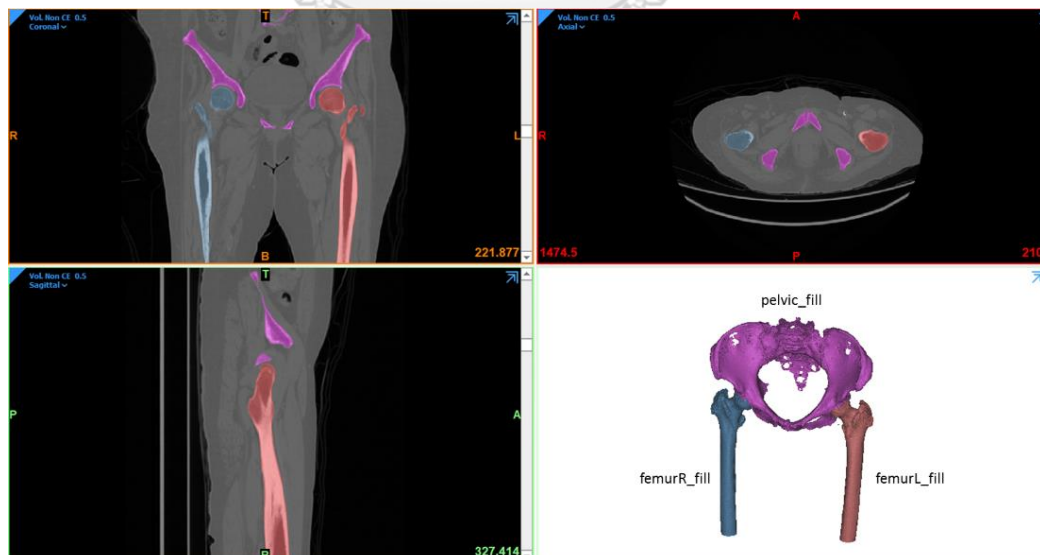
รูปที่ 32 การแยกชิ้นส่วนกระดูกแต่ละชิ้น



(ก) ก่อนเติมเต็มช่องว่างกระดูก

(ข) หลังเติมเต็มช่องว่างกระดูก

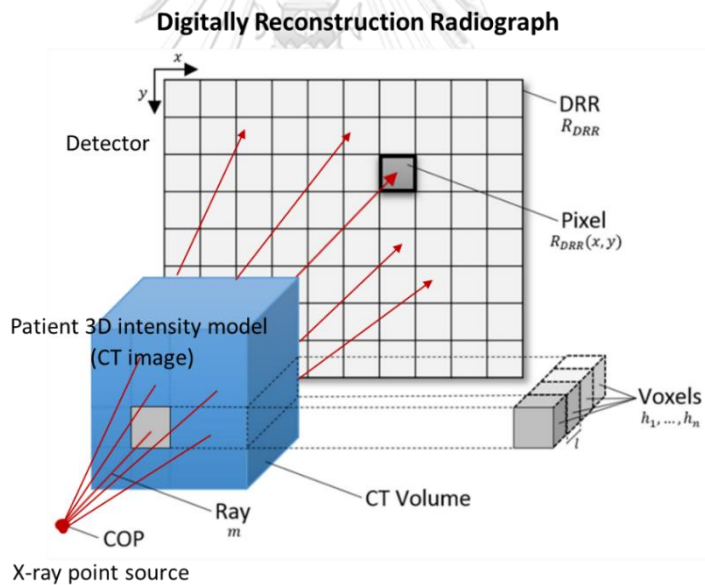
รูปที่ 33 การเติมเต็มช่องว่างในกระดูก



รูปที่ 34 กระดูกต้นขาที่ถูกแบ่งส่วนเรียบร้อยแล้ว (Complete Femur Mask)

### 3.2.2 การสร้างภาพ Scout film

งานวิจัยนี้สนใจที่จะนำภาพ X-ray มาใช้ในการสร้างโมเดลสร้างมิติของกระดูกต้นขา แต่โดยเทคนิคการสร้างโมเดลสามมิติด้วยวิธี machine learning จำเป็นที่จะต้องใช้ภาพ X-ray (ข้อมูลขาเข้า) ที่มีมุมมองตรงกับโมเดลสามมิติที่ได้จากการ segmentation ของภาพ CT (ข้อมูลขาออก) จากการสืบค้นข้อมูลทางจากโรงพยาบาลจุฬาลงกรณ์มหาวิทยาลัยพบว่าไม่สามารถหาข้อมูลภาพ X-ray ที่มีมุมมองตรงกับภาพ CT ได้เลย จะมีเพียงแต่ภาพ CT เท่านั้น และการจะเก็บข้อมูลภาพ X-ray พร้อมกับภาพ CT ของผู้ป่วยกระดูกโพงหักใหม่ให้มีเพียงพอสำหรับการสอนของเครื่องนั้นจะต้องใช้เวลาานาน 5-10 ปี และมีค่าใช้จ่ายที่สูงมาก ดังนั้นในงานวิจัยนี้จึงมีความจำเป็นที่จะต้องสังเคราะห์ภาพ X-ray จากภาพ CT ที่ได้เก็บข้อมูลมา ซึ่งเรียกว่าภาพ Scout film หรือภาพถ่ายรังสีสังเคราะห์แบบดิจิทัล (Digitally reconstructed radiograph, DRR) [20] โดยจะใช้หลักการเกิดภาพ X-ray ตามกฎของ Beer-Lambert [16] ที่ได้กล่าวไปแล้วในหัวข้อ 2.2 รูปที่ 35 แสดงแผนภาพการสังเคราะห์ Scout film ซึ่งมีรายละเอียดดังต่อไปนี้



รูปที่ 35 แผนภาพการสังเคราะห์ Scout film [20]

- 1) เริ่มต้นจากภาพ CT จากไฟล์ DICOM ต้นฉบับ ( $V(x,y,z)$ ) โดยปกติแล้วไฟล์ DICOM จะมีความเข้มที่แตกต่างกันไปขึ้นอยู่กับเครื่อง CT scan ที่ใช้เก็บภาพ ไฟล์ DICOM จะเก็บค่าคุณสมบัติ (DICOM Attribute) ของความชัน (Rescale Slope) และค่าจุดตัดแกน (Rescale Interception) แสดงดังตารางที่ 3 เราจะสามารถใช้สองค่าคุณสมบัตินี้เพื่อ

แปลงค่าความเข้มของภาพ CT ต้นฉบับ ให้กลายเป็น Voxel ในหน่วย Hounsfield ( $HU_{(x,y,z)}$ ) ได้จากสมการที่ 2

ตารางที่ 3 DICOM Attribute ของค่า Rescale Intercept และ Slope

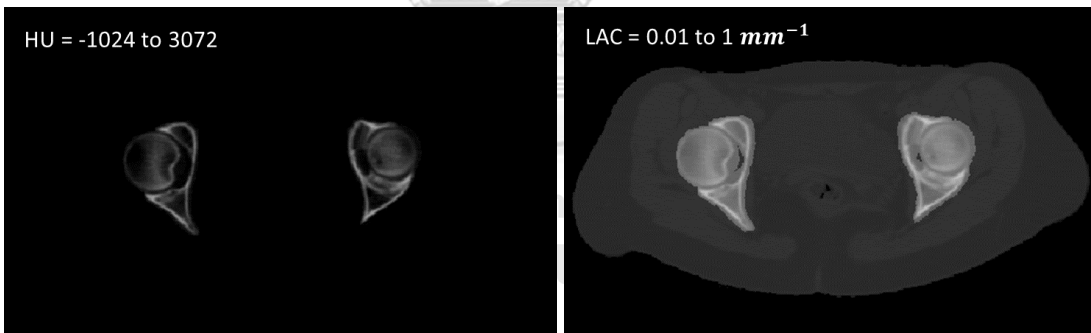
Attribute Name	Tag	Type	Attribute Description
>Rescale Intercept	(0028,1052)	1	The value b in relationship between stored values (SV) and the output units. Output units = m*SV + b.
>Rescale Slope	(0028,1053)	1	m in the equation specified by Rescale Intercept (0028,1052).

$$HU_{(x,y,z)} = RescaleSlope \times V_{(x,y,z)} + RescaleIntercept \quad \text{สมการที่ 2}$$

- 2) เราจะทำการแปลงภาพ CT ที่มีค่าความเข้มในหน่วย Hounsfield  $HU_{(x,y,z)}$  ให้เป็นภาพ CT ที่มีค่าเป็นสัมประสิทธิ์การลดทอนเชิงเส้น  $\mu_{(x,y,z)}$  หน่วย  $mm^{-1}$  จากความสัมพันธ์สมการที่ 3 หรือสมการที่ 4 ผลลัพธ์จากการคำนวณปริมาตรในหน่วยของ LAC แสดงดังรูปที่ 36

$$HU_{(x,y,z)} = \frac{\mu_{(x,y,z)} - \mu_w}{\mu_w} \times 1000 \quad \text{สมการที่ 3}$$

$$\mu_{(x,y,z)} = \frac{(HU_{(x,y,z)})(\mu_w)}{1000} + \mu_w \quad \text{สมการที่ 4}$$



(ก) ภาพ CT ในหน่วย Hounsfield

(ข) ภาพ CT ในหน่วยของ LAC

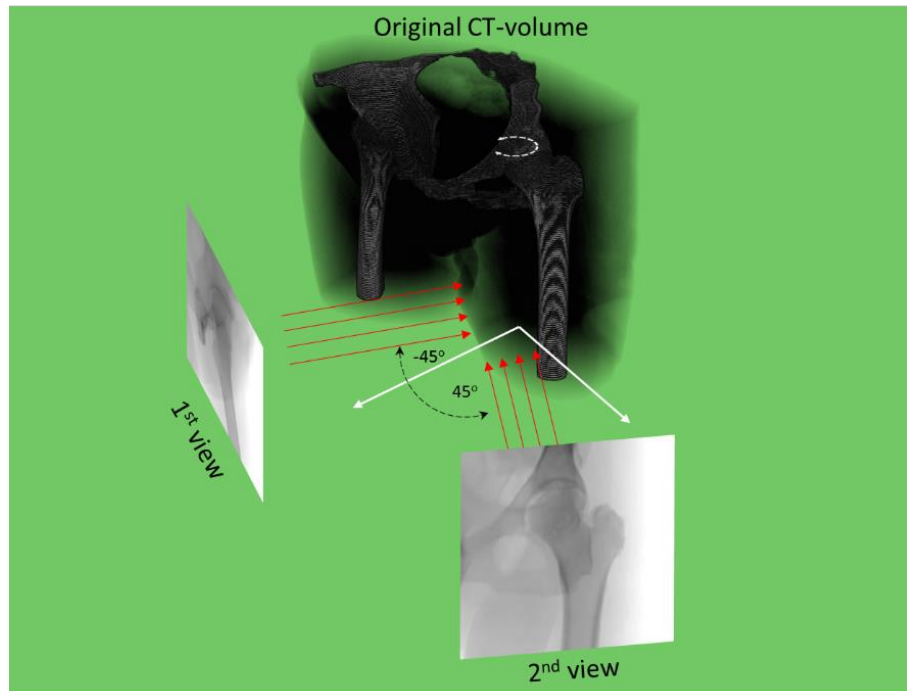
รูปที่ 36 ผลลัพธ์จากการคำนวณ LAC

- 3) สังกะสี Scout film หรือภาพรังสีแบบดิจิทัลด้วยเทคนิค Ray casting โดยดัดแปลงกฎของ Beer Lambert [16] ตามสมการที่ 1 ให้อยู่ในรูปของตัวแปรไม่ต่อเนื่อง (Discrete Variables) ตามสมการที่ 5 โดยใช้ภาพ CT ที่มีค่าเป็นสัมประสิทธิ์การลดทอนเชิงเส้น  $\mu_{(x,y,z)}$  และ  $l_{n,m}$  คือความละเอียดของ voxel เรากำหนดให้  $I_0$  มีค่าเริ่มต้นเท่ากับ 1 เพื่อให้  $I_{(x,z)}$  มีค่าอยู่ในช่วง 0-1 ผลลัพธ์ที่ได้แสดงดังรูปที่ 37 Scout film นี้จะสามารถ

สังเคราะห์จากทิศทางใดๆ ของภาพ CT ก็ได้ และถูกใช้เป็นข้อมูลขาเข้าของโมเดลสร้างภาพสามมิติ ซึ่งจะถูกอธิบายอีกครั้งในหัวข้อที่ 3.3

$$I_{(x,z)} = I_0 \exp \left( \sum_{j=1}^N \mu_{(x,j,z)} l_{n,m} \right)$$

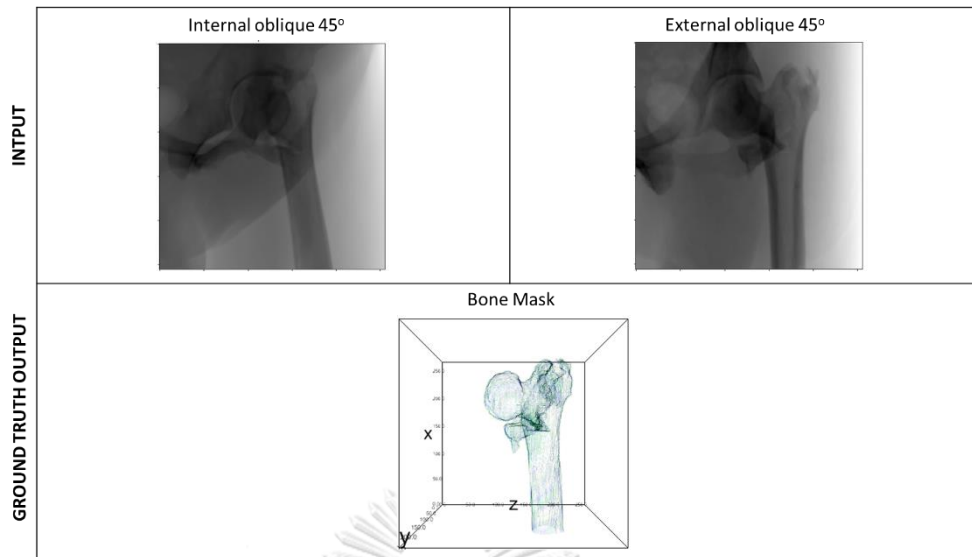
สมการที่ 5



รูปที่ 37 ผลลัพธ์การสังเคราะห์ Scout film หรือ ภาพรังสีแบบดิจิทัล

ข้อดีของการใช้ Scout film คือ เราสามารถสังเคราะห์แบบ X-ray จากทิศทางไหนของภาพ CT ก็ได้ และได้ภาพ X-ray ที่มีตำแหน่งของกระดูกหรืออวัยวะเป้าหมายที่สอดคล้องกับภาพ CT ซึ่งง่ายต่อการนำไปใช้สำหรับการประมวลผลภาพหรือคอมพิวเตอร์วิทัศน์ด้วยวิธีการเรียนรู้ของเครื่อง

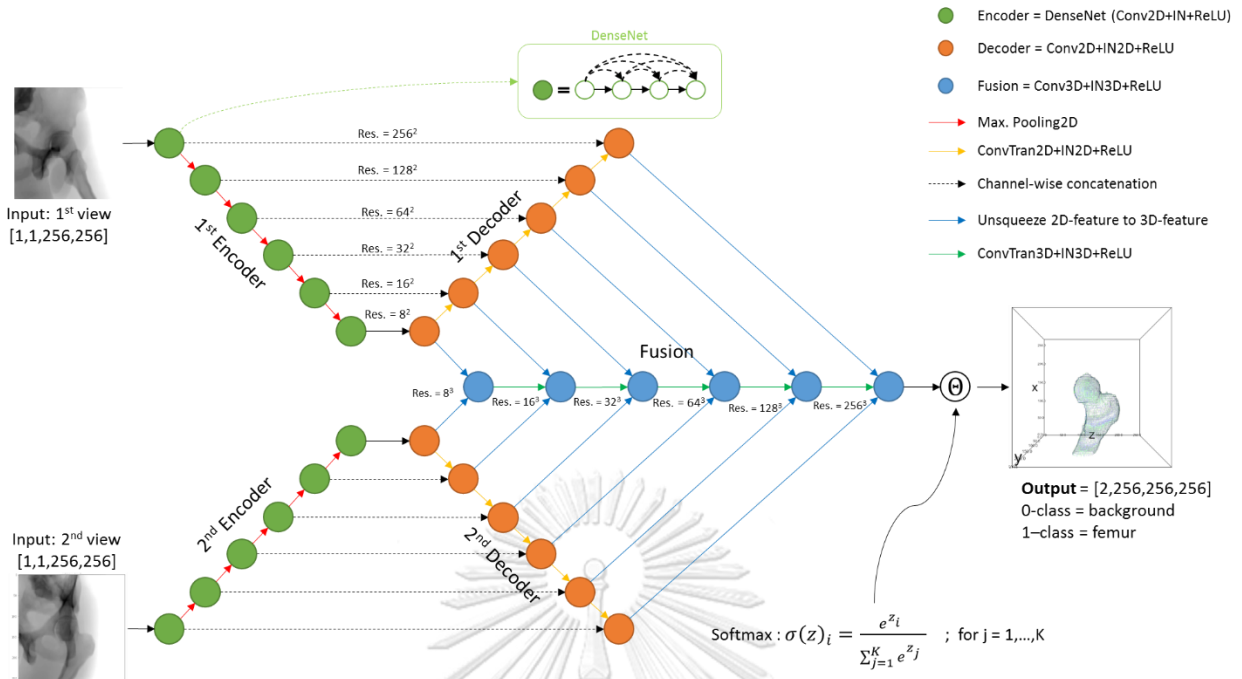
สรุปข้อมูลสำหรับฝึกสอนโมเดลสร้างรูปทรงสามมิติจะประกอบไปด้วย Scout film ที่มีมุมมองแบบ internal/external oblique view 45 องศา จำนวนทั้งหมด 2 ภาพ (ทั้งสองภาพทำมุม 90 องศา) และคลาสของรูปทรงสามมิติของพื้นหลังและกระดูกต้นขาตั้งรูปที่ 38



รูปที่ 38 ตัวอย่างสำหรับฝึกสอนโมเดลสร้างรูปทรงสามมิติของกระดูกต้นขา

### 3.3 โมเดลสร้างรูปทรงสามมิติ (3D Reconstruction Network)

โมเดลสร้างรูปทรงสามมิติจะถูกพัฒนาบน Pytorch framework ด้วยภาษา Python ภายในโมเดลจะใช้ตัวแปรแบบเทนเซอร์ (Tensor) เพื่อจัดเก็บค่าและประมวลผล Input และ Output ของตัวอย่างสอน ซึ่งอยู่ในรูปของ 4D-Tensor คือ  $[N,C,H,W]$  และ 5D-Tensor คือ  $[N,C,D,H,W]$  (โดยที่ N คือจำนวนชุดตัวอย่าง, C คือจำนวนคุณสมบัติที่อยู่ในปริภูมิสองมิติหรือสามมิติ (Channel) ส่วน D คือมิติความลึก, H คือมิติความสูง และ W คือมิติความกว้างของปริภูมิสองหรือสามมิติ) โมเดลการเรียนรู้ของเครื่องที่เราเลือกใช้จะเป็น Convolutional neural network (CNN) เพราะเหมาะสมกับการเรียนรู้บนโดเมนของภาพ ใช้จำนวนตัวแปรที่เรียนรู้ได้ (Learnable parameter) น้อยกว่า Perceptron neural network (NN) และมีคุณสมบัติการเรียนรู้โดยไม่ขึ้นกับตำแหน่งของวัตถุหรือสิ่งที่สนใจ (translational invariance) ในการแก้โจทย์การสร้างโมเดลสามมิติ เราจะใช้ Input เป็น Scout film ซึ่งสังเคราะห์จากสองมุมมองได้แก่ มุม  $-45$  และ  $+45$  องศา รอบแกนแนวตั้งของภาพ CT บริเวณกระดูกต้นขา (Internal/external oblique view 45 degree) ดังรูปที่ 37 และทำการปรับขนาดของ Input และ Ground Truth Output ให้มีขนาด  $256 \times 256$  (Tensor ขนาด  $[1,1,256,256]$ ) และ  $256 \times 256 \times 256$  (Tensor ขนาด  $[1,1,256,256,256]$ ) ตามลำดับ



รูปที่ 39 แผนภาพโมเดลสร้างรูปทรงสามมิติของกระดูกต้นขา (3DReconNet)

โมเดลสร้างรูปทรงสามมิติที่น่าเสนอจะประกอบไปด้วยตัวประมวลผล 3 ส่วนหลัก คือ ตัวเข้ารหัส (Encoder), ตัวถอดรหัส (Decoder) และตัวผสม (Fusion) แสดงดังรูปที่ 39 เราจะขอตั้งชื่อโมเดลนี้ว่า 3DReconNet ในแต่ละส่วนของตัวประมวลผลจะมีการประมวลผลแบบ down-sampling หรือ up-sampling เป็นระยะๆ เพื่อลดหรือเพิ่มความละเอียดของคุณสมบัติ (features) เป็นจำนวนหนึ่งเท่าของความละเอียดเดิม ช่วยให้ประหยัดพื้นที่ใน RAM โดยในแต่ละชั้นที่มีความละเอียดของ features เท่าๆ กัน จะเรียกว่าเป็น ระดับชั้น (level,  $i$ ) ในโมเดลจะมีทั้งหมด  $L = 7$  ระดับชั้น ( $i \in [0, L - 1]$ ) อธิบายดังต่อไปนี้

1) ตัวเข้ารหัส (Encoder)

เริ่มต้นรับภาพ Scout film หรือข้อมูลขาเข้า  $x_v$  แต่ละมุมมอง  $v$  โดยที่  $v \in [1, 2]$  คือมุมมองที่ 1 หรือ 2 ของภาพ Scout film สำหรับในระดับชั้นแรก  $i = 0$  จะเริ่มต้นประมวลผลการด้วย Conv2D{filter=3x3, stride=1, padding=1} ให้ได้ features จำนวน 16 channel (Tensor ขนาด  $[N, 16, 256, 256]$ ) ตามสมการที่ 6 จะได้ข้อมูลขาเข้าของ Encoder ในระดับชั้นที่ 0 คือ  $ei_v^{(0)}$

$$ei_v^{(0)} = \text{Conv}_{3 \times 3} 2d(x_v) \tag{สมการที่ 6}$$

ในทุกๆ ระดับชั้น  $0 \leq i \leq L - 1$  ของ Encoder จะประมวลผลด้วยโครงข่ายคอนโวลูชันแบบ DenseNet เพื่อให้การฝึกสอนมีประสิทธิภาพและช่วยลดการจางหายไปของ

ค่าอนุพันธ์ของความสูญเสียนั้น (gradient vanishing) [32] โดยใช้หลักการประมวลผลเป็นระดับชั้นๆ (Layer) แต่ละชั้นประกอบด้วย InstanceNorm2d (IN) + ReLU+ Conv2d{filter=3x3, stride=1, padding=1} จากนั้นนำ features จากชั้นที่แล้วมาทำการเชื่อมต่อกันในมิติของจำนวนคุณสมบัติ (channel-wise concatenation) ด้วยอัตราการขยาย (Growth rate) ของ features ที่  $k = 4 - 16$  ทำซ้ำจำนวน  $N = 4$  ครั้ง ( $l \in [0, N - 1]$ ) ได้จะข้อมูลที่ถูกรหัส (encoded features)  $e_v^{(i)}$  แสดงดังสมการที่ 7

$$e_v^{(i)} = DenseNet(e_i^{(i)}) \quad \text{สมการที่ 7}$$

โดยที่ฟังก์ชัน *DenseNet* คำนวณได้จากสมการที่ 8 โดยที่  $e_i^{(i,j)}$  คือคุณสมบัติขาเข้าของ *DenseNet* ระดับชั้นที่  $i$  และระดับชั้นที่  $j$  โดยที่สัญลักษณ์  $[ ]_c$  แสดงถึง channel-wise concatenation และ  $e_i^{(i,0)} = e_i^{(i)}$  ที่ระดับชั้น  $j = 0$

$$DenseNet(e_i^{(i,j)}) = \begin{cases} \left[ IN \left( ReLU \left( Conv2d_{3 \times 3} \left( e_i^{(i,0)} \right) \right) \right), e_i^{(i,0)} \right]_c, & \text{เมื่อ } j = 0 \\ \left[ IN \left( ReLU \left( Conv2d_{3 \times 3} \left( DenseNet \left( e_i^{(k,l)} \right) \right) \right) \right), DenseNet \left( e_i^{(k,0)} \right) \right]_c, & \text{เมื่อ } j \in [0,3] \end{cases} \quad \text{สมการที่ 8}$$

ในขั้นตอนสุดท้ายก่อนจะประมวลผลในระดับชั้นถัดไป จะประมวลผลด้วย Max.Pooling{filter=2x2, stride=2} เพื่อลดขนาดของ encoded features  $e_v^{(i)}$  ลงครึ่งหนึ่งเพื่อประหยัด RAM ตามสมการที่ 9 จะได้ข้อมูลขาเข้าของ Encoder ในระดับชั้นถัดไป  $e_i^{(i+1)}$  โดยจะทำการประมวลผลด้วย Max.Pooling ที่ระดับชั้น  $0 \leq i < L - 1$  เท่านั้น

$$e_i^{(i+1)} = MaxPool2d(e_v^{(i)}) \quad \text{สมการที่ 9}$$

## 2) ตัวถอดรหัส (Decoder)

รับ encoded features ที่ระดับชั้นต่างๆ จาก Encoders มาถอดรหัส โดยข้อมูลขาเข้าของ Decoder ในระดับชั้นที่  $i = L - 1$  หรือ  $d_i^{(L-1)}$  จะเป็น  $e_v^{(L-1)}$  ส่วนในระดับชั้นที่  $0 \leq i < L - 1$  หรือ  $d_i^{(i)}$  จะเป็น encoded features ที่ระดับชั้นที่  $i + 1$  (decoded features)  $d_v^{(i+1)}$  มาทำการ up-sampling ให้มีความละเอียดที่สูงขึ้นหนึ่งเท่าจากความละเอียดเดิม ด้วย ConvolutionTranspose2d {filter=2x2, stride=2, padding=0} + InstanceNorm2d + ReLU และทำ Channel-wise concatenation กับ

encoded features หรือ  $e_v^{(i)}$  ที่ในระดับชั้นมีความละเอียดเท่ากัน เรียกว่าเป็นการเชื่อมต่อด้านข้าง (Lateral connection) ทำให้เกิดเป็นโครงสร้างแบบ Feature Pyramid Network (FPN) เพื่อให้การฝึกสอนมีประสิทธิภาพและช่วยลดการจางหายไปของค่าอนุพันธ์ของความสูญเสีย (gradient vanishing) ผลลัพธ์ที่ได้จะเป็นข้อมูลขาเข้าของ Decoder หรือ  $di_v^{(i)}$  เมื่อ  $0 \leq i < L - 1$  ตามสมการที่ 10

$$di_v^{(i)} = \left( e_v^{(i)}, \text{ReLU} \left( \text{IN} \left( \text{ConvTrans}_{2 \times 2} 2dU \left( d_v^{(i+1)} \right) \right) \right) \right)_C \quad \text{สมการที่ 10}$$

จากนั้นทำการถอดรหัสโดยการประมวลผลด้วย Conv2d{filter=3x3, stride=1, padding=1} + InstanceNorm2d + ReLU ในทุกๆ ระดับชั้น จะได้คุณสมบัติที่ถูกถอดรหัส (decoded features) หรือ  $d_v^{(i)}$  ตามสมการที่ 11

$$d_v^{(i)} = \text{ReLU} \left( \text{IN} \left( \text{Conv2d} \left( di_v^{(i)} \right) \right) \right) \quad \text{สมการที่ 11}$$

การประมวลผลด้วย Conv2d ทั้งหมดของ Decoder จะถูกกำหนดขนาด channel ขาออกให้มีค่าเท่ากับความละเอียดของ features (H หรือ W) ในระดับชั้นนั้นๆ เพื่อเราจะได้แปลงมิติ channel (C) ของ Decoded features  $d_v^{(i)}$  ให้กลายเป็นมิติความลึก (D) แสดงถึงปริมาตรในปริภูมิสามมิติหรือ 5D-Tensor ( $D_v^{(i)}$ ) แสดงดังสมการที่ 12

$$[N, C, H, W] \leftrightarrow [N, 1, D, H, W] \quad \text{สมการที่ 12}$$

### 3) ตัวผสม (Fusion)

ตัวผสมทำหน้าที่รวม Decoded features หรือ  $d_v^{(i)}$  แต่ละมุมมอง  $v$  เข้าด้วยกัน เพื่อสร้างปริมาตรสามมิติที่แสดงถึงรูปทรงของกระดูกต้นขาที่เราสนใจ ข้อมูลขาเข้าของ Fusion หรือ  $X_v^{(i)}$  สำหรับในระดับชั้น  $0 \leq i < L - 1$  จะเป็นการ channel-wise concatenation กันระหว่าง Decoded features หรือ  $d_v^{(i)}$  และ fused features ของระดับชั้นที่ลึกกว่า ( $i + 1$ ) หรือ  $F_v^{(i+1)}$  ที่ถูก up-sampling แล้วของแต่ละมุมมอง  $v$  ตามสมการที่ 13 ส่วนระดับชั้นสุดท้ายที่  $i = L - 1$  ข้อมูลขาเข้าของ Fusion คือ  $X_v^{(L-1)} = d_v^{(L-1)}$

$$X_v^{(i)} = \left( d_v^{(i)}, \text{ReLU} \left( \text{IN} \left( \text{ConvTrans}_{2 \times 2} 3d \left( F_v^{(i+1)} \right) \right) \right) \right)_C \quad \text{สมการที่ 13}$$

สำหรับการคำนวณคุณสมบัติที่ถูกผสมแล้ว (Fused features) ที่ระดับชั้นใดๆ หรือ  $F_v^{(i)}$  จะเป็นการนำข้อมูลขาเข้าของแต่ละมุมมองของตัวผสม ได้แก่  $X_1^{(i)}$  และ  $X_2^{(i)}$  มา



รวมกันเพื่อสร้างโมเดลสามมิติ เนื่องจากข้อมูลขาเข้ามุมมองที่ 2 ( $X_2^{(i)}$ ) มีองศาการถ่ายภาพรังสีที่ต่างกับข้อมูลขาเข้ามุมมองที่ 1 ( $X_1^{(i)}$ ) อยู่ 90 องศารอบแนวตั้ง ดังนั้นจึงต้องทำการสลับลำดับของเทนเซอร์(Permutation)  $X_2^{(i)}$  เพื่อให้  $X_1^{(i)}$  และ  $X_2^{(i)}$  มีมุมมองที่ตรงกัน ก่อนที่จะนำ  $X_2^{(i)}$  ไปถัวเฉลี่ยกับ  $X_1^{(i)}$  ดังรูปที่ 40 โดยใช้ Conv3d{filter=3x3x3, stride=1, padding=1} + InstanceNorm3d(IN) + ReLU จะได้ fused features  $F_v^{(i)}$  ตามสมการที่ 14

$$F^{(i)} = \text{ReLU} \left( \text{IN3d} \left( \text{Conv3d}_{3 \times 3 \times 3} \left( \frac{X_1^{(i)} + T(X_2^{(i)})}{2} \right) \right) \right) \quad \text{สมการที่ 14}$$

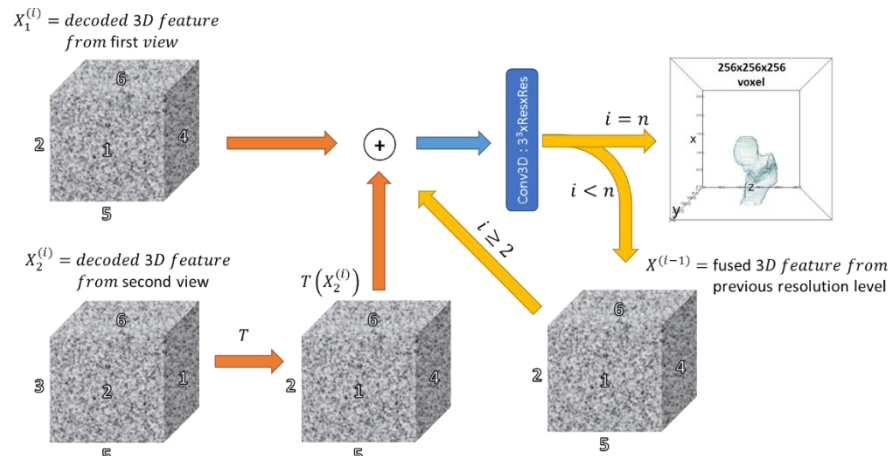
โดยที่  $T$  แสดงถึงการสลับลำดับของเทนเซอร์ เราขอตั้งชื่อวิธีการนี้ว่า Averaging Permutation

เมื่อทำการประมวลผลครบทุกระดับชั้น  $0 \leq i \leq L - 1$  แล้ว พิจารณาที่ Fused feature ระดับที่ชั้นที่  $i = 0$  หรือ  $F_v^{(i)}$  จะถูกประมวลผลอีกครั้งด้วย Conv3d{filter=1x1x1, stride=1, padding=0} + InstanceNorm3d(IN3D) + ReLU เพื่อเป็นการแยกประเภทในระดับ Voxel (voxel-wise classification) และสุดท้ายทำการคำนวณด้วยฟังก์ชัน SoftMax บนมิติของ channel  $C$  สุดท้ายจะได้เทนเซอร์ขนาด  $[N, 3, 256, 256, 256]$  ที่มีจำนวน channel เป็น 3 โดยที่ channel แรกจะเป็นที่อยู่ของพื้นหลัง, channel ที่สองคือรูปทรงของกระดูก และ channel สุดท้ายคือรูปทรงของรอยแตกของกระดูก

$$Y = \sigma \left( \text{ReLU} \left( \text{IN3d} \left( \text{Conv3d}_{1 \times 1 \times 1} (F^{(i)}) \right) \right) \right) \quad \text{สมการที่ 15}$$

โดยที่  $\sigma$  คือฟังก์ชัน SoftMax สามารถคำนวณได้จากสมการที่ 16 โดยที่  $z_i$  คือตำแหน่งใดๆ บนปริภูมิสามมิติ

$$\sigma(z)_i = \frac{e^{z_i}}{\sum_{j=0}^2 e^{z_j}} \quad \text{สมการที่ 16}$$



รูปที่ 40 การรวมคุณสมบัติ (Fusion)  $X_v^{(i)}$  ของแต่ละมุมมองเข้าด้วยกัน

### 3.4 การฝึกสอนโมเดล (Model training)

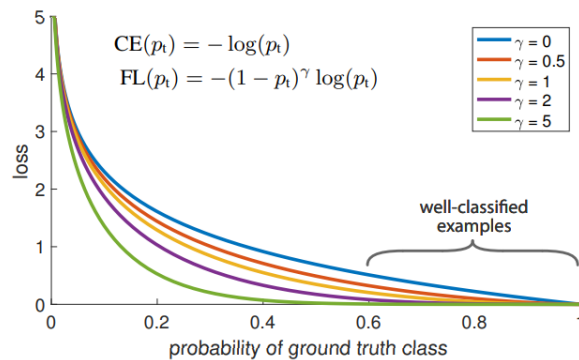
จากโมเดลที่เสนอมาในหัวข้อที่ 3.3 เราจะใช้เครื่อง Server computer ในการฝึกสอน ทดสอบ และอนุมานผลลัพธ์ ซึ่งมีคุณสมบัติดังต่อไปนี้

- Mainboard: TYAN Server computer B8021G88V2HR-2T-RM-N
- CPU : AMD EPYC™ 7002 Series Processor 16 cores
- GPU : ASUS GeForce RTX 3090 Blower 24GB  
ZOTAC GeForce RTX 2080 Ti Blower 11GB
- RAM : DDR4 ECC RDIMM/LRDIMM 2400 32GB
- ROM : 970 EVO NVMe M.2 SSD 1TB

และเราจะใช้ Focal loss เป็นฟังก์ชันความสูญเสียตามสมการที่ 17 โดยเป็นฟังก์ชันที่ปรับปรุงมาจาก Binary cross entropy loss (BCE loss) เหมาะกับตัวอย่างสอนที่มีความไม่สมดุลระหว่างคลาสสูง (Highly unbalance) อย่างเช่นตัวอย่างสอนของเรามีสัดส่วนคลาสพื้นหลังต่อคลาสกระดูกต้นขาอยู่ที่ 49 ต่อ 1 โดยจะมี focusing parameter เพิ่มเข้ามาในพจน์แรก ช่วยให้การถ่วงน้ำหนักของค่าความสูญเสียของวอกเซลที่ทำนายถูกต้อง (well-classified voxel) เข้าใกล้ศูนย์ (เมื่อเทียบ BCE loss จะยังมีค่าความสูญเสียของอยู่มาก) ส่วนวอกเซลที่ทำนายผิด (wrong-classified voxel) จะมีค่าสูงกว่า well-classified voxel อย่างมากดังรูปที่ 41 ด้วยคุณสมบัตินี้ทำให้ focal loss สามารถใช้ฝึกสอนโมเดลที่มีรูปทรงที่ซับซ้อนได้ดีกว่า BCE อย่างมีนัยสำคัญ [33]

$$FocalLoss = \sum_{t=1}^n -w_c(1 - p_t)^{\gamma} \log(p_t)$$

สมการที่ 17



รูปที่ 41 กราฟค่าความสูญเสีย Focal loss ต่อค่าความน่าจะเป็น [33]

เราจะทำการฝึกสอนโมเดลสร้างรูปทรงสามมิติที่เสนอไปในหัวข้อที่ 3.3 ด้วยขั้นตอนตาม Pseudo Code ต่อไปนี้

- 1: Initialized learnable weight of the model
- 2: For e < epoch, do
  - For i in number of training batch, do (Training loop)
    - Randomly sampling training batch
    - Compute forward-propagation for 3D Reconstruction model
    - Calculate the Multiclass Focus Loss
    - Calculate reconstruction accuracy as IoU matrix
    - Compute back-propagation
    - Update learnable weight by Adam Optimizer [34]
  - End For i
  - For j in number of validation batch, do (Validation loop)
    - Randomly sample of validation batch
    - Compute forward-propagation using previously weight-updated model
    - Calculate reconstruction accuracy as IoU matrix
    - If the current validation accuracy is the best accuracy, do
      - save the best trained model
  - End If
  - End For j
  - adjust learning rate by scheduler considering validation accuracy
- End For e

### 3.5 การวัดความแม่นยำเชิงปริมาณ (Quantitative evaluation)

การวัดความแม่นยำของผลลัพธ์ที่ได้ในเชิงปริมาณจะทำได้โดยการนำรูปทรงกระดูกต้นขาที่ถูกสร้างขึ้น (output) มาทำการเปรียบเทียบกับรูปทรงที่แท้จริง (ground-truth) ด้วย 2 metrics ดังต่อไปนี้

- 1) Intersection over Union (IoU) เป็นหนึ่งใน overlap-based evaluation metrics จะเป็นการนำปริมาณของ output และ ground-truth มาคำนวณหาบริเวณที่ซ้อนทับกัน (Intersection) แล้วหารด้วยบริเวณรวมกัน (Union) ในหน่วยของ voxel ที่เป็นบวก ตามสมการที่ 18 ค่า IoU จะมีค่าอยู่ในช่วง 0 – 1 โดยจะมีค่าใกล้เคียงกับ 1 เมื่อปริมาณของ output และ ground-truth มีบริเวณที่ซ้อนทับกันมากๆ และมีเข้าใกล้ 0 เมื่อปริมาณทั้งสองมีบริเวณที่ซ้อนทับกันน้อยๆ

$$IOU = \frac{TP}{FP + TP + FN}$$

สมการที่ 18

- 2) Average Symmetric Surface Distance (ASSD) เป็นหนึ่งใน distance-based evaluation metric จะเป็นการหารระยะทางที่สั้นที่สุดระหว่างจุดบนพื้นผิวของรูปทรงสามมิติที่ถูกสร้างขึ้น (output,  $S_M$ ) และรูปทรงที่แท้จริง (ground truth,  $S_G$ ) โดยจะนำรูปทรงสามมิติทั้งสองมาทำการแปลงจาก Voxel ให้อยู่ในรูปของ Cloud point โดยใช้อัลกอริทึม Marching cubes [35] จะได้เมช (Mesh) ซึ่งประกอบไปด้วยจุด (vertices) และพื้นผิว (faces) จากนั้นจะทำการหารระยะกระจัด (Euclidean distance error) ระหว่าง vertices บนพื้นผิวของรูปทรงสามมิติทั้งสองด้วยระยะทางที่สั้นที่สุด (Nearest neighbours) โดยจะหารระยะทางที่สั้นที่สุดทั้งจาก ground truth ไปยัง output และจาก output ไปยัง ground truth ตามสมการ

$$ASSD = \frac{1}{|S_M| + |S_G|} \left( \sum_{x \in S_M} d(x, S_G) + \sum_{y \in S_G} d(y, S_M) \right)$$

สมการที่ 19

โดยที่  $d(x, S_G) | x \in S_M$  คือระยะเชิงกระจัดระหว่างจุดใดๆ บนพื้นผิวของ output ไปยัง ground truth และ  $d(x, S_M) | x \in S_G$  คือระยะเชิงกระจัดระหว่างจุดใดๆ บนพื้นผิวของ ground truth ไปยังพื้นผิวของ output

### 3.6 ผลลัพธ์ (Result)

ในการวัดผลลัพธ์ของโมเดลสร้างรูปทรงสามมิติ เราแบ่งการวัดผลออกเป็นสองส่วน คือ การวัดผลเชิงปริมาณและการวัดผลเชิงคุณภาพ

#### การวัดผลเชิงปริมาณ (Quantitative result evaluation)

เราทำการวัดความแม่นยำของโมเดลที่เราพัฒนาขึ้นมาด้วยการวัดค่าความแม่นยำของรูปทรงที่ถูกสร้างขึ้น (Output) เทียบกับรูปทรงที่แท้จริง (Ground Truth) ของชุดตัวอย่างทดสอบ ค่าความแม่นยำของรูปทรงสามมิติจะถูกวัดด้วย mIoU และ mASSD แสดงดังตารางที่ 4

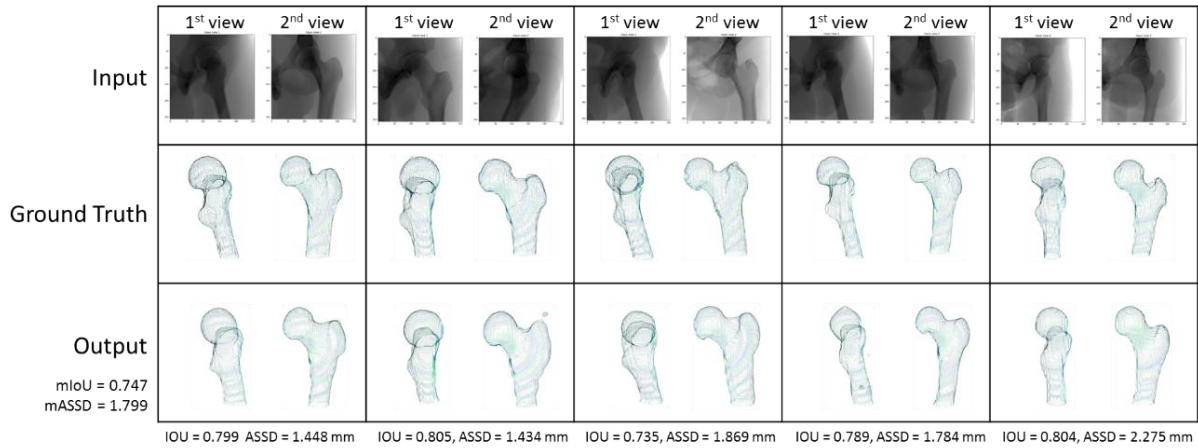
ตารางที่ 4 ความแม่นยำของโมเดลสร้างรูปทรงสามมิติของกระดูกชนิดต่างๆ

Sample type	mIoU (SD)	mASSD (SD)
Intact	0.785 (0.033)	1.799 (0.640)
Nondisplaced	0.786 (0.041)	1.907 (0.460)
Displaced	0.584 (0.101)	2.726 (0.738)
Overall	0.718 (0.116)	2.114 (0.766)

#### การวัดผลเชิงคุณภาพ (Qualitative result evaluation)

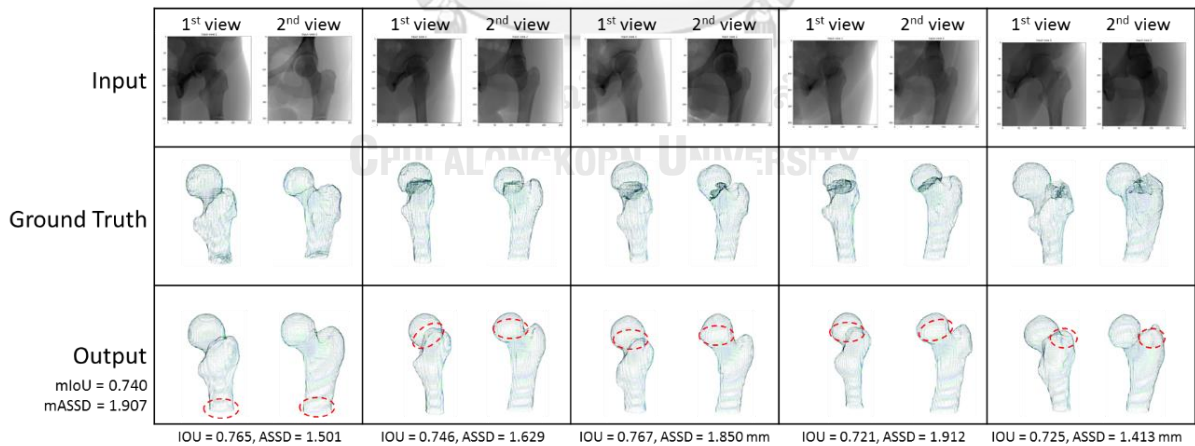
การวัดผลเชิงปริมาณจะเป็นการวิเคราะห์รายละเอียดของรูปทรงที่ถูกสร้างขึ้นว่ามีความใกล้เคียงกับรูปทรงที่แท้จริงอย่างไรจากชุดตัวอย่างทดสอบ ผลลัพธ์ที่ได้จากสร้างรูปทรงสามมิติของกระดูกต้นขาปกติ กระดูกต้นขาที่แตกหักแบบไม่เคลื่อนที่และแตกหักแบบเคลื่อนที่แสดงดังรูปที่ 42 ถึง รูปที่ 44 ตามลำดับ

สำหรับตัวอย่างทดสอบของกระดูกปกติผลลัพธ์ตามรูปที่ 42 จากการวิเคราะห์รายละเอียดของรูปทรงสามมิติที่ถูกสร้างขึ้นจะพบว่ารูปทรงกระดูกต้นขาปกตินั้นมีรายละเอียดที่ใกล้เคียงกับความเป็นจริงพอสมควรทั้งเชิงตำแหน่งและรูปทรงของกระดูก ซึ่งยังสามารถปรับปรุงให้มีความแม่นยำได้มากกว่านี้



รูปที่ 42 ผลลัพธ์การสร้างรูปทรงสามมิติของกระดูกต้นขาปกติ































สำหรับตัวอย่างทดสอบของกระดูกหักแบบไม่เคลื่อนที่ผลลัพธ์ตามรูปที่ 43 จากการวิเคราะห์รายละเอียดของรูปทรงสามมิติที่ถูกสร้างขึ้นจะพบว่ารูปทรงภายนอกโดยรวมแล้วมีรายละเอียดของกระดูกที่ครบถ้วนใกล้เคียงกับความเป็นจริงอยู่พอสมควร เนื่องจากการกระดูกที่แตกหักแบบไม่เคลื่อนที่จะมีรูปทรงที่แตกต่างจากกระดูกปกติไม่มากนัก ทำให้การสร้างรูปทรงภายนอกของกระดูกทำได้อย่างแม่นยำใกล้เคียงกับกระดูกปกติ แต่รายละเอียดของรอยแตกหักนั้นยังไม่แสดงออกมาในผลลัพธ์ อาจจะเป็นเพราะจำนวนตัวอย่างสอนไม่เพียงพอ และรายละเอียดของรอยแตกนั้นเป็นส่วนที่เล็กมากๆ ทำให้การเรียนรู้ทำได้ยาก ซึ่งผู้วิจัยจะหาวิธีการปรับปรุงในขั้นตอนถัดไป



รูปที่ 43 ผลลัพธ์การสร้างรูปทรงสามมิติของกระดูกต้นขาหักแบบไม่เคลื่อนที่

สำหรับตัวอย่างทดสอบของกระดูกหักแบบเคลื่อนที่ได้ผลลัพธ์ตามรูปที่ 44 จากการวิเคราะห์รายละเอียดของรูปทรงสามมิติที่ถูกสร้างขึ้นจะพบว่ารูปทรงภายนอกของกระดูกนั้นมีรูปร่างที่ไม่ถูกต้องเท่าไรนัก เนื่องจากการกระดูกที่แตกหักแล้วมีการเคลื่อนที่ออกจากตำแหน่งเดิมไป

ค่อนข้างมากทำให้รูปร่างภายนอกเปลี่ยนแปลงไปจากกระดูกปกติอย่างมาก เป็นผลให้การสร้างรูปทรงสามมิติทำได้ไม่ดีเท่าที่ควร ส่วนรายละเอียดของรอยแตกหักก็ยังมีแนวโน้มของตำแหน่งที่ถูกต้องอยู่บ้างแต่ยังต้องปรับปรุงอีกมาก ปัญหาในส่วนนี้เป็นเรื่องที่ทำหายอย่างมากสำหรับงานวิจัยนี้ เราจะพัฒนาโมเดลหรือเทคนิคเสริมอื่นๆ ต่อไปเพื่อแก้ปัญหาดังกล่าว ซึ่งมีแนวทางการแก้ปัญหาในหัวข้อถัดไป

Input										
Ground Truth										
Output										
	IOU = 0.512, ASSD = 3.192 mm		IOU = 0.512, ASSD = 3.192 mm		IOU = 0.655, ASSD = 2.603		IOU = 0.677, ASSD = 2.390		IOU = 0.581, ASSD = 2.769 mm	

รูปที่ 44 ผลลัพธ์การสร้างรูปทรงสามมิติของกระดูกต้นขาหักแบบเคลื่อนที่

## บทที่ 4

### วิธีการที่นำเสนอ (Proposed Method)

จากการทบทวนวรรณกรรมงานวิจัยในบทที่ 2 และผลลัพธ์การสร้างรูปทรงสามมิติเบื้องต้นในบทที่ 3 พบว่ารูปทรงสามมิติที่ถูกสร้างขึ้นสำหรับกระดูกปกติมีความแม่นยำของรูปทรงในระดับปานกลาง แต่ในกรณีกระดูกหักทั้งแบบไม่เคลื่อนที่และเคลื่อนที่นั้นรูปทรงยังไม่ถูกต้องและรูปทรงกระดูกที่สร้างขึ้นยังไม่แสดงรอยแตกหักใดๆ เลย เราจึงสรุปได้ว่าสาเหตุการหายไปของรายละเอียดรอยแตกหักนั้นอาจเกิดจากจำนวนกลุ่มตัวอย่างกระดูกที่แตกหักอาจยังไม่เพียงพอต่อการเรียนรู้รอยแตกหัก และไม่มีการฝึกสอนด้วยข้อมูลของรอยแตกหักโดยเฉพาะจากกลุ่มตัวอย่างที่ป้อนให้กับโมเดล ดังนั้นเราจึงมีแนวคิดที่จะปรับปรุงความแม่นยำให้สูงขึ้นโดยเฉพาะรายละเอียดของรอยแตกหักดังต่อไปนี้

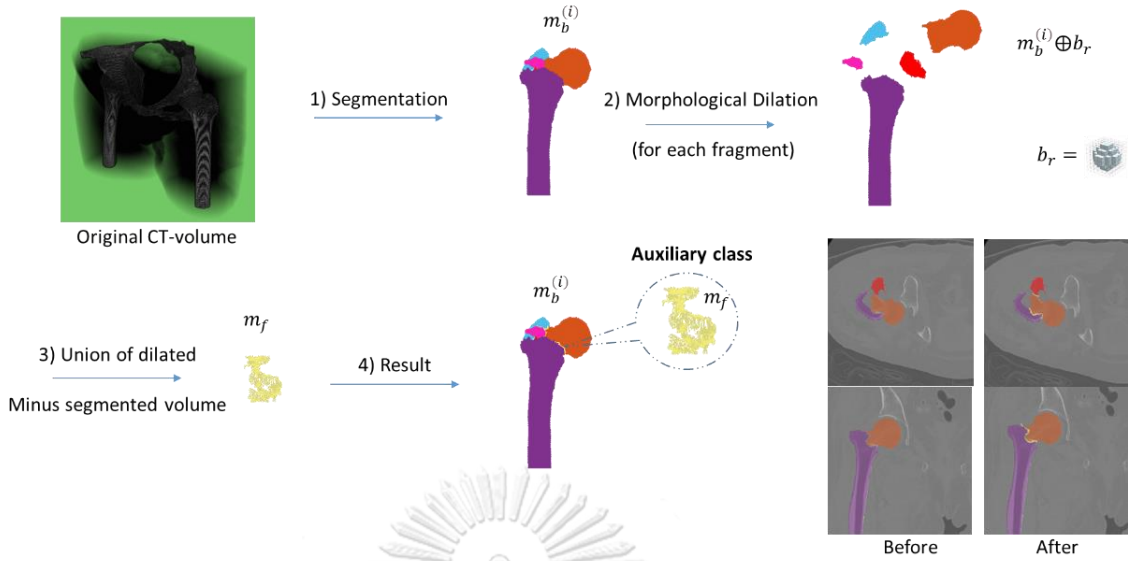
1. เทคนิคการเรียนรู้ด้วยคลาสช่วยฝึกสอน (Auxiliary class) เป็นเพื่อข้อมูลขาออก (3<sup>rd</sup> output class) ซึ่งเป็นตัวแทนของรอยแตกหักที่สังเคราะห์มาจาก bone class เพื่อให้โมเดลเรียนรู้ที่จะสร้างรอยแตกหักของกระดูกไปพร้อมๆ กับสร้างรูปทรงของกระดูกต้นขา
2. เทคนิคการเพิ่มขนาดกลุ่มตัวอย่างของกระดูกที่แตกหัก (Fractural augmentation) เพื่อเพิ่มจำนวนกลุ่มตัวอย่างที่กระดูกมีการแตกหัก ใช้สำหรับฝึกสอน

ดังนั้นในบทนี้เราจะอธิบายถึงสองเทคนิคใหม่ที่คิดค้นขึ้นมาโดยเฉพาะเพื่อแก้ปัญหาการสร้างรูปทรงสามมิติของกระดูกต้นขาที่มีการแตกหัก มีรายละเอียดดังต่อไปนี้

#### 4.1 เทคนิคการเรียนรู้ด้วยคลาสช่วยฝึกสอน (Auxiliary class)

จากการสร้างรูปทรงสามมิติด้วยโมเดลที่เรานำเสนอไปเบื้องต้น จะพบว่าในกรณีกลุ่มตัวอย่างที่กระดูกหักแบบไม่เคลื่อนที่และหักแบบเคลื่อนที่แสดงดังรูปที่ 43 และรูปที่ 44 ตามลำดับรูปร่างกระดูกที่ถูกสร้างขึ้นมาจะไม่มีรายละเอียดของรอยแตกหักปรากฏขึ้นมาเลย ซึ่งโมเดลที่เราคิดค้นขึ้นพยายามที่จะสร้างรูปทรงภายนอกเพียงอย่างเดียว โดยไม่สนใจรายละเอียดของรอยแตกหักเลย ดังนั้นเราจึงมีแนวคิดที่จะสร้างรอยแตกหัก หรือ ปริมาตรของรอยแตกหักขึ้นมา เพื่อให้เป็นอีกหนึ่งคลาส (3<sup>rd</sup> output class or Auxiliary class) ใช้สำหรับฝึกสอนโมเดลให้เรียนรู้ที่จะสร้างรอยแตกหักของกระดูกไปพร้อมๆ กับสร้างรูปทรงภายนอกของกระดูกต้นขา นิยามคืออาณาบริเวณที่อยู่ระหว่างกระดูกที่แตกหักแต่ละชิ้นที่อยู่ใกล้เคียงกัน Auxiliary class จะถูกสังเคราะห์ขึ้นมาจากรูปทรงสามมิติของกระดูกหัก (Bone class) ที่ผ่านการ segmentation จากภาพ CT เรียบร้อยแล้ว โดยมีขั้นตอนดังต่อไปนี้ (รูปที่ 45)





รูปที่ 45 ขั้นตอนการสังเคราะห์คลาสช่วยฝึกสอน (Auxiliary class) จากโมเดลกระดูกที่แตกหัก

- i. เริ่มต้นจากรูปทรงสามมิติของกระดูกที่แตกหักที่ถูก segmentation ตามขั้นตอนที่กล่าวไปในหัวข้อที่ 3.2.1 เรียบร้อยแล้วจะได้กระดูกที่แตกหักละชิ้น  $m_b^{(i)}$  (รูปที่ 45-1) โดยที่  $i \in [1, n]$  และ  $n$  เป็นจำนวนชิ้นทั้งหมดของกระดูกต้นขาที่แตกหัก
- ii. จากนั้นนำกระดูกแต่ละชิ้น  $m_b^{(i)}$  มาทำ morphological dilation ด้วย structural element แบบทรงกลมขนาด 2 หน่วย (รูปที่ 45-2) เพื่อขยายปริมาตรของกระดูกแต่ละชิ้นออกรอบๆ ตัว เพื่อให้แต่ละชิ้นมีปริมาตรส่วนที่ซ้อนทับกัน
- iii. หาปริมาตรที่ซ้อนทับกัน (Intersection) ของกระดูกที่ถูกขยายขนาดด้วย dilation แล้ว และลบออกด้วย  $m_b^{(i)}$  จะได้คลาสช่วยฝึกสอน หรือ Auxiliary class ( $m_f$ ) ตามที่ต้องการ

ขั้นตอนทั้งหมดนี้สามารถอธิบายด้วยสมการที่ 20 Auxiliary class ( $m_f$ ) จะใช้เป็นคลาสที่สามของ Ground Truth สำหรับการฝึกสอน

$$m_f = \sum_{i=1}^{n-1} (m_b^{(i)} \oplus b_r) \cap (m_b^{(i+1)} \oplus b_r) - \sum_{i=1}^n m_b^{(i)} \quad \text{สมการที่ 20}$$

โดยที่  $\oplus$  คือ morphological dilation และ  $b_r$  คือ Structural element แบบทรงกลม 2 หน่วย

หลังจากการสังเคราะห์ Auxiliary class แล้ว โมเดลสร้างรูปทรงสามมิติในหัวข้อที่ 3.3 (รูปที่ 39) จะถูกดัดแปลงเพื่อให้สามารถเรียนรู้ 3<sup>rd</sup> class เพียงเล็กน้อยเท่านั้น โดยการเพิ่มจำนวน channel

ของข้อมูลขาออก(Output) จาก 2-class เป็น 3-class ประกอบไปด้วย 1<sup>st</sup> Background class, 2<sup>nd</sup> Bone class และ 3<sup>rd</sup> Auxiliary class

#### 4.2 เทคนิคการเพิ่มขนาดกลุ่มตัวอย่าง (Data Augmentation)

โดยปกติแล้ว เทคนิคการเพิ่มขนาดกลุ่มตัวอย่างจะเป็นการนำชุดข้อมูลที่มีอยู่แล้วมาทำการประมวลผลภาพทางสัณฐานวิทยา(Morphological Image Processing) ตัวอย่างเช่น การเลื่อนพิกเซล (Pixel shifting), การหมุน (Affine transform), ตัดกรอบ (Cropping), การกลับด้าน (Flipping) และเพิ่ม/ลดขนาด (Scaling) เป็นต้น เพื่อให้มีความหลากหลายและมีลักษณะเฉพาะ (Unique) จะได้ชุดข้อมูลฝึกสอนที่มีขนาดใหญ่ขึ้น อีกทั้งการทำ Data augmentation ยังช่วยลดการเกิด overfitting หรือ high variance ของโมเดล หรือทำให้โมเดลมีความเป็น Generalize มากยิ่งขึ้น [36]

แต่เนื่องจากตัวอย่างที่เราใช้ฝึกสอนมีลักษณะเฉพาะ คือ รูปทรงสามมิติของกระดูกขา ประกอบกับรอยแตกหัก ทำให้การ augmentation ด้วยวิธีการต่างๆไปอย่างที่ได้ยกตัวอย่างไป ยังไม่สามารถเพิ่มขนาดสำหรับชุดตัวอย่างที่เฉพาะเจาะจงเช่นนี้ได้ เพราะไม่สามารถสร้างความหลากหลายของรอยแตกหักที่เกิดขึ้นได้ ดังนั้นเราจึงมีแนวคิดที่จะพัฒนาการเพิ่มขนาดตัวอย่างฝึกสอนที่มีลักษณะการแตกหัก (Fracture augmentation) ขึ้น เพื่อเพิ่มความหลากหลายสัณฐานวิทยาในเชิงการแตกหักของกระดูกต้นขาให้ครอบคลุมกับกลุ่มประชากรมากขึ้น โดยนำตัวอย่างที่มีกระดูกต้นขาปกติมาทำให้แตกหักด้วยวิธีการสุ่ม (Randomization)

ในการเพิ่มขนาดตัวอย่างกระดูกที่แตกหักนั้นจะเป็นการนำกระดูกปกติมาทำให้แตกหักด้วยการสร้างพื้นผิวของรอยแตกแบบสุ่ม (Randomly fractural surface) ขึ้นมาเพื่อนำไปลบออกจากรูปทรงสามมิติของกระดูกปกติ การสร้างพื้นผิวแบบสุ่มเราจะประยุกต์ใช้วิธีการของคุณ Mana Mahboob Kanafi (2021) [37] เป็นการสร้าง surface roughness จากสมการ power spectrum density ( $B_q$ ) ตามสมการที่ 21 (รูปที่ 46) โดย Power spectrum density นี้ถูกใช้เป็นขนาด(Magnitude) บนโดเมนความถี่ (Frequency domain) ของพื้นผิวแบบสุ่ม

$$B_q = \sqrt{\frac{4\pi^2 mn C_q}{PixelWidth^2}} \quad \text{สมการที่ 21}$$

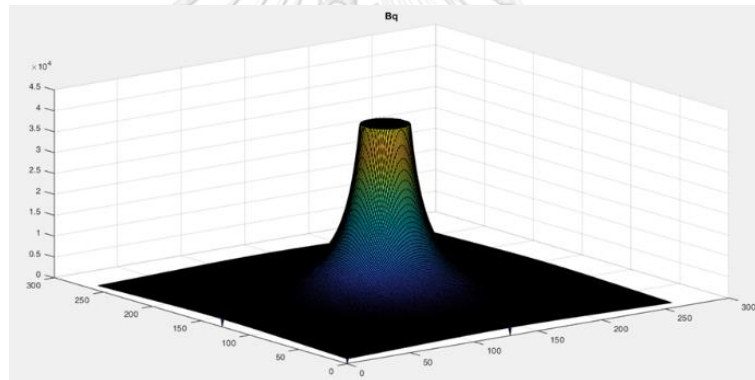
โดยที่  $B_q$  สมการ Power spectrum density,  $m$  กับ  $n$  เป็นขนาดกว้างxยาวของ surface, PixelWidth คือความละเอียดของพิกเซล และ  $C_q$  คือสมการทรงกรวยปานตัดแบบยกกำลังบนโดเมนทรงกระบอกตาม โดยที่  $q(x,y)(k+1) = \frac{2\pi k}{m}$  คือสมการทรงกรวยบนโดเมนเชิงพื้นที่

$$C_q = \begin{cases} q_r^{-2(H+1)} & ; \rho < q_r \\ \rho^{-2(H+1)} & ; else \end{cases} \quad \text{สมการที่ 22}$$

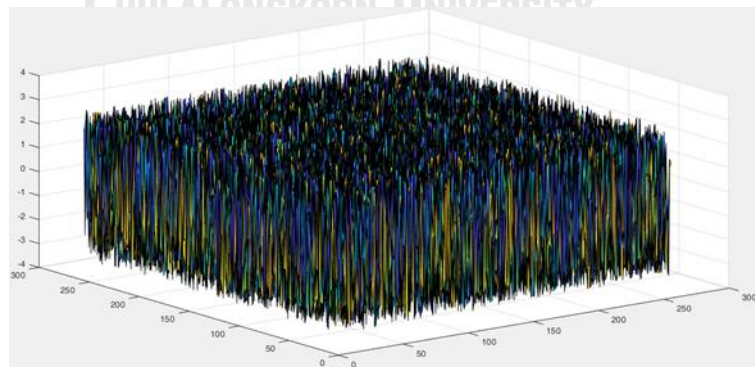
ส่วนความถี่ (Frequency) ของพื้นผิวจะถูกสุ่มขึ้นมาแบบสม่ำเสมอ (Uniformly distributed random) บนโดเมนความถี่ตามสมการที่ 23 (รูปที่ 47) ความถี่ของพื้นผิวจะถูกสุ่มใหม่ทุกครั้งที่มีสร้างพื้นผิวใหม่ จากนั้นนำขนาดและความถี่บนดังกล่าวมาแปรรูปให้อยู่บนโดเมนเชิงพื้นที่ (Spatial domain) โดยใช้ Inverse Fourier Transform จะได้ผลลัพธ์เป็นพื้นผิวแบบสุ่มตามรูปที่ 48

$$\theta = -\pi + 2\pi * \text{random}(0,1) \quad \text{สมการที่ 23}$$

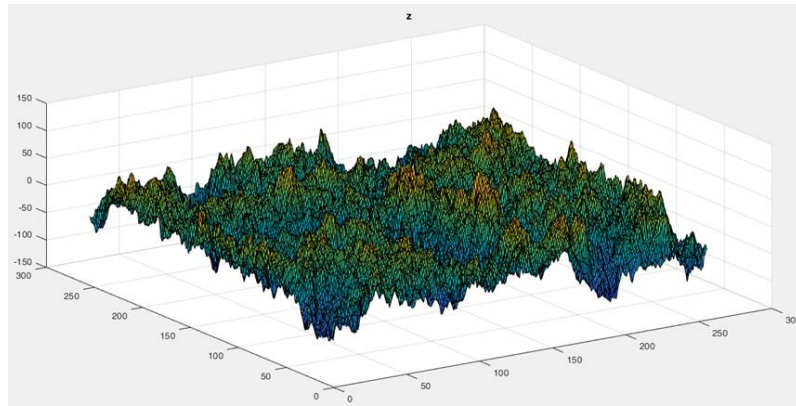
โดย  $\theta$  คือความถี่ของ random surface และ  $\text{random}(0,1)$  คือการสุ่มค่าแบบสม่ำเสมอในช่วง 0.0 - 1.0



รูปที่ 46 กราฟ Power spectrum density

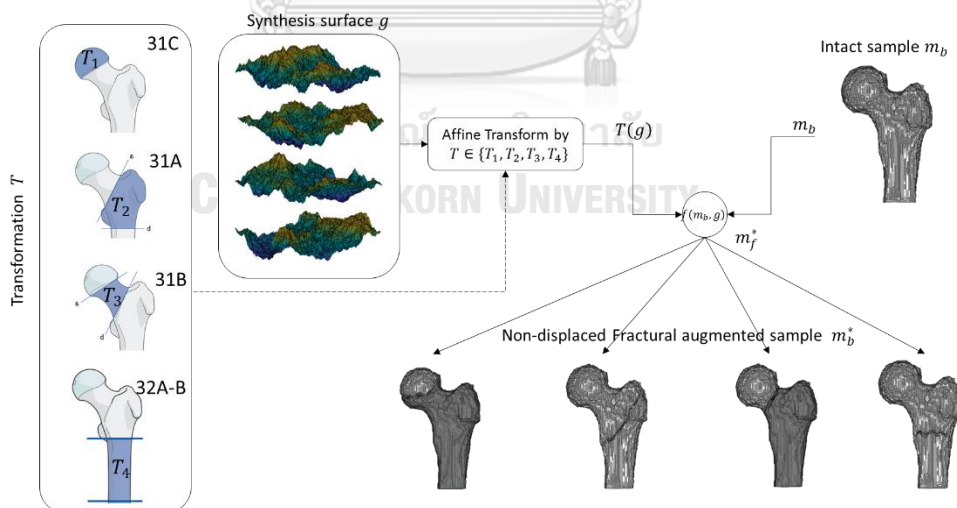


รูปที่ 47 กราฟ Uniformly random distributed frequency



รูปที่ 48 Fractal surface สำหรับ Data Augmentation

หลังจากที่สร้าง Fracture surface ได้เรียบร้อยแล้ว เราจะนำพื้นผิวที่ได้มาทำ Affine transform ให้อยู่ในตำแหน่งที่เกิดจากแตกหักได้บ่อยๆ บนกระดูกต้นขา ได้แก่ Head, Neck, Intertrochanter และ Sub-trochanter ตามที่ได้ศึกษาไว้ในหัวข้อที่ 2.1 รูปทรงทรงกระดูกเดิม ( $m_b$ ) จะถูกหักลบออกด้วยปริมาตรของพื้นผิวแบบสุ่มที่ถูกสร้างขึ้นมา จะได้กระดูกต้นขาที่แตกหักจากการทำ Fractal augmentation ( $m_b^*$ ) เครื่องหมาย \* แสดงถึงการทำให้ Fracture augmentation และปริมาตรส่วนที่ซ้อนทับของระหว่างพื้นผิวแบบสุ่มกับรูปทรงกระดูกต้นขาจะถูกใช้เป็น Auxiliary class ( $m_f^*$ ) แสดงดังรูปที่ 49



รูปที่ 49 การเพิ่มขนาดตัวอย่างกระดูกที่แตกหัก (Fracture augmentation)

### 4.3 ผลลัพธ์จากการใช้ Proposed method

ในหัวข้อนี้เราจะแสดงผลลัพธ์จากการใช้เทคนิคพิเศษต่างๆ ในการฝึกสอนโมเดลสร้างรูปทรงสามมิติของกระดูกต้นขาที่แตกหัก โดยจะเป็นการเปรียบเทียบกันระหว่างการฝึกสอนด้วยโมเดลดังต่อไปนี้

- 1) 3DReconNet การใช้โมเดลเพียงอย่างเดียวไม่ได้เพิ่มเทคนิคต่างๆ เข้าไป
- 2) 3DReconNet-AC การใช้เทคนิคคลาสช่วยฝึกสอน (Auxiliary class)
- 3) FracReconNet การใช้ทั้งสองเทคนิคพิเศษร่วมกัน (Auxiliary class + Augmentation)

และจะทำการเปรียบเทียบระหว่างกลุ่มตัวอย่างกระดูกปกติ กระดูกหักแบบไม่เคลื่อนที่ และแบบเคลื่อนที่ ผลลัพธ์เชิงปริมาณด้วยค่า mean IoU (mIoU) และ mean ASSD (mASSD) ผลลัพธ์ที่ได้ดังแสดงตารางที่ 5 ประกอบกับการทดสอบแบบจับคู่ (Paired T-test) ด้วยความเชื่อมั่นที่ 95% แสดงดังตารางที่ 6 และตารางที่ 7 แสดงการเปรียบเทียบจำนวนพารามิเตอร์, vRAM ที่ใช้, ระยะเวลาในการฝึกสอนและอนุมานผลลัพธ์

ตารางที่ 5 เปรียบเทียบค่า mIoU และ mASSD ด้วยโมเดลสร้างภาพสามมิติต่างๆ

Model	Sample types			Overall	Evaluation metrics
	Intact	Nondisplaced	Displaced		
3DReconNet	0.747±0.068	0.740±0.050	0.625±0.099	0.707±0.095	mIoU
3DReconNet-AC	0.873±0.041	0.872±0.014	0.764±0.083	0.838±0.075	
FracReconNet	0.885±0.036	0.872±0.014	0.770±0.081	0.846±0.074	
3DReconNet	1.799±0.640	1.907±0.460	2.726±0.738	2.114±0.766	mASSD (mm)
3DReconNet-AC	0.902±0.482	0.860±0.148	1.448±0.385	1.066±0.480	
FracReconNet	0.729±0.288	0.748±0.189	1.241±0.370	0.895±0.381	

ตารางที่ 6 การทดสอบแบบจับคู่ (Paired T-test)

เปรียบเทียบความแตกต่างระหว่างผลลัพธ์ที่ได้จากเทคนิคพิเศษต่าง

Baseline Technique	3DReconNet	3DReconNet-AC	FracReconNet	Evaluation metrics
3DReconNet		-53.627 ( $<0.05$ )*	-54.474 ( $<0.05$ )*	mIoU
3DReconNet-AC	53.627 ( $<0.05$ )*		-8.884 ( $<0.05$ )*	
FracReconNet	54.474 ( $<0.05$ )*	8.884 ( $<0.05$ )*		
Baseline Technique	3DReconNet	3DReconNet-AC	FracReconNet	Evaluation metrics
3DReconNet		42.548 ( $<0.05$ )*	45.555 ( $<0.05$ )*	mASSD (mm) Lower is better
3DReconNet-AC	-42.548 ( $<0.05$ )*		13.737 ( $<0.05$ )*	
FracReconNet	-45.555 ( $<0.05$ )*	-13.737 ( $<0.05$ )*		

ตารางที่ 7 เปรียบเทียบจำนวน learnable parameters, vRAM ที่ใช้,

เวลาในการฝึกสอนและอนุมานผลลัพธ์

Model	Learnable parameters	vRAM required (GB)	Training time	Inference time
3DReconNet	6,491,517	10.48 GB	18 min./epoch	1.5 sec/sample
3DReconNet-AC	6,491,965	10.59 GB	18 min./epoch	1.5 sec/sample
FracReconNet	6,491,965	10.59 GB	58 min./epoch	1.7 sec/sample

สำหรับตัวอย่างทดสอบกระดูกปกติ (รูปที่ 50) จะพบว่าโมเดล FracReconNet ที่เพิ่มเทคนิคต่างๆเข้าไป สามารถสร้างรูปทรงสามมิติของกระดูกปกติทำได้แม่นยำขึ้นอย่างมีนัยสำคัญ ( $p$ -value  $< 0.05$ ) ตามตารางที่ 6 เมื่อเทียบกับ 3DReconNet แต่โมเดล 3DReconNet-AC ที่ใช้เทคนิคฝึกสอนด้วย Auxiliary class อาจจะมี noise เกิดขึ้นได้บ้าง ในบ้างตัวอย่างทดสอบ

สำหรับตัวอย่างทดสอบกระดูกหักแบบไม่เคลื่อนที่ (รูปที่ 51) จะพบว่าถึงแม้โมเดล FracReconNet จะให้ค่า mIoU และ mASSD สูงที่สุดอย่างมีนัยสำคัญ การใช้เทคนิค Auxiliary

class ร่วมกับ Augmentation data จะให้รายละเอียดบริเวณรอยแตกหักที่ชัดเจนกว่าโมเดล 3DReconNet และ 3DReconNet-AC อย่างเห็นได้ชัด

สุดท้ายสำหรับตัวอย่างทดสอบกระดูกหักแบบเคลือบ (รูปที่ 52) จะพบว่าการใช้เทคนิคต่างๆ เสริมเข้าด้วยกัน หรือ FracReconNet จะช่วยให้การสร้างรูปทรงสามมิติทำได้ดีกว่าการฝึกสอนด้วย 3DReconNet อย่างมีนัยสำคัญ ( $p\text{-value} < 0.05$ ) ตามตารางที่ 6 ส่วนในเชิงคุณภาพ จะเห็นว่า การใช้เทคนิคแต่ละแบบนี้ให้คุณภาพผลลัพธ์ที่ไม่ได้แตกต่างกันเท่าไรนัก และยังมี noise เกิดขึ้นได้ในทุกๆ โมเดล

	1 <sup>st</sup> view	2 <sup>nd</sup> view	1 <sup>st</sup> view	2 <sup>nd</sup> view	1 <sup>st</sup> view	2 <sup>nd</sup> view	1 <sup>st</sup> view	2 <sup>nd</sup> view
<b>Input</b>								
<b>Ground Truth</b>								
<b>3DReconNet</b> mIoU = 0.747 mASSD = 1.799 mm								
	IOU = 0.799, ASSD = 1.448 mm		IOU = 0.805, ASSD = 1.434 mm		IOU = 0.735, ASSD = 1.869		IOU = 0.789, ASSD = 1.784 mm	
<b>3DReconNet-AC</b> mIoU = 0.873 mASSD = 0.902 mm								
	IOU = 0.885, ASSD = 0.777 mm		IOU = 0.912, ASSD = 0.617 mm		IOU = 0.860, ASSD = 0.876		IOU = 0.882, ASSD = 1.244 mm	
mIoU = 0.885 mASSD = 0.729 mm								
	IOU = 0.888, ASSD = 0.777 mm		IOU = 0.923, ASSD = 0.549 mm		IOU = 0.876, ASSD = 0.773		IOU = 0.893, ASSD = 0.619 mm	

รูปที่ 50 เปรียบเทียบผลลัพธ์การสร้างรูปทรงสามมิติของกระดูกต้นขาปกติ

	1 <sup>st</sup> view	2 <sup>nd</sup> view	1 <sup>st</sup> view	2 <sup>nd</sup> view	1 <sup>st</sup> view	2 <sup>nd</sup> view	1 <sup>st</sup> view	2 <sup>nd</sup> view
<b>Input</b>								
<b>Ground Truth</b>								
<b>3DReconNet</b> mIoU = 0.740 mASSD = 1.907 mm	 IOU = 0.765, ASSD = 1.501		 IOU = 0.746, ASSD = 1.629		 IOU = 0.767, ASSD = 1.850 mm		 IOU = 0.721, ASSD = 1.912	
<b>3DReconNet-AC</b> mIoU = 0.872 mASSD = 0.860 mm	 IOU = 0.875, ASSD = 0.736		 IOU = 0.866, ASSD = 0.789		 IOU = 0.888, ASSD = 1.041 mm		 IOU = 0.844, ASSD = 0.907	
<b>FracReconNet</b> mIoU = 0.872 mASSD = 0.748 mm	 IOU = 0.895, ASSD = 0.571		 IOU = 0.882, ASSD = 0.629		 IOU = 0.878, ASSD = 0.691 mm		 IOU = 0.866, ASSD = 0.681	

รูปที่ 51 เปรียบเทียบผลลัพธ์การสร้างรูปทรงสามมิติของกระดูกต้นขาหักแบบไม่เคลื่อนที่

	1 <sup>st</sup> view	2 <sup>nd</sup> view	1 <sup>st</sup> view	2 <sup>nd</sup> view	1 <sup>st</sup> view	2 <sup>nd</sup> view	1 <sup>st</sup> view	2 <sup>nd</sup> view
<b>Input</b>								
<b>Ground Truth</b>								
<b>3DReconNet</b> mIoU = 0.625 mASSD = 2.726 mm	 IOU = 0.512, ASSD = 3.192 mm		 IOU = 0.512, ASSD = 3.192 mm		 IOU = 0.655, ASSD = 2.603		 IOU = 0.677, ASSD = 2.390	
<b>3DReconNet-AC</b> mIoU = 0.746 mASSD = 1.448 mm	 IOU = 0.726, ASSD = 1.556 mm		 IOU = 0.605, ASSD = 2.496 mm		 IOU = 0.751, ASSD = 1.500		 IOU = 0.814, ASSD = 1.189	
<b>FracReconNet</b> mIoU = 0.770 mASSD = 1.241 mm	 IOU = 0.696, ASSD = 1.654 mm		 IOU = 0.638, ASSD = 2.211 mm		 IOU = 0.778, ASSD = 1.141		 IOU = 0.816, ASSD = 1.090	

รูปที่ 52 เปรียบเทียบผลลัพธ์การสร้างรูปทรงสามมิติของกระดูกต้นขาหักแบบเคลื่อนที่

สรุป โมเดล FracReconNet สามารถสร้างรูปทรงสามมิติของกระดูกต้นขาจากภาพถ่ายรังสีมุมมองแบบ internal-external oblique  $\pm 45^\circ$  ได้ดีที่สุดในทุกกลุ่มตัวอย่าง เป็นผลจากการใช้เทคนิคพิเศษ auxiliary class ช่วยฝึกสอนให้โมเดลสามารถเรียนรู้รายละเอียดของรอยแตกหัก และ fractural augmentation ช่วยเพิ่มกลุ่มตัวอย่างฝึกสอนที่มีรอยแตกหักของกระดูก ซึ่งทั้งสอง



เทคนิคนี้ทำงานร่วมกันได้อย่างมีประสิทธิภาพ อย่างไรก็ตามโมเดลที่กล่าวมาข้างต้นสามารถทำงานได้ดีกับภาพรังสีสองภาพที่มีมุมการฉากรังสีตั้งฉากกันพอดีเท่านั้น ซึ่งในกรณีที่ภาพถ่ายรังสีมีความคลาดเคลื่อนหรือไม่ตั้งฉากกัน จะกล่าวถึงในบทถัดไป



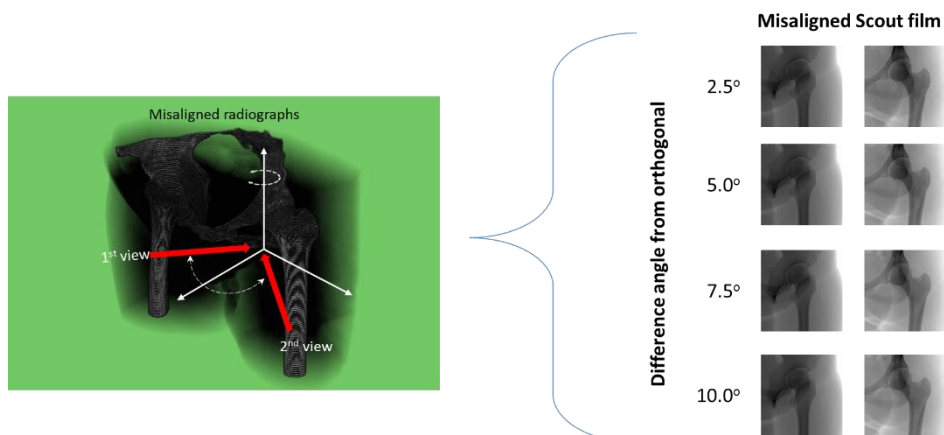
## บทที่ 5

### การปรับปรุงประสิทธิภาพของโมเดลสร้างรูปทรงสามมิติ

เนื่องจากผ่านมาเราทำการฝึกสอนโมเดลสร้างรูปทรงสามมิติโดยใช้ Input เป็น Scout film สองมุมมองที่ทำมุมตั้งฉากกันพอดีเท่านั้น แต่ในการใช้งานจริงภาพถ่ายรังสีที่บันทึกออกมามีโอกาสที่แต่ละมุมมองจะทำมุมกันไม่พอดีกับมุม 90 องศา หรือมีความคลาดเคลื่อนเชิงมุม (Rotational error) เกิดขึ้น ซึ่งความคลาดเคลื่อนนั้นเกิดขึ้นได้หลายสาเหตุ เช่น การจัดตำแหน่งการนอนของคนไข้ที่ไม่แม่นยำ, การตั้งวางการฉายรังสีที่ผิดพลาด เป็นต้น ส่งผลให้ผลลัพธ์ที่ได้มีความแม่นยำลดลง ดังนั้นเพื่อเป็นปรับปรุงประสิทธิภาพของโมเดลเราจึงจำเป็นต้องฝึกสอนโมเดลด้วยภาพ Scout film สองภาพที่มีความคลาดเคลื่อนเชิงมุม และทำการปรับแต่งพารามิเตอร์(Hyperparameter) ที่เกี่ยวข้องเพื่อให้โมเดลสร้างรูปทรงสามมิติสามารถให้ผลลัพธ์ที่แม่นยำแม้ว่า Input ที่รับเข้ามามีความคลาดเคลื่อนเชิงมุมก็ตาม

#### 5.1 การสร้างภาพถ่ายรังสีแบบดิจิทัลแบบมุมคลาดเคลื่อน (Misaligned Digitally Reconstructed Radiograph)

ในขั้นตอนนี้เราจะใช้ Scout film สองภาพที่มีมุมมองคลาดเคลื่อนกัน  $90 \pm 10$  องศา โดยการนำภาพ CT ที่มีค่าเป็นสัมประสิทธิ์การลดทอนเชิงเส้น  $\mu_{(x,y,z)}$  หน่วย  $\text{mm}^{-1}$  ที่ได้ จากความสัมพันธ์สมการที่ 3 หรือสมการที่ 4 ในหัวข้อที่ 3.2.2 มาใช้ในการสังเคราะห์ Scout film โดยจะเปลี่ยนมุมที่รังสีวิ่งผ่านภาพ CT เป็นสองมุมมองที่ทำมุมตั้งฉากกัน และมีความคลาดเคลื่อนเชิงมุมตั้งแต่ 0 – 10 องศา หลังจากนั้นทำการสังเคราะห์ภาพรังสีแบบดิจิทัลด้วยเทคนิค Ray casting โดยดัดแปลงกฎของ Beer Lambert [16] ดังที่ได้กล่าวไว้แล้วในสมการที่ 5 จะได้ผลลัพธ์ดังแสดงดังรูปที่ 53 ชุดข้อมูลที่สังเคราะห์ขึ้นจะถูกนำไปใช้สำหรับการฝึกสอนโมเดลสร้างรูปทรงสามมิติในหัวข้อถัดไป



รูปที่ 53 การสังเคราะห์ Scout film ที่มีความคลาดเคลื่อนเชิงมุมที่  $90 \pm 10$  องศา

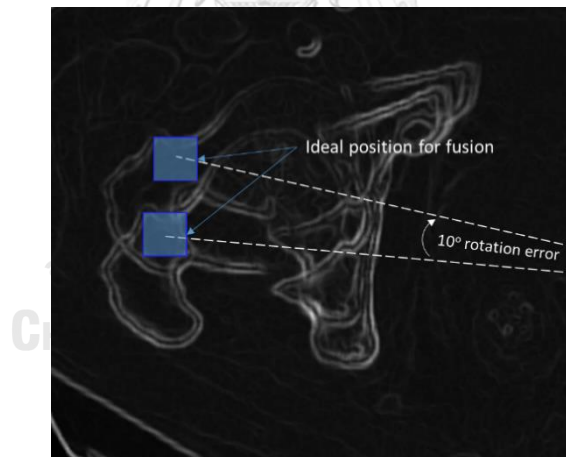
## 5.2 การปรับปรุงโมเดลการสร้างภาพสามมิติที่สามารถประมวลผลภาพถ่ายรังสีสองมุมมองที่มีความคลาดเคลื่อนเชิงมุม

การแก้ปัญหาเรื่องภาพถ่ายรังสีสองมุมมองที่ทำให้มุมคลาดเคลื่อนกัน สามารถทำได้หลากหลายวิธีดังนี้

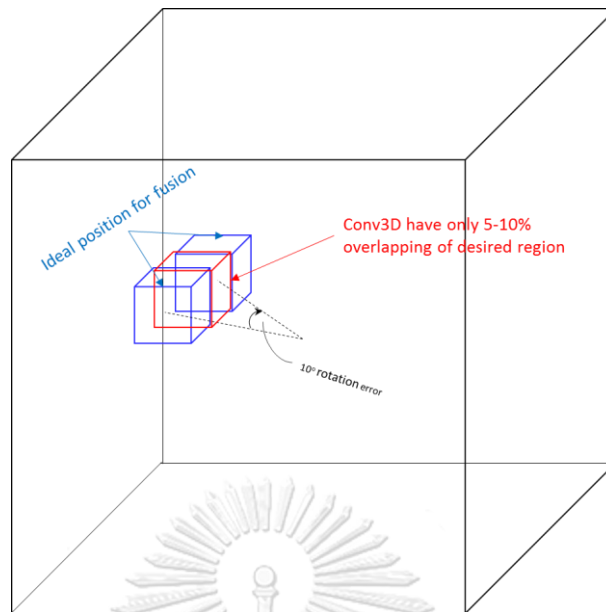
- 1) การปรับมุมมองของภาพถ่ายรังสีที่เป็นข้อมูลขาเข้า: เพื่อปรับมุมมองของภาพถ่ายรังสีแต่ละภาพให้มีมุมมองตรงกับ internal/external oblique view มุม 45 องศา โดยประยุกต์ใช้โครงข่ายที่ออกแบบมาสำหรับการหมุนภาพในสามมิติซึ่งใช้โครงข่ายแบบ GAN [38, 39] ก่อนนำภาพที่ถูกหมุนแล้วไปใช้เป็นข้อมูลขาเข้าของโมเดลการสร้างรูปทรงสามมิติ การใช้วิธีนี้มีข้อดี คือ ช่วยให้เราสามารถปรับภาพถ่ายรังสีให้อยู่ในมุมมองที่เราต้องการได้อย่างสมจริง แต่วิธีนี้มีโอกาสลดทอนรายละเอียดของรอยแตกหักของภาพถ่ายรังสีให้มีผิดเพี้ยนจากภาพต้นฉบับได้ ซึ่งอาจส่งผลให้ผลลัพธ์จากการการสร้างรูปทรงสามมิติสร้างรอยแตกหักของกระดูกต้นขาที่มีความผิดพลาดมากขึ้น
- 2) การปรับเปลี่ยนตัวประมวลผลพื้นฐานของโครงข่าย: ใช้ตัวประมวลผลที่มีคุณสมบัติที่ไม่ขึ้นอยู่กับการหมุนของวัตถุ (Rotational invariant) จากงานวิจัย CyCNN [40] วิธีการนี้เป็นการปรับปรุงตัวประมวลผลแบบ Conv2D ให้สามารถรับรองการหมุนของภาพสองมิติขาเข้าได้ โดยแปลงทั้งภาพและตัวประมวลผลจากระบบพิกัดคาร์ทีเซียน (Cartesian coordinate system) ให้เป็นระบบพิกัดเชิงขั้ว (Polar coordinate system) เพื่อให้โครงข่ายสามารถเรียนรู้คุณสมบัติจากรูปภาพขาเข้าที่การหมุนได้ แต่อย่างไรก็ตามวิธีการนี้สามารถนำไปใช้ได้กับการหมุนของภาพในสองมิติเท่านั้น ดังนั้นวิธีการนี้จึงยังไม่สามารถนำมาประยุกต์กับภาพที่มีความคลาดเคลื่อนจากการหมุนในสามมิติของทิศทางการฉายรังสีได้
- 3) การปรับปรุงการทำงานของ fusion module: ให้สามารถผสมคุณสมบัติเชิงปริมาตร (volumetric feature) ที่มีความคลาดเคลื่อนเชิงมุมได้ โดยการปรับปรุงรูปแบบการประมวลผลภายในใหม่ วิธีนี้อาจเป็นวิธีที่ง่ายและซับซ้อนน้อยที่สุดเนื่องจากไม่จำเป็นต้องออกแบบ Encoder กับ Decoder ใหม่ โดยยังคงใช้ภาพถ่ายรังสีต้นฉบับ (ซึ่งไม่ถูกลดทอนรายละเอียดจากการใช้วิธีการแรกได้) ดังนั้นผู้วิจัยจึงเลือกใช้วิธีการนี้ในการแก้ปัญหา โดยมีรายละเอียดดังต่อไปนี้

เริ่มต้นจากการวิเคราะห์ปัญหาที่เกิดขึ้นจาก Fusion module เดิม ซึ่งเป็นการนำ Decoded features หรือ  $D_v^{(i)}$  ที่ได้จากขั้นตอนการถอดรหัสในแต่ละมุมมอง  $v \in \{1,2\}$  มารวมกัน

โดยมีการสลับลำดับของเทนเซอร์ (Permutation,  $P$ ) ของ  $D_2^{(i)}$  หรือการหมุน 90 องศารอบแกนแนวตั้ง ให้มีมุมมองที่ตรงกับ  $D_1^{(i)}$  ก่อน จากนั้นใช้ Conv3D{filter=3x3x3, stride=1, padding=1} + InstanceNorm3d(IN) + ReLU ในการผสมคุณสมบัติ  $D_1^{(i)}$  กับ  $P(D_2^{(i)})$  เข้าด้วยกันดังที่ได้อธิบายไว้ในหัวข้อ 3.3 (รูปที่ 40) หากพิจารณาตัวอย่างของการผสมคุณสมบัติ  $D_1^{(i)}$  กับ  $P(D_2^{(i)})$  จากภาพถ่ายรังสีที่มีมุมมองคลาดเคลื่อนกัน 10 องศาตามรูปที่ 53 การผสมคุณสมบัติ Decoded features  $D_v^{(i)}$  ทั้งสองมุมมองเข้าด้วยกันในแบบอุดมคติ (Idea fusion process) จะเกิดขึ้นเมื่อเป็นการผสมคุณสมบัติ ณ ตำแหน่ง key-point บน anatomical structure ที่สอดคล้องกันของแต่ละปริภูมิสามมิติ  $D_1^{(i)}$  และ  $D_2^{(i)}$  (ตามกรอบสีน้ำเงินแสดงดังรูปที่ 54 และรูปที่ 55) แต่ทว่าการใช้ 3D Convolution operation ที่มีขนาด 3x3x3 ในการผสมคุณสมบัติ Decoded features  $D_v^{(i)}$  จะสามารถผสมคุณสมบัติทั้งสองได้ประมาณ 20% ของการผสมคุณสมบัติในอุดมคติเท่านั้น (ตามกรอบสีแดงในรูปที่ 55) เนื่องจาก 3D Convolution จะสามารถผสมคุณสมบัติได้ที่ตำแหน่งบนปริภูมิสามมิติเท่านั้น ไม่สามารถผสมคุณสมบัติที่สอดคล้องกันระหว่างสองตำแหน่งได้อย่างครอบคลุม ซึ่งเกิดขึ้นจากการคลาดเคลื่อนเชิงมุมในระหว่างการถ่ายภาพรังสีจริงๆ ซึ่งไม่สามารถระบุได้ว่าความคลาดเคลื่อนเชิงมุมดังกล่าวมีค่าเท่าไรและมากน้อยเพียงไร

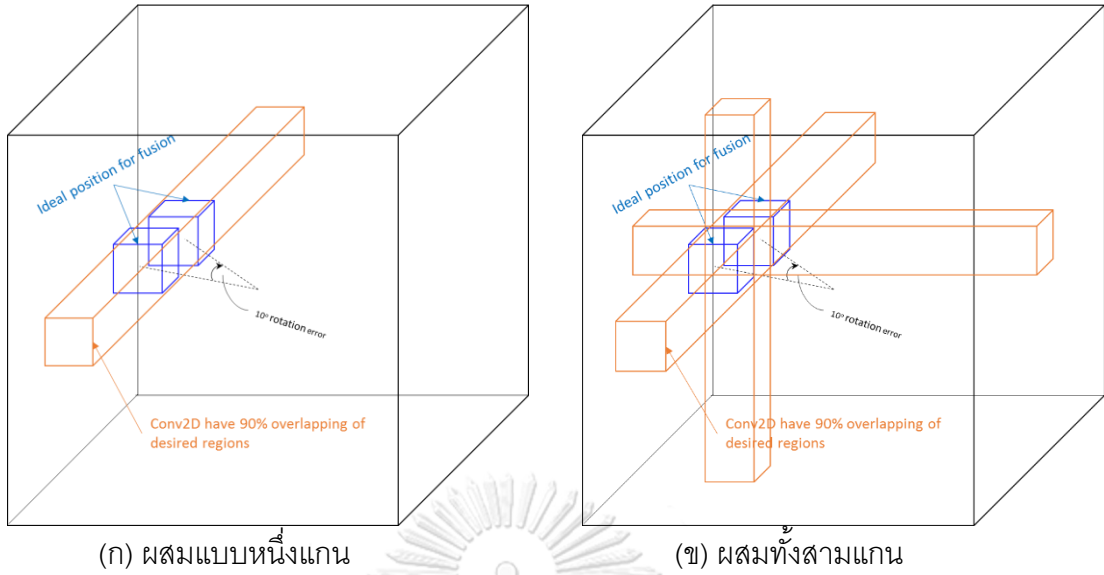


รูปที่ 54 ตัวอย่างภาพคุณสมบัติ  $D_1^{(i)}$  กับ  $P(D_2^{(i)})$  ที่ซ้อนทับกันของภาพรังสีสองมุมมองที่มีความคลาดเคลื่อนเชิงมุม 10 องศา กรอบสีน้ำเงินแสดงถึงตำแหน่ง key-point บน anatomical structure เดียวกันของแต่ละมุมมอง



รูปที่ 55 การผสมคุณสมบัติของ *Native fusion module* ในหัวข้อที่ 3.3 หมายถึงการบดสีน้ำเงิน ทั้งสองแสดงตำแหน่งการผสมกันในอุดมคติ ส่วนกรอบสีแดงแสดงตำแหน่งการผสมกันที่เกิดขึ้น ขณะใช้ *Conv3D*

จากการวิเคราะห์ปัญหาของความคลาดเคลื่อนในการผสมคุณสมบัติที่เกิดขึ้น เราจึงมีความจำเป็นที่จะต้องปรับปรุงวิธีการผสมคุณสมบัติให้สามารถประมวลผลได้อย่างมีประสิทธิภาพ ถึงแม้ว่าจะมีการคลาดเคลื่อนเชิงมุมเกิดขึ้นระหว่างคุณสมบัติ  $d_1^{(i)}$  และ  $d_2^{(i)}$  ก็ตาม เป้าหมายของการผสมคุณสมบัติคือการเลือกใช้ตัวประมวลผลที่สามารถรวบรวมคุณสมบัติทั้งสองตำแหน่งได้ในตำแหน่งของการประมวลผลเดียวกัน  $(i, j, k)$  เราจะแก้ปัญหานี้โดยการใช้ 2D Convolution ขนาด  $3 \times 3$  (กรอบสีเหลือง) ในการผสมคุณสมบัติแทนการใช้ 3D Convolution ขนาด  $3 \times 3 \times 3$  ตามรูปที่ 56 เนื่องจาก 2D Convolution จะประมวลผลตามมิติ channel ใดๆ บนเทนเซอร์สองมิติ ซึ่งเราสามารถกำหนดให้ 2D Convolution ทำการคอนโวลูชันบนมิติใดก็ได้ เราจะขอเรียกว่า “การผสมคุณสมบัติแบบหนึ่งแกน” แสดงดังรูปที่ 56(ก) จะเห็นว่า 2D Convolution สามารถวิ่งผ่านบริเวณการผสมกันแบบอุดมคติ (กรอบสีน้ำเงิน) ซึ่งจะครอบคลุมปริมาตรมากกว่า 90% ของปริมาตรการผสมแบบอุดมคติ และเราจะใช้ 2D Convolution สำหรับแต่ละแกนของปริมาตรเพื่อให้ครอบคลุมการคลาดเคลื่อนในทุกๆ แกนหมุน ซึ่งเราจะตั้งชื่อขั้นตอนการผสมคุณสมบัตินี้ว่า “การผสมแบบแกน (Axial-fusion module)” แสดงดังรูปที่ 56(ข) ซึ่งมีรายละเอียดดังต่อไปนี้



รูปที่ 56 การผสมคุณสมบัติด้วย 2D Convolution แบบ Axial-fusion

นำ decoded features  $d_{\nu}^{(i)}$  ของแต่ละระดับชั้น  $0 \leq i \leq L - 1$  มาทำการผสมเข้าด้วยกัน โดย จะถูกสลับลำดับของเทนเซอร์เพื่อหมุน 90 องศารอบแนวตั้งก่อน จากนั้นนำมาทำการเชื่อมต่อคุณสมบัติ (channel-wise concatenation) ในมิติ depth, height และ width ของเทนเซอร์ แล้วทำการ Conv2D { filter=3x3x3, stride=1, padding=1} บนแต่ละแกนมิติของเทนเซอร์ ตามลำดับ โดยแต่ละมิติจะถูกกำหนดให้ channel ของข้อมูลขาออกมีขนาดเท่ากับ resolution ของระดับชั้นนั้นๆ แสดงดังสมการที่ 24

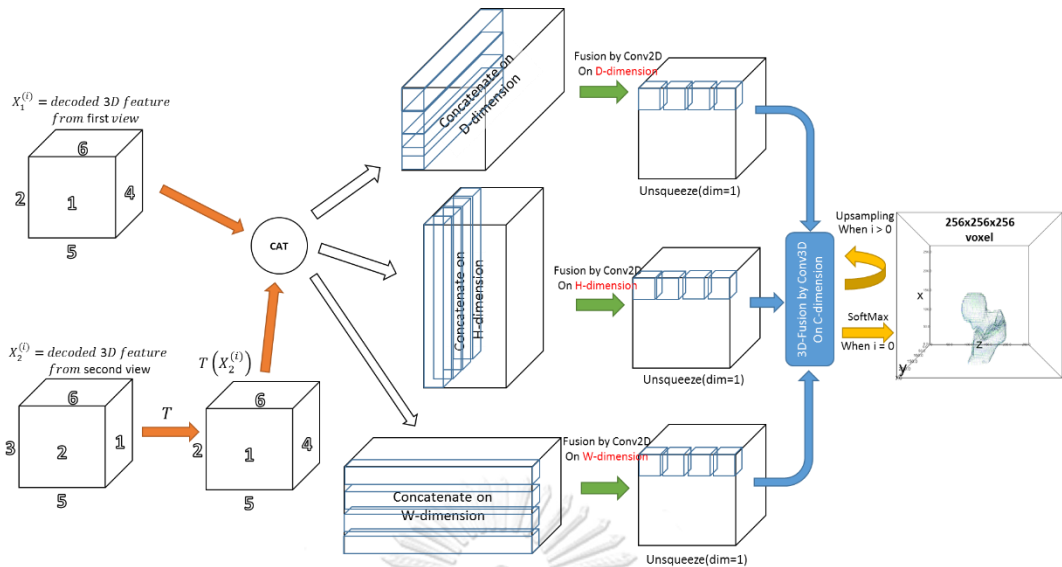
$$D_{axis}^{(i)} \equiv d_{axis}^{(i)} = \mathcal{H}_{2D_{axis}} \left[ d_1^{(i)}, P \left( d_2^{(i)} \right) \right]_{axis} \quad \text{สมการที่ 24}$$

โดยที่  $axis \in \{depth, height, width\}$  คือ แกนของมิติ ความลึก, สูง และกว้าง ตามลำดับ,  $i \in [0, L - 1]$  คือระดับชั้น, เครื่องหมาย  $\equiv$  คือการแปลงรูป 2D features ให้อยู่ในรูปของ 3D features และ  $[ ]_{axis}$  แสดงถึงการทำ concatenation บนแกนมิติใด

จากนั้นนำ  $D_{axis}^{(i)}$  แต่ละแกนมิติมาทำการผสมกันด้วย Conv3d {filter=3x3x3, stride=1, padding=1} + InstanceNorm3d(IN) + ReLU ตามสมการที่ 25 จะได้  $F^{(i)}$  และทำการ Up-sampling เพื่อเพิ่ม resolution เป็น  $F^{(i+1)}$  ให้นำไปผสมกับ  $D_{axis}^{(i+1)}$  ที่อยู่ในระดับชั้นถัดๆ ไป

$$F^{(i)} = \begin{cases} \mathcal{H}_{3D} \left( \left[ D_{depth}^{(i)}, D_{height}^{(i)}, D_{width}^{(i)}, \mathcal{U}_{3D}(F^{(i+1)}) \right] \right), & i < L - 1 \\ \mathcal{H}_{3D} \left( \left[ D_{depth}^{(i)}, D_{height}^{(i)}, D_{width}^{(i)} \right] \right), & i = L - 1 \end{cases} \quad \text{สมการที่ 25}$$

โดยที่  $\mathcal{H}_{3D}$  แสดงการทำ 3D Convolution,  $i = L - 1$  คือระดับชั้นต่างๆ และ  $[ ]$  คือแสดงถึงการทำ concatenation



รูปที่ 57 โครงสร้าง Axial-fusion module ใช้สำหรับเพิ่มประสิทธิภาพการรวมคุณสมบัติที่มีมุมมองคลาดเคลื่อนกัน

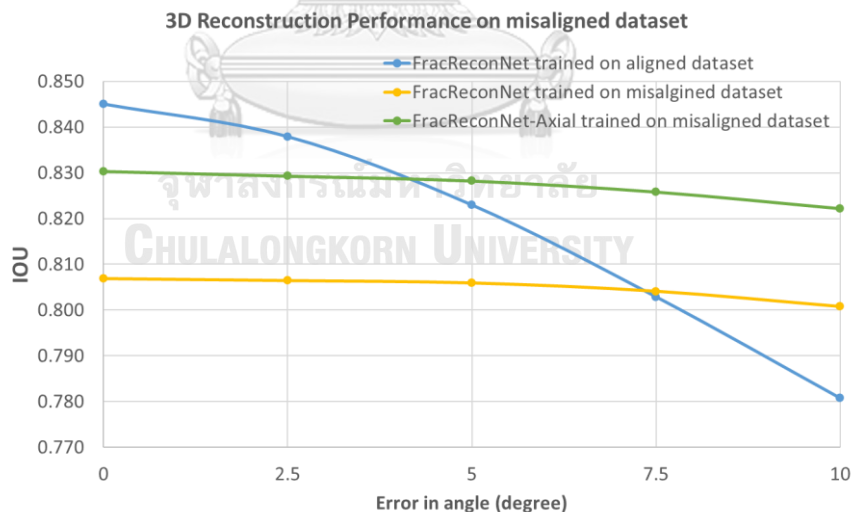
### 5.3 ผลลัพธ์จากการปรับปรุงโมเดลสร้างรูปทรงสามมิติด้วย Axial-fusion module

เพื่อทำการเปรียบเทียบผลลัพธ์ที่จากการสอนโมเดลสร้างรูปทรงสามมิติจากภาพถ่ายรังสีสองมุมมองที่คลาดเคลื่อนกัน เราทำการทดสอบกับชุดตัวอย่างทดสอบ และทำการเปรียบเทียบผลลัพธ์จาก 3 โมเดล ดังต่อไปนี้

- FracReconNet-Aligned ซึ่งถูกสอนกับชุดข้อมูลภาพถ่ายรังสีสองมุมมองที่ตั้งฉากกันพอดี (Aligned dataset)
- FracReconNet-Misaligned ซึ่งถูกสอนกับชุดข้อมูลภาพถ่ายรังสีสองมุมมองที่คลาดเคลื่อนกัน (Misaligned dataset)
- FracReconNetAxial-Misaligned ใช้ Axial-fusion module ที่คิดค้นขึ้นในหัวข้อที่ 5.2 และถูกสอนกับชุดข้อมูลภาพถ่ายรังสีสองมุมมองที่คลาดเคลื่อนกัน (Misaligned dataset)

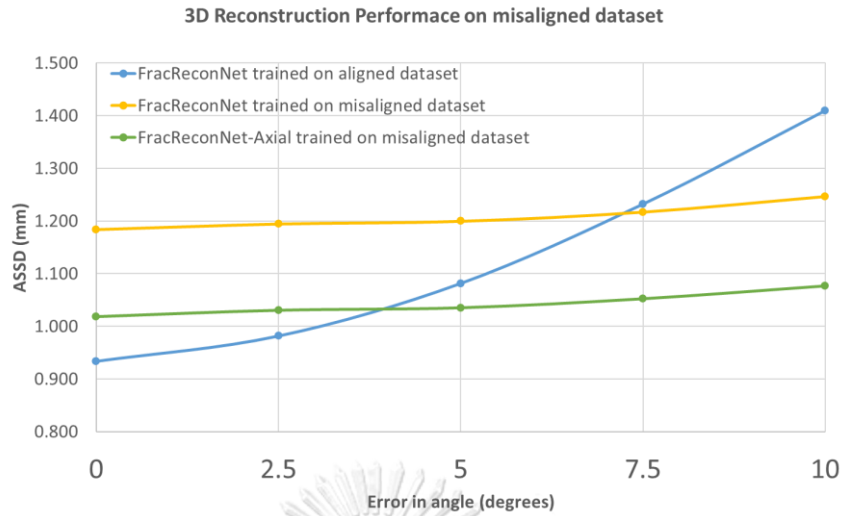
ในการประเมินเราจะใช้ IoU และ ASSD metrics ในการวัดความแม่นยำของรูปทรงสามมิติที่ถูกสร้างขึ้น แสดงดังรูปที่ 58 และรูปที่ 59 ตามลำดับ และการทดสอบแบบจับคู่ (Paired T-test) เปรียบเทียบความแตกต่างระหว่างผลลัพธ์ที่ได้ของแต่ละโมเดล ด้วยความเชื่อมั่นที่ 95% แสดงดังตารางที่ 8 และตารางที่ 9 (ผลลัพธ์อย่างละเอียดแสดงในภาคผนวก A ถึง C) สามารถอธิบายผลลัพธ์ได้ดังนี้

- 1) พิจารณาผลลัพธ์กรณีที่ไม่มี ความคลาดเคลื่อนเชิงมุม จะพบว่าโมเดล A จะให้ผลลัพธ์ที่แม่นยำที่สุด (IoU=0.846, ASSD=0.934 mm) เนื่องจากถูกฝึกสอนให้ประมวลผลบนภาพรังสีสองมุมมองที่ตั้งฉากกันพอดีเท่านั้น จึงทำให้โมเดลดังกล่าวทำงานได้ดีที่สุดเมื่อชุดตัวอย่างทดสอบไม่มีความคลาดเคลื่อนเชิงมุม ส่วนโมเดล B และ C ให้ผลลัพธ์ที่แย่กว่า เพราะถูกฝึกสอนด้วยข้อมูลภาพถ่ายรังสีที่มีความคลาดเคลื่อนเชิงมุมที่หลากหลาย ทำให้การอนุมานผลลัพธ์จากข้อมูลภาพที่ไม่มี ความคลาดเคลื่อนทำให้ได้แย่กว่าโมเดล A
- 2) พิจารณาผลลัพธ์กรณีที่เมื่อมีความคลาดเคลื่อนเชิงมุม จะพบว่าที่มุมคลาดเคลื่อนต่ำกว่า 4 องศา โมเดล A จะยังให้ผลลัพธ์ที่ดีที่สุด (IoU=0.838-0.846, ASSD=0.934-0.982 mm) เนื่องจากโมเดล A ถูกฝึกสอนด้วยภาพรังสีที่ไม่มี ความคลาดเคลื่อนเชิงมุมเท่านั้น จึงทำให้ แต่เมื่อภาพรังสีมีความคลาดเคลื่อนเชิงมุมเกิน 4 องศาขึ้นไป โมเดล C จะทำงานได้แม่นยำมากที่สุด (IoU=0.826, ASSD=1.052 mm) ทั้งโมเดล B และ C ที่ถูกสอนข้อมูลภาพรังสีที่มีความคลาดเคลื่อนเชิงมุมจะมีอัตราการลดลงของค่าความแม่นยำทั้ง IoU และ ASSD อย่างช้าๆ ส่วนโมเดล A ที่ถูกสอนด้วยภาพรังสีที่ตั้งฉากพอดีนั้น จะมีอัตราการลดลงของค่าความแม่นยำทั้ง IoU และ ASSD อย่างรวดเร็วเมื่อความคลาดเคลื่อนเชิงมุมมีค่ามากยิ่งขึ้น (IoU=0.781-0.823, ASSD=1.082-1.410 mm)



รูปที่ 58 กราฟแสดงค่า Average IoU ของผลลัพธ์การสร้างรูปทรงสามมิติ ที่ความคลาดเคลื่อนเชิงมุมต่าง (2.5 ถึง 10 องศา)





รูปที่ 59 กราฟแสดงค่า Average ASSD ของผลลัพธ์การสร้างรูปทรงสามมิติ ที่ความคลาดเคลื่อนเชิงมุมต่าง (2.5 ถึง 10 องศา)

ตารางที่ 8 เปรียบเทียบความแตกต่างระหว่างผลลัพธ์ที่ได้จากโมเดล FracReconNet-Aligned, FracReconNet-Misaligned และ FracReconNetAxial-Misaligned ด้วย IoU

		<b>Model Comparison on mIoU</b>			
		Baseline	FracReconNet Aligned	FracReconNet Misaligned	FracReconNet-Axial Misaligned
T-score (P-value)	Technique				
	FracReconNet Aligned			9.385 (<0.05) *	-9.339 (<0.05) *
	FracReconNet Misaligned		-9.385 (<0.05) *		-24.185 (<0.05) *
	FracReconNet-Axial Misaligned		9.339 (<0.05) *	24.185 (<0.05) *	

ตารางที่ 9 เปรียบเทียบความแตกต่างระหว่างผลลัพธ์ที่ได้จากโมเดล FracReconNet-Aligned, FracReconNet-Misaligned และ FracReconNetAxial-Misaligned ด้วย ASSD

		<b>Model Comparison on mASSD</b>			
		Baseline	FracReconNet Aligned	FracReconNet Misaligned	FracReconNet-Axial Misaligned
T-score (P-value)	Technique				
	FracReconNet Aligned			-8.292 (<0.05) *	10.777 (<0.05) *
	FracReconNet Misaligned		8.292 (<0.05) *		22.500 (<0.05) *
	FracReconNet-Axial Misaligned		-10.777 (<0.05) *	-22.500 (<0.05) *	

ตารางที่ 10 เปรียบเทียบจำนวนพารามิเตอร์ (Learnable parameters), การใช้ vRAM, ระยะเวลาฝึกสอน (Training time) และระยะเวลาอนุมานผลลัพธ์ (Inference time) ระหว่างโมเดล FracReconNet ที่ใช้ Fusion module แบบเดิมกับแบบ Axial-fusion

Fusion module	Learnable parameters	vRAM required (GB)	Training time	Inference time
Native	6,491,965	10.59 GB	~1 hr./epoch	~1.7 sec/sample
Axial-fusion	17,482,977	34.42 GB	~9 hr./epoch	~6.5 sec/sample

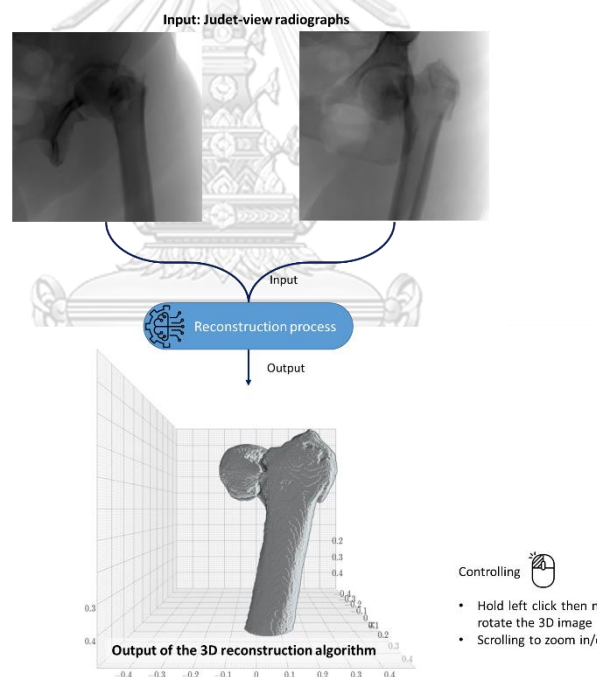
จากผลลัพธ์ที่ได้จะเห็นว่า การใช้ Axial-fusion module จะให้ผลลัพธ์การสร้างรูปทรงสามมิติได้ดีมากยิ่งขึ้นเมื่อภาพถ่ายรังสีมีความคลาดเคลื่อนเชิงมุม แต่ก็ต้องมีข้อสังเกตที่ต้องพิจารณาเพิ่มเติม จากตารางที่ 10 จะพบว่า Axial-fusion module จะมีจำนวน learnable parameters เพิ่มขึ้นและต้องใช้ vRAM มากกว่าเดิมประมาณ 2 เท่า และระยะเวลาในการฝึกสอนและการอนุมานผลลัพธ์เพิ่มขึ้นประมาณ 3.5 เท่า ส่งผลให้ต้องใช้หน่วยประมวลผลกราฟิกส์ (Graphics processing unit: GPU) ที่มีความจุ Virtual RAM ที่มากขึ้น ซึ่งมีราคาแพง และต้องการพื้นที่ติดตั้งเพิ่มเติมบนคอมพิวเตอร์

## บทที่ 6

### การประเมินเชิงคุณภาพของโมเดลสร้างรูปทรงสามมิติ

#### 6.1 วิธีการประเมินเชิงคุณภาพ

ในการประเมินเชิงคุณภาพของโมเดลสร้างรูปทรงสามมิติ จะให้คุณหมอผู้เชี่ยวชาญในการวิเคราะห์ภาพถ่ายรังสี ได้แก่ รังสีแพทย์ (Radiologist), ศัลยแพทย์ทางออร์โธปิดิกส์ (Orthopaedic surgery) หรือผู้อื่นๆที่เกี่ยวข้องเป็นผู้ทำการประเมินคุณภาพของผลลัพธ์ โดยผู้เชี่ยวชาญที่เข้าร่วมทำการประเมินจะได้รับตัวอย่างทดสอบจำนวน 12 ตัวอย่าง แต่ละตัวอย่างทดสอบจะประกอบไปด้วยภาพถ่ายรังสี (2D radiographs) ในมุมมอง Internal/External oblique ที่มุม  $-45$  และ  $+45$  องศา บริเวณกระดูกต้นขา (เป็น Input ของอัลกอริทึม) จำนวน 2 ภาพต่อ 1 ตัวอย่างทดสอบ และรูปทรงสามมิติของกระดูกที่ถูกสร้างขึ้น (3D reconstructed image) ซึ่งเป็น Output ของอัลกอริทึม แสดงดังรูปที่ 60

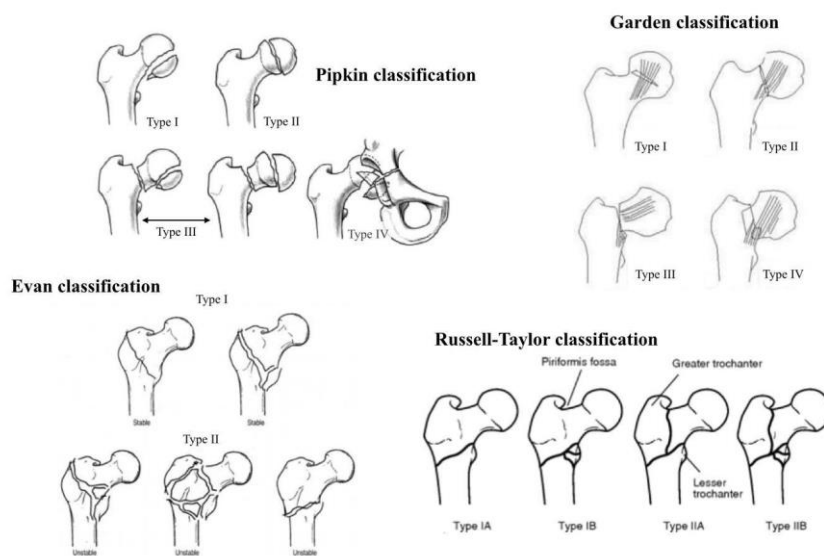


รูปที่ 60 ตัวอย่างทดสอบ ซึ่งประกอบไปด้วย ภาพถ่ายรังสี (Input) และรูปทรงสามมิติที่ถูกสร้างขึ้น (Output) เพื่อใช้สำหรับประเมินเชิงคุณภาพ

จากนั้นให้จะให้ผู้เชี่ยวชาญทำการเลือกชนิดการแตกหักของกระดูกต้นขาตามการจำแนกประเภทหลัก (Main-categories) ของ Pipkin, Garden, Evan และ Russell-Taylor fracture classification หรืออื่นๆ (ตามรูปที่ 61) ซึ่งแต่ละประเภทหลักจะมีการจำแนกประเภทย่อย (Sub-categories) หรือชนิดของการแตกหัก (Type) ที่แตกต่างกันออกไป โดยวิเคราะห์จากภาพสามมิติ

ของกระดูกต้นขาที่อัลกอริทึมสร้างขึ้น และทำการประเมินเพื่อเปรียบเทียบภาพสามมิติของกระดูกที่อัลกอริทึมสร้างขึ้นกับภาพถ่ายรังสีสองมุมมอง ว่ามีความสอดคล้องกันหรือไม่ โดยทำการประเมินความพึงพอใจของผลลัพธ์เชิงคุณภาพของภาพสามมิติที่อัลกอริทึมสร้างขึ้น ผ่าน Google Doc ด้วยเกณฑ์การให้คะแนนดังต่อไปนี้

- I. Correctness of the femoral shape: ความถูกต้องของรูปทรงกระดูกต้นขา
- II. Correctness of fracture details: ความถูกต้องของร่องรอยการแตกหักของกระดูกต้นขา
- III. Correctness of the bone fragment position: ความถูกต้องของตำแหน่งการวางตัวของกระดูกที่แตกหักละชิ้น
- IV. Beneficial in diagnosis of femoral fracture: ประโยชน์ต่อการวินิจฉัยการแตกหักของกระดูกต้นขา
- V. Benefits for preoperative planning of femoral fracture: ประโยชน์ต่อการนำไปใช้ในการวางแผนการผ่าตัด
- VI. Satisfaction: ความพึงพอใจโดยรวมในการใช้อัลกอริทึมนี้



รูปที่ 61 การจำแนกการแตกหักของกระดูกต้นขาแบบต่างๆ

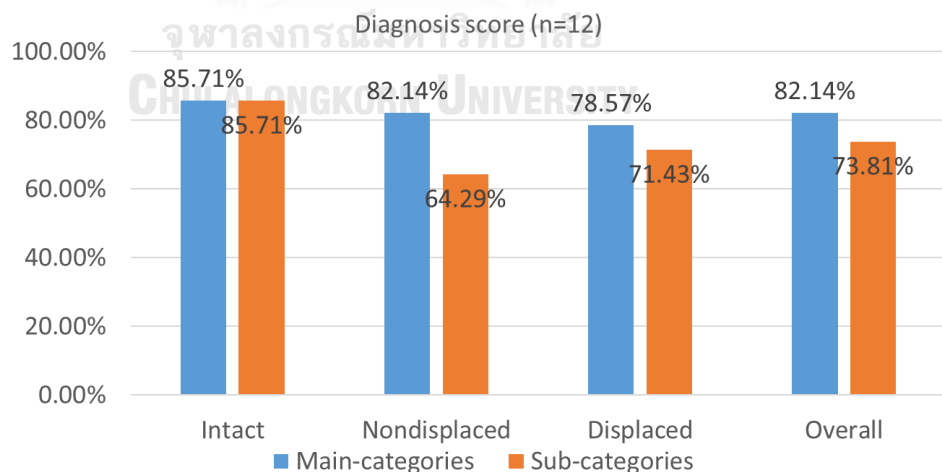
โดยการประเมินจะเป็นแบบ blind test กล่าวคือ ผู้ทำการประเมินจะไม่ทราบรูปทรงสามมิติที่แท้จริง หรือ Ground truth ของตัวอย่างทดสอบนั้นๆ แต่จะได้รับรูปทรงสามมิติที่ถูกสร้างขึ้นโดยอัลกอริทึมเท่านั้น และตัวอย่างทดสอบจะถูกแบ่งออกเป็น 3 กลุ่มตัวอย่างทดสอบ ได้แก่ Intact, nondisplaced และ displaced samples โดยจะใช้เกณฑ์ประเมินตามข้อ I-IV ที่กล่าวไว้ข้างต้น

## 6.2 ผลลัพธ์จากการประเมินเชิงคุณภาพ

จากการประเมินเชิงคุณภาพในหัวข้อที่ 6.1 เราสามารถสรุปผลได้ดังต่อไปนี้

### 1) ความสามารถนำไปวินิจฉัยเพื่อจำแนกการแตกหักของกระดูกต้นขา

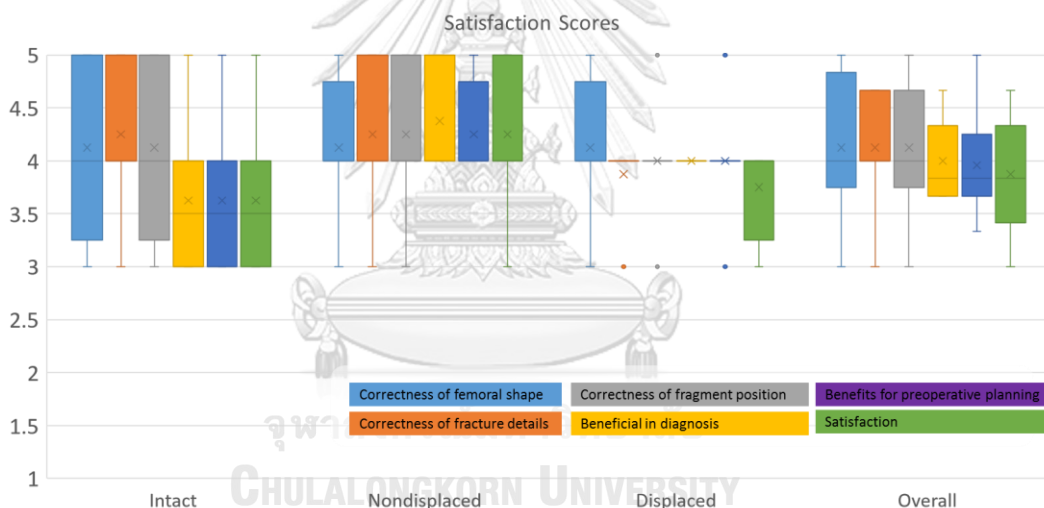
หลังจากให้ผู้เชี่ยวชาญทำการทดลองวินิจฉัยเพื่อจำแนกประเภทของการแตกหักของกระดูกต้นขา จะพบว่าความถูกต้องของการวินิจฉัย (Diagnosis scores) สำหรับกลุ่มตัวอย่าง Intact, Nondisplaced และ Displaced samples จะอยู่ที่ 89.29%, 78.57% และ 79.02% ตามลำดับ และมีความถูกต้องโดยรวมอยู่ที่ 71.88% ตามเกณฑ์การจำแนกประเภทหลัก (Main-categories) แสดงดังกราฟแท่งสีน้ำเงินแสดงดังรูปที่ 62 หากพิจารณาตามเกณฑ์การจำแนกประเภทย่อย (Sub-categories) จะได้คะแนนความถูกต้องของการวินิจฉัยของแต่ละกลุ่มลดน้อยลงมาอยู่ที่ 84.38% (Intact), 62.05% (Nondisplaced), 64.29% (Displaced) และ 70.24% (Overall) แสดงดังกราฟแท่งสีส้มแสดงดังรูปที่ 62 จะเห็นว่ามี ความแตกต่างของผลลัพธ์ระหว่างการจำแนกประเภทหลักกับการจำแนกประเภทย่อย เนื่องจากการจำแนกประเภทหลักสามารถทำได้ง่ายกว่า เพราะจำแนกโดยการระบุตำแหน่งหรือบริเวณที่มีการแตกหักของกระดูกเท่านั้น ส่วนจำแนกประเภทย่อยจะต้องพิจารณาถึงชนิดหรือรูปแบบของการแตกหักร่วมด้วยซึ่งทำได้ยากกว่า ซึ่งในประเด็นนี้เราจะทำการศึกษาและพัฒนาอัลกอริทึมต่อไปในอนาคต เพื่อให้สามารถใช้ในการจำแนกประเภทการแตกหักอย่างละเอียดได้ดียิ่งขึ้น



รูปที่ 62 ความถูกต้องของการวินิจฉัยเพื่อจำแนกประเภทการแตกหักของกระดูกต้นขา

## 2) ผลจากการประเมินความพึงพอใจแบบกลุ่ม

ในการประเมินแบบกลุ่ม เราได้แบ่งตัวอย่างออกเป็น 3 กลุ่ม (Group I: Intact, Group I: Nondisplaced และ Group III: Displaced samples) จากความเห็นของผู้เชี่ยวชาญ (ตามรูปที่ 63) คะแนนความถูกต้องของรูปทรงโดยรวมที่  $4.13 \pm 0.67S.D.$ , คะแนนมีความถูกต้องของรอยแตกหักโดยรวมที่  $4.13 \pm 0.56S.D.$ , คะแนนความถูกต้องของตำแหน่งการวางตัวของกระดูกที่แตกหักละชิ้น  $4.13 \pm 0.64S.D.$ , คะแนนความสามารถในการนำไปใช้ประโยชน์ต่อการวินิจฉัยการแตกหักของกระดูกต้นขา  $4.00 \pm 0.4S.D.$ , คะแนนความสามารถในการนำไปใช้ประโยชน์ในการวางแผนการผ่าตัดของกระดูกต้นขา  $3.96 \pm 0.52S.D.$ , และได้คะแนนความพึงพอใจโดยรวมที่  $3.88 \pm 0.56S.D.$  ซึ่งโดยภาพรวมแล้วสามารถสรุปได้ว่าว่าอัลกอริทึมนี้มีประโยชน์ในทางคลินิกและความพึงพอใจในการทำงานที่ประเมินโดยคุณหมอมือเชี่ยวชาญอยู่ในเกณฑ์ที่ดี



รูปที่ 63 ผลลัพธ์จากการประเมินเชิงคุณภาพของโมเดลสร้างรูปทรงสามมิติ

หากพิจารณาเป็นรายกลุ่มตัวอย่าง จะพบว่าตัวอย่างกลุ่มที่ 1 (Intact samples) ได้รับคะแนนเชิงคุณภาพตามเกณฑ์ I-III ใกล้เคียงกัน ส่วนคะแนนเชิงคุณภาพตามเกณฑ์ IV-VI มีค่าน้อยกว่าอาจเนื่องมาจากกลุ่มตัวอย่างนี้ไม่มีรอยแตกหักใดๆ ปรากฏให้เห็น อาจทำให้แพทย์สงสัยว่ารอยแตกหักไม่มีอยู่จริงๆ ใช่หรือไม่

ส่วนกลุ่มตัวอย่างกลุ่มที่ 2 (Nondisplaced samples) จะได้รับคะแนนเชิงคุณภาพตามเกณฑ์ IV-VI ที่สูงกว่ากลุ่มอื่นๆ อาจเนื่องมาจากกลุ่มตัวอย่างที่แตกหักแบบไม่เคลื่อน สามารถวิเคราะห์ได้ยากหากให้วินิจฉัยจากภาพรังสีสองมิติ การมีรูปทรงสามมิติ

ประกอบการตัดสินใจจะเป็นประโยชน์ต่อการวินิจฉัยการแตกหักของกระดูกในกลุ่มนี้มากยิ่งขึ้น

ส่วนกลุ่มตัวอย่างกลุ่มที่ 3 (Displaced samples) จะได้รับคะแนนเชิงคุณภาพตามเกณฑ์ I, III-V อยู่ในเกณฑ์ที่ดีใกล้เคียงกัน ส่วนคะแนนตามเกณฑ์ VI จะทำได้น้อยกว่า อาจเนื่องมาจากรูปทรงของผลัฟท์ยังมีความแม่นยำโดยรวมอยู่ในระดับปานกลาง และยังมีรอยแตกหักที่ไม่ชัดเจนมากนัก ซึ่งสอดคล้องกับผลัฟท์จากการประเมินเชิงปริมาณในหัวข้อที่ 4.3 และ 5.3

สรุปจากการประเมินเชิงคุณภาพ จะพบว่าอัลกอริทึมที่พัฒนาขึ้นสามารถช่วยให้คุณหมอผู้เชี่ยวชาญวินิจฉัยการแตกหักและวางแผนการผ่าตัดได้ในระดับที่ดี ซึ่งเป็นที่ยืนยันถึงประสิทธิภาพของอัลกอริทึมที่น่าเสนอนี้ ซึ่งผู้วิจัยคิดว่าอัลกอริทึมนี้ยังต้องถูกปรับปรุงเรื่องของความถูกต้องแม่นยำให้มากขึ้นไปอีกในอนาคต

## บทที่ 7

### สรุปผลการวิจัยและข้อเสนอแนะงานวิจัยในอนาคต

#### 7.1 สรุปผลการวิจัย

ในงานวิจัยนี้ ผู้วิจัยได้ทำการพัฒนาอัลกอริทึมหรือโมเดลการสร้างรูปทรงสามมิติ จากภาพถ่ายรังสีสองมุมมองของกระดูกต้นขาทั้งแบบกระดูกปกติ, แดกหักแบบไม่เคลื่อนที่ และ แดกหักแบบเคลื่อนที่ได้สำเร็จ ด้วยการคิดค้นโมเดลสร้างรูปทรงสามมิติ หรือ FracReconNet-Axial ซึ่งเป็นโมเดลการเรียนรู้ของเครื่องแบบ FPN (เป็น CNN ชนิดหนึ่ง) ประกอบกับเทคนิคคลาส ตัวแทนของบริเวณที่แตกหักของกระดูก และเทคนิคการเพิ่มจำนวนตัวอย่างสอนของกระดูกที่มี รอยการแตกหัก ช่วยให้โมเดลสามารถสร้างรูปทรงสามมิติที่มีรายละเอียดการแตกหักได้อย่างมีประสิทธิภาพ อีกทั้งยังได้คิดค้น Axial-fusion module เสริมความสามารถให้โมเดลสามารถ รวบรวมข้อมูลจากภาพรังสีแต่ละมุมมองที่มีความคลาดเคลื่อนเชิงมุมได้ถึง 10 องศา ซึ่งช่วยลด ปัญหาที่อาจเกิดขึ้นจากการจัดมุมมองการถ่ายภาพรังสี จากการทดสอบความแม่นยำของผลลัพธ์ ที่ได้จะมีค่า mIoU เท่ากับ  $0.827 \pm 0.083SD$ . และ mASSD เท่ากับ  $1.043 \pm 0.481SD$ . จากการ ประเมินเชิงคุณภาพหรือทดสอบการใช้งานเบื้องต้น พบว่าอัลกอริทึมที่นำเสนอช่วยให้คุณหมอ ผู้เชี่ยวชาญสามารถวินิจฉัยเพื่อจำแนกประเภทของการแตกหักของกระดูกต้นขาได้เป็นอย่างดี อย่างไรก็ตามผู้ใช้งานควรควบคุมความคลาดเคลื่อนเชิงมุมให้น้อยกว่า 2-3 องศา เพื่อให้ได้ ผลลัพธ์ที่ดีกว่า

ประโยชน์ที่คาดว่าจะได้รับ ถึงแม้ว่าในปัจจุบันอัลกอริทึมยังมีความแม่นยำในการสร้าง รูปทรงสามมิติไม่สูงนัก แต่ก็ยังเป็นประโยชน์ต่อการนำไปประยุกต์ใช้ในการเรียนการสอนของนิสิต แพทย์ และการประยุกต์ใช้ในการสื่อสารระหว่างคุณหมอกับผู้ป่วยหรือญาติผู้ป่วยเพื่อให้เข้าใจถึง สภาวะกระดูกแตกหักที่เกิดขึ้นกับตัวผู้ป่วย อีกทั้งยังสามารถนำอัลกอริทึมที่พัฒนาขึ้นไปใช้เป็น โมเดลต้นแบบสำหรับนำไปต่อยอดเพื่อพัฒนาอัลกอริทึมเพื่อสร้างรูปทรงสามมิติของกระดูกต้นขา ที่แตกหัก ที่สามารถใช้ในงานที่ซับซ้อนยิ่งขึ้นและมีความแม่นยำที่สูง และในอนาคตหาก อัลกอริทึมได้รับการฝึกสอนด้วยชุดข้อมูลที่มีขนาดใหญ่และมีความหลากหลายจนมีประสิทธิภาพ ที่สูงขึ้น ผู้วิจัยคาดหวังว่าจะสามารถนำอัลกอริทึมไปใช้ประกอบการวิเคราะห์สภาวะกระดูกต้นขา ที่แตกหัก เพื่อสนับสนุนการวินิจฉัยและการวางแผนการผ่าตัดของแพทย์ให้มีประสิทธิภาพยิ่งขึ้น



## 7.2 ข้อเสนอแนะงานวิจัยในอนาคต

เนื่องจากงานวิจัยนี้ยังมีข้อจำกัดในการทำวิจัยอยู่มาก เช่น จำนวนตัวอย่างสำหรับฝึกสอนที่มีไม่มากพอ, การใช้ภาพถ่ายรังสีสังเคราะห์ในการฝึกสอนโมเดลการเรียนรู้เท่านั้น, การศึกษาวิจัยและเก็บข้อมูลการเก็บข้อมูลภาพถ่ายรังสีแบบย้อนหลัง (Retrospective study) จึงไม่สามารถกำหนดขั้นตอนการเก็บข้อมูลภาพ X-ray และ CT ในมุมมองและตำแหน่งที่สอดคล้องกันได้ อีกทั้งยังมีข้อจำกัดในเรื่องของเวลาและงบประมาณสำหรับการเก็บข้อมูล ดังนั้นเราจึงขอเสนอแนะแนวทางการทำวิจัยในอนาคต ดังต่อไปนี้

1. การศึกษาไปข้างหน้า (Prospective study) เพื่อให้สามารถกำหนดขั้นตอนการเก็บข้อมูลภาพ X-ray หรือ Fluoroscopy และ CT ที่มีมุมมองและตำแหน่งการถ่ายภาพรังสีที่สอดคล้องกัน การทำเช่นนี้จะสามารถกำหนดหรือควบคุมตัวแปรการถ่ายภาพรังสี เช่น view positioning, patient arrangement, radiation dose (eV) เป็นต้น ให้เป็นมาตรฐานเดียวกันหรือให้มีความหลากหลายได้ ซึ่งสามารถนำไปใช้ประโยชน์ในการฝึกสอนโมเดลการสร้างรูปทรงสามมิติที่มีความแม่นยำมากยิ่งขึ้น โดยงานวิจัยเรื่อง X2CT-GAN ให้คำแนะนำไว้ว่าจำเป็นจะต้องใช้ตัวอย่างฝึกสอนประมาณ 1000-2000 ตัวอย่างเพื่อให้ได้ผลลัพธ์ที่แม่นยำเพียงพอที่จะนำไปใช้งานจริงได้
2. การเรียนรู้โดยการถ่ายโอนโมเดล (Transfer learning) เป็นการนำโมเดลที่ถูกฝึกสอนไว้แล้ว เช่น FracReconNet-Axial หรือ Pre-trained model บนโดเมนภาพถ่ายรังสีสังเคราะห์มาทำการฝึกสอนใหม่บนโดเมนภาพ X-ray, fluoroscopy หรืออื่นๆ ให้โมเดลเกิด generalization เพื่อให้สามารถจำไปให้ได้ดีกับสถานการณ์จริง
3. การทำปรับแต่ง Hyper parameters ต่างๆ ของโมเดล เช่น ขนาด filter ของ Convolution, ขนาดคุณสมบัติขาออกของ Convolution แต่ละส่วน, จำนวนชั้น ของ Convolution กับ Normalization layer และระดับความชันหรือความละเอียดของการประมวลผล เป็นต้น เพื่อเพิ่มประสิทธิภาพ (Maximization) ในการสร้างรูปทรงสามมิติขั้นสุด
4. การสร้างรูปทรงสามมิติแบบแยกส่วน (3D instance reconstruction) เป็นการสร้างรูปทรงสามมิติของแต่ละชิ้นส่วนของกระดูกหรืออวัยวะที่สนใจแต่ละชิ้นส่วนแยกเป็นคนละวัตถุกัน ซึ่งแต่ละวัตถุจะเป็น label ที่แตกต่างกันในปริภูมิสามมิติ เพื่อให้ผู้ใช้งานสามารถแยกส่วน, เคลื่อนย้าย, นับจำนวน และวิเคราะห์หรือกระดูกแต่ละชิ้นได้ ซึ่งจะเป็นประโยชน์ต่อการวางแผนการผ่าตัดอย่างมาก

\*หมายเหตุ ในงานวิจัยนี้จะเป็นการสร้างรูปทรงสามมิติของกระดูกต้นขาทั้งหมดพร้อมๆ กัน (3D semantic reconstruction) โดยแต่ละชิ้นส่วนจะร่วมเป็นวัตถุเดียวกันในปริภูมิสามมิติ ไม่สามารถแยกกันได้

5. การทดลองทางคลินิก (Clinical trial) เป็นการนำอัลกอริทึมดังกล่าวไปทดลองใช้ในขั้นตอนการวินิจฉัยและวางแผนการรักษา ซึ่งพิจารณาในเรื่องของผลลัพธ์ทางคลินิก (Clinical outcome) เช่น ระยะเวลาที่ใช้ในการผ่าตัด, ปริมาณเลือดที่ผู้ป่วยเสียไประหว่างผ่าตัด, จำนวนภาพถ่ายรังสีระหว่างผ่าตัด เป็นต้น



## ภาคผนวก A

## ผลลัพธ์การสร้างรูปทรงสามมิติ วัดผลด้วย IoU metric

ตารางที่ 11 แสดงผลลัพธ์การสร้างรูปทรงสามมิติด้วยโมเดล FracReconNet  
ซึ่งถูกฝึกสอนด้วยภาพถ่ายรังสีสองมุมมองที่ตั้งฉากกัน (aligned data)

		Intersection-over-Union (IoU)			
		Intact	Nondisplaced	Displaced	Overall
Diff. angles	Sample Types				
	Baseline (FracReconNet)	0.885±0.036	0.872±0.014	0.770±0.081	0.846±0.074
	0.0	0.885±0.036	0.872±0.017	0.770±0.080	0.846±0.074
	2.5	0.879±0.038	0.868±0.018	0.759±0.084	0.838±0.077
	5.0	0.865±0.044	0.857±0.023	0.740±0.093	0.823±0.084
	7.5	0.847±0.052	0.840±0.032	0.715±0.100	0.803±0.092
	10.0	0.824±0.062	0.823±0.039	0.691±0.107	0.781±0.098

ตารางที่ 12 แสดงผลลัพธ์การสร้างรูปทรงสามมิติด้วยโมเดล FracReconNet  
ซึ่งถูกฝึกสอนด้วยภาพถ่ายรังสีสองมุมมองที่มีความคลาดเคลื่อนเชิงมุม (misaligned data)

		Intersection-over-Union (IoU)			
		Intact	Nondisplaced	Displaced	Overall
Diff. angles	Sample Types				
	Baseline (FracReconNet)	0.885±0.036	0.872±0.014	0.770±0.081	0.846±0.074
	0.0	0.859±0.043	0.816±0.015	0.726±0.098	0.807±0.088
	2.5	0.858±0.044	0.818±0.017	0.722±0.097	0.806±0.088
	5.0	0.858±0.044	0.818±0.020	0.721±0.100	0.806±0.089
	7.5	0.856±0.046	0.816±0.024	0.719±0.102	0.804±0.090
	10.0	0.854±0.048	0.812±0.028	0.714±0.107	0.801±0.094

ตารางที่ 13 แสดงผลลัพธ์การสร้างรูปทรงสามมิติด้วยโมเดล *FracReconNet-Axial* ซึ่งถูกฝึกสอนด้วยภาพถ่ายรังสีสองมุมมองที่มีความคลาดเคลื่อนเชิงมุม (*misaligned data*)

		Intersection-over-Union (IoU)			
		Intact	Nondisplaced	Displaced	Overall
Diff. angles	Sample Types				
	Baseline (FracReconNet)	0.885±0.036	0.872±0.014	0.770±0.081	0.846±0.074
	0.0	0.873±0.045	0.854±0.018	0.753±0.087	0.830±0.081
	2.5	0.872±0.047	0.855±0.017	0.750±0.084	0.829±0.080
	5.0	0.872±0.048	0.854±0.018	0.748±0.088	0.828±0.082
	7.5	0.871±0.049	0.851±0.020	0.744±0.092	0.826±0.085
	10.0	0.868±0.052	0.848±0.022	0.738±0.097	0.822±0.089

## ภาคผนวก B

## ผลลัพธ์การสร้างรูปทรงสามมิติ วัดผลด้วย ASSD metric

ตารางที่ 14 แสดงผลลัพธ์การสร้างรูปทรงสามมิติด้วยโมเดล FracReconNet  
ซึ่งถูกฝึกสอนด้วยภาพถ่ายรังสีสองมุมมองที่ตั้งฉากกัน (aligned data)

**Mean Average Symmetric Surface Distance (mASSD) (mm)**

Diff. angles	Sample Types	Intact	Nondisplaced	Displaced	Overall
	<b>Baseline (FracReconNet)</b>		0.729±0.288	0.748±0.189	1.241±0.370
0.0		0.782±0.319	0.778±0.175	1.255±0.399	0.934±0.393
2.5		0.820±0.323	0.808±0.181	1.329±0.412	0.982±0.410
5.0		0.919±0.376	0.886±0.214	1.442±0.467	1.082±0.457
7.5		1.065±0.454	1.013±0.293	1.614±0.552	1.232±0.532
10.0		1.267±0.567	1.139±0.361	1.786±0.610	1.410±0.606

ตารางที่ 15 แสดงผลลัพธ์การสร้างรูปทรงสามมิติด้วยโมเดล FracReconNet  
ซึ่งถูกฝึกสอนด้วยภาพถ่ายรังสีสองมุมมองที่มีความคลาดเคลื่อนเชิงมุม (misaligned data)

**Mean Average Symmetric Surface Distance (mASSD) (mm)**

Diff. angles	Sample Types	Intact	Nondisplaced	Displaced	Overall
	<b>Baseline (FracReconNet)</b>		0.729±0.288	0.748±0.189	1.241±0.370
0.0		0.973±0.377	1.100±0.185	1.537±0.455	1.184±0.457
2.5		0.980±0.393	1.098±0.194	1.573±0.463	1.194±0.470
5.0		0.988±0.397	1.099±0.192	1.577±0.476	1.200±0.475
7.5		1.005±0.415	1.115±0.210	1.596±0.495	1.217±0.490
10.0		1.034±0.438	1.144±0.224	1.629±0.539	1.247±0.516

ตารางที่ 16 แสดงผลลัพธ์การสร้างรูปทรงสามมิติด้วยโมเดล *FracReconNet-Axial* ซึ่งถูกฝึกสอนด้วยภาพถ่ายรังสีสองมุมมองที่มีความคลาดเคลื่อนเชิงมุม (*misaligned data*)

**Mean Average Symmetric Surface Distance (mASSD) (mm)**

Sample Types Diff. angles	Intact	Nondisplaced	Displaced	Overall
<b>Baseline (FracReconNet)</b>	0.729±0.288	0.748±0.189	1.241±0.370	0.895±0.381
0.0	0.865±0.402	0.854±0.189	1.338±0.457	1.018±0.450
2.5	0.875±0.420	0.857±0.193	1.366±0.448	1.030±0.459
5.0	0.875±0.423	0.868±0.193	1.375±0.474	1.035±0.470
7.5	0.879±0.439	0.890±0.212	1.408±0.510	1.052±0.495
10.0	0.900±0.466	0.912±0.225	1.440±0.548	1.076±0.523

## ภาคผนวก C

ผลลัพธ์จากการสร้างรูปทรงสามมิติแบบ 3D Volume และ Surface distance error

Input					
Ground-truth					
FracReconNet (Align train)	 IoU = 0.888, ASSD = 0.776 mm	 IoU = 0.887, ASSD = 0.723 mm	 IoU = 0.878, ASSD = 0.690 mm	 IoU = 0.696, ASSD = 1.654 mm	 IoU = 0.778, ASSD = 1.141 mm
FracReconNet (misalign train)	 IoU = 0.874, ASSD = 0.854 mm	 IoU = 0.828, ASSD = 1.106 mm	 IoU = 0.830, ASSD = 0.949 mm	 IoU = 0.592, ASSD = 2.224 mm	 IoU = 0.725, ASSD = 1.413 mm
FracReconNet-Axial (misalign train)	 IoU = 0.881, ASSD = 0.797 mm	 IoU = 0.845, ASSD = 0.991 mm	 IoU = 0.851, ASSD = 0.795 mm	 IoU = 0.683, ASSD = 1.605 mm	 IoU = 0.727, ASSD = 1.462 mm

รูปที่ 64 แสดงผลลัพธ์แบบ 3D volume จากการสร้างรูปทรงสามมิติของโมเดลต่างๆ จากภาพถ่ายรังสีสองมุมมองที่ตั้งฉากกัน

Input					
Ground-truth					
FracReconNet (Align train)	 IoU = 0.859, ASSD = 0.957 mm	 IoU = 0.845, ASSD = 0.990 mm	 IoU = 0.871, ASSD = 0.771 mm	 IoU = 0.627, ASSD = 2.268 mm	 IoU = 0.728, ASSD = 1.470 mm
FracReconNet (misalign train)	 IoU = 0.871, ASSD = 0.878 mm	 IoU = 0.852, ASSD = 0.951 mm	 IoU = 0.832, ASSD = 0.940 mm	 IoU = 0.591, ASSD = 2.343 mm	 IoU = 0.730, ASSD = 1.423 mm
FracReconNet-Axial (misalign train)	 IoU = 0.877, ASSD = 0.826 mm	 IoU = 0.877, ASSD = 0.781 mm	 IoU = 0.849, ASSD = 0.821 mm	 IoU = 0.687, ASSD = 1.598 mm	 IoU = 0.742, ASSD = 1.455 mm

รูปที่ 65 แสดงผลลัพธ์แบบ 3D volume จากการสร้างรูปทรงสามมิติของโมเดลต่างๆ จากภาพถ่ายรังสีสองมุมมองที่มีความคลาดเคลื่อนเชิงมุม 5 องศา

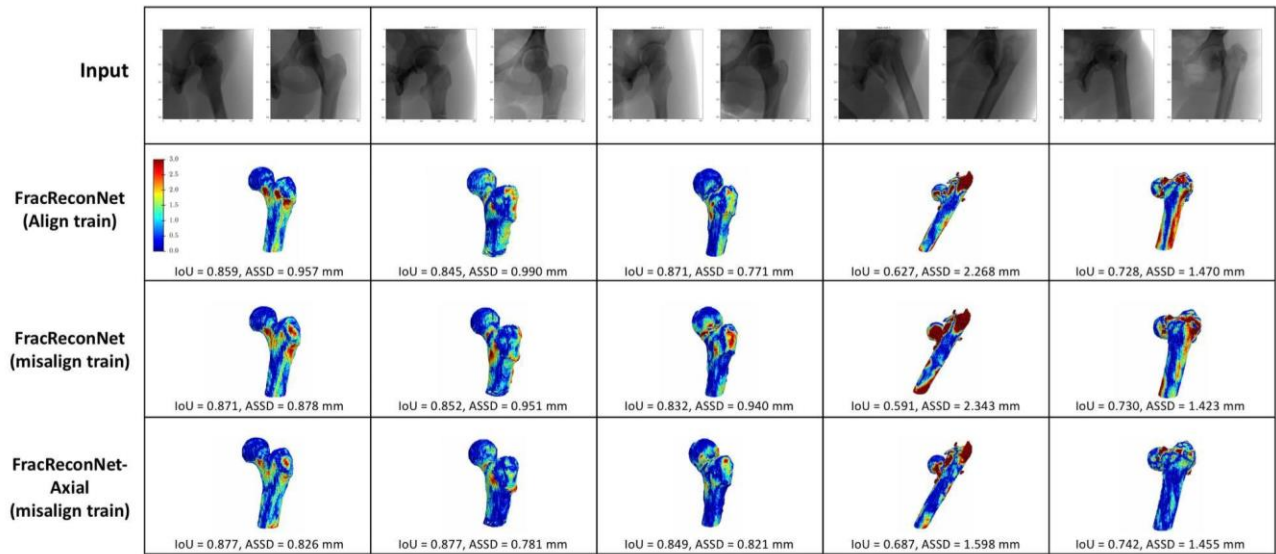
<b>Input</b>					
<b>Ground-truth</b>					
<b>FracReconNet (Align train)</b>	 IoU = 0.821, ASSD = 1.243 mm	 IoU = 0.818, ASSD = 1.188 mm	 IoU = 0.858, ASSD = 0.827 mm	 IoU = 0.548, ASSD = 3.015 mm	 IoU = 0.661, ASSD = 1.900 mm
<b>FracReconNet (misalign train)</b>	 IoU = 0.870, ASSD = 0.886 mm	 IoU = 0.847, ASSD = 0.987 mm	 IoU = 0.835, ASSD = 0.922 mm	 IoU = 0.578, ASSD = 2.560 mm	 IoU = 0.718, ASSD = 1.501 mm
<b>FracReconNet-Axial (misalign train)</b>	 IoU = 0.871, ASSD = 0.863 mm	 IoU = 0.878, ASSD = 0.773 mm	 IoU = 0.846, ASSD = 0.826 mm	 IoU = 0.681, ASSD = 1.622 mm	 IoU = 0.737, ASSD = 1.379 mm

รูปที่ 66 แสดงผลลัพธ์แบบ 3D volume จากการสร้างรูปทรงสามมิติของโมเดลต่างๆ จากภาพถ่ายรังสีสองมุมมองที่มีความคลาดเคลื่อนเชิงมุม 10 องศา

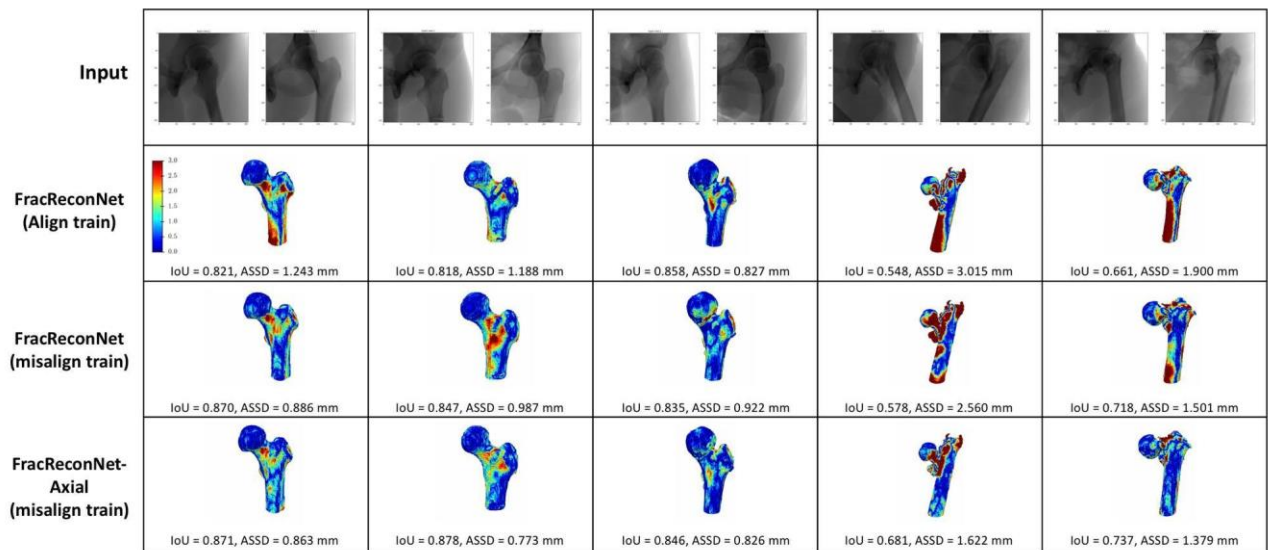
<b>Input</b>					
<b>FracReconNet (Align train)</b>	 IoU = 0.888, ASSD = 0.776 mm	 IoU = 0.887, ASSD = 0.723 mm	 IoU = 0.878, ASSD = 0.690 mm	 IoU = 0.696, ASSD = 1.654 mm	 IoU = 0.778, ASSD = 1.141 mm
<b>FracReconNet (misalign train)</b>	 IoU = 0.874, ASSD = 0.854 mm	 IoU = 0.828, ASSD = 1.106 mm	 IoU = 0.830, ASSD = 0.949 mm	 IoU = 0.592, ASSD = 2.224 mm	 IoU = 0.725, ASSD = 1.413 mm
<b>FracReconNet-Axial (misalign train)</b>	 IoU = 0.881, ASSD = 0.797 mm	 IoU = 0.845, ASSD = 0.991 mm	 IoU = 0.851, ASSD = 0.795 mm	 IoU = 0.683, ASSD = 1.605 mm	 IoU = 0.727, ASSD = 1.462 mm

รูปที่ 67 แสดงผลลัพธ์แบบ Surface distance error จากการสร้างรูปทรงสามมิติของโมเดลต่างๆ จากภาพถ่ายรังสีสองมุมมองที่ตั้งฉากกัน





รูปที่ 68 แสดงผลลัพธ์แบบ Surface distance error จากการสร้างรูปทรงสามมิติของโมเดลต่างๆ จากภาพถ่ายรังสีสองมุมมองที่มีความคลาดเคลื่อนเชิงมุม 5 องศา



รูปที่ 69 แสดงผลลัพธ์แบบ Surface distance error จากการสร้างรูปทรงสามมิติของโมเดลต่างๆ จากภาพถ่ายรังสีสองมุมมองที่มีความคลาดเคลื่อนเชิงมุม 10 องศา

## ภาคผนวก D

## เอกสารจริยธรรมงานวิจัยในมนุษย์



COA No. 541/2021

IRB No. 249/64

## INSTITUTIONAL REVIEW BOARD

Faculty of Medicine, Chulalongkorn University

1873 Rama 4 Road, Pathumwan, Bangkok 10330, Thailand, Tel 662-256-4493

## Certificate of Approval

The Institutional Review Board of the Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand, has approved the following study which is to be carried out in compliance with the International guidelines for human research protection as Declaration of Helsinki, The Belmont Report, CIOMS Guideline and International Conference on Harmonization in Good Clinical Practice (ICH-GCP)

Study Title	: Two-View 3D Reconstruction of Proximal Femoral Fracture Using Deep Learning
Study Code	: -
Principal Investigator	: Mr. Danupong Buttongkum
Affiliation of PI	: Faculty of Engineering, Chulalongkorn University.
Review Method	: Expedited
Continuing Report	: At least once annually or submit the final report if finished.
Document Reviewed	: <ol style="list-style-type: none"> <li>1. Research Proposal Version 02 Date 05 April 2021</li> <li>2. Protocol Synopsis Version 02 Date 04 April 2021</li> <li>3. Curriculum Vitae and GCP Training           <ul style="list-style-type: none"> <li>- Mr. Danupong Buttongkum</li> <li>- Asst.Prof. Pairat Tangpompasert, Ph.D.</li> <li>- Asst.Prof. Chanyaphan Virulsri, Ph.D.</li> </ul> </li> </ol>

Approval granted is subject to the following conditions: (see back of this Certificate)



- Asst.Prof. Numphung Numkarunarunrote, M.D.
- Chavarin Amarase, M.D.

Signature ..... *Tada Sueblinong* .....  
(Emeritus Professor Tada Sueblinong MD)  
Chairperson  
The Institutional Review Board

Signature ..... *Onanong Kulaputana* .....  
(Associate Professor Onanong Kulaputana MD, PhD)  
Member and Assistant Secretary, Acting Secretary  
The Institutional Review Board

Date of Approval : April 21, 2021  
Approval Expire Date : April 20, 2022

Approval granted is subject to the following conditions: (see back of this Certificate)

All approved investigators must comply with the following conditions:

1. Strictly conduct the research as required by the protocol;
2. Use only the information sheet, consent form (and recruitment materials, if any), interview outlines and/or questionnaires bearing the Institutional Review Board's seal of approval ; and return one copy of such documents of the first subject recruited to the Institutional Review Board (IRB) for the record;
3. Report to the Institutional Review Board any serious adverse event or any changes in the research activity within five working days;
4. Provide reports to the Institutional Review Board concerning the progress of the research upon the specified period of time or when requested;
5. If the study cannot be finished within the expire date of the approval certificate, the investigator is obliged to reapply for approval at least one month before the date of expiration.
6. If the research project is completed, the researcher must be form the Faculty of Medicine, Chulalongkorn University.

\* A list of the Institutional Review Board members (names and positions and expertises) present at the meeting of Institutional Review Board on the date of approval of this study has been attached. All approved documents will be forwarded to the principal investigator.

## บรรณานุกรม

- [1] S. J. A. S. Amin, E. J. Atkinson, S. Khosla and L. J. Melton, "Trends in Fracture Incidence: A Population-Based Study Over 20 Years," *J Bone Miner Res.*, vol. 29, pp. 581-589, 2015.
- [2] N. T. Khunying Kobchitt Limpaphayom, Unnop Jaisamrarn, "Prevalence of osteopenia and osteoporosis in Thai women," *The Journal of The North American Menopause Society*, vol. 8, pp. 65-69, 2001.
- [3] S. L. Tanawat Vaseenon, Prasit Wongtriratanachai and Sattaya Rojanasthien, "Long-Term Mortality After Osteoporotic Hip Fracture in Chiang Mai, Thailand," *Journal of Clinical Densitometry: Assessment of Skeletal Health*, vol. 13, pp. 63-67, 2010.
- [4] B. W. Dietmar Krappinger, Dietmar Dammerer, Martin Thaler, Peter Schwendinger, Richard A. Lindtner, "Risk factors for nonunion after intramedullary nailing of subtrochanteric femoral fractures," *Archives of Orthopaedic and Trauma Surgery*, vol. 139, pp. 769-777, 2019.
- [5] J. Y. S. Scott E. Sheehan, Michael J. Weaver, Aaron D. Sodickson, Bharti Khurana, "Proximal Femoral Fractures: What the Orthopedic Surgeon Wants to Know," *Radiographics*, vol. 35, 17 July 2015.
- [6] S. Z. Y. Zhang, S. Wang, H. Zhang, W. Zhang, P. Liu, J. Ma, N. Pervaiz, J. Wang, "Long and short intramedullary nails for fixation of intertrochanteric femur fractures (OTA 31-A1, A2 and A3): A systematic review and meta-analysis," *Orthopaedics & Traumatology: Surgery & Research*, vol. 103, pp. 685-690, 2017.
- [7] H. K. Ji Wan Kim, Chang-wug Oh, Joon-Woo Kim, Oog-Jin Shon, Young-Soo Byun, Jung Jae Kim, Hyoung Keun Oh, Hiroaki Minehara, Kyu-Tae Hwang, Ki Chul Park, "Surgical outcomes of intramedullary nailing for diaphyseal atypical femur fractures: is it safe to modify a nail entry in bowed femur?," *Archives of Orthopaedic and Trauma Surgery*, vol. 137, pp. 1515-1522, 2017.

- [8] T. Okada *et al.*, "Computer-Assisted Preoperative Planning for Reduction of Proximal Femoral Fracture Using 3-D-CT Data," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 3, pp. 749-759, 2009.
- [9] I. Radiological Society of North America. Radiation Dose in X-Ray and CT Exams [Online]. Available: <https://www.radiologyinfo.org/en/info.cfm?pg=safety-xray>
- [10] A. Wilharm, I. Marintshev, G. O. Hofmann, and F. Gras, "2D-fluoroscopic based navigation for Gamma 3 nail insertion versus conventional procedure- a feasibility study," (in eng), *BMC musculoskeletal disorders*, vol. 14, pp. 74-74, 2013.
- [11] K. Z. Dongdong Wang, Minfei Qiang, Xiaoyang Jia and Yanxi Chen, "Computer-assisted preoperative planning improves the learning curve of PFNA-II in the treatment of intertrochanteric femoral fractures," *BMC Musculoskeletal Disorders*, 2020.
- [12] X. Jia, K. Zhang, M. Qiang, Y. Wu, and Y. Chen, "Association of Computer-Assisted Virtual Preoperative Planning With Postoperative Mortality and Complications in Older Patients With Intertrochanteric Hip Fracture," (in eng), *JAMA Netw Open*, vol. 3, no. 8, p. e205830, Aug 3 2020.
- [13] H. Kasban, M. El-Bendary, and D. H. Salama, "A Comparative Study of Medical Imaging Techniques," *International Journal of Information Science and Intelligent System*, vol. 4, no. 2, pp. 37-58, 2015.
- [14] M. N. Kulachote, "Hip Fracture-Dislocation and Femur Fracture," in *Fracture of proximal humerus, clinical evaluation, imaging and classification (new development)*, 2015.
- [15] D. V. Sonawane, "Classifications of Intertrochanteric fractures and their Clinical Importance," *Trauma International Review Article* vol. 1, no. 1, 2015.
- [16] G. Wypych, "1 - PHOTOPHYSICS," in *Handbook of Material Weathering (Fifth Edition)*, G. Wypych, Ed. Oxford: Elsevier, 2013, pp. 1-25.
- [17] H. B. A. Akara, U. Cevikc, F. Korkmazb, N.T. Okumusoglu, "Measurement of attenuati on coefficients for bone, muscle, fat and water at 140, 364 and 662 keV

- g-ray energies," *Journal of Quantitative Spectroscopy & Radiative Transfer*, vol. 102, pp. 203 – 211, 2006.
- [18] D. Wang, K. Zhang, M. Qiang, X. Jia, and Y. Chen, "Computer-assisted preoperative planning improves the learning curve of PFNA-II in the treatment of intertrochanteric femoral fractures," *BMC Musculoskeletal Disorders*, vol. 21, no. 1, p. 34, 2020/01/16 2020.
- [19] K.-L. D. G. a. G. Penn, "Radiation in Medicine: A Need for Regulatory Reform," p. 111.
- [20] C. J. F. Reyneke, M. Lüthi, V. Burdin, T. S. Douglas, T. Vetter, and T. E. M. Mutsvangwa, "Review of 2-D/3-D Reconstruction Using Statistical Shape and Intensity Models and X-Ray Image Synthesis: Toward a Unified Framework," *IEEE Reviews in Biomedical Engineering*, vol. 12, pp. 269-286, 2019.
- [21] T. Whitmarsh, L. Humbert, M. D. Craene, L. M. D. R. Barquero, and A. F. Frangi, "Reconstructing the 3D Shape and Bone Mineral Density Distribution of the Proximal Femur From Dual-Energy X-Ray Absorptiometry," *IEEE Transactions on Medical Imaging*, vol. 30, no. 12, pp. 2101-2114, 2011.
- [22] J. J. Fei-Fei Li, Serena Yeung. CS231n Convolutional Neural Networks for Visual Recognition [Online].
- [23] S. H. Hyeonwoo Noh, Bohyung Han, "Learning Deconvolution Network for Semantic Segmentation," *Computer Vision and Pattern Recognition*, 17 May 2015.
- [24] T. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature Pyramid Networks for Object Detection," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 936-944, 21-26 July 2017 2017.
- [25] A. Z. Karen Simonyan, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *ICLR 2020 Conference Paper*, 2015.
- [26] D. A. Wei Liu, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, Alexander C. Berg. SSD: Single Shot MultiBox Detector [Online]. Available: <https://arxiv.org/abs/1512.02325>

- [27] P. F. Olaf Ronneberger, Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation [Online]. Available: <https://arxiv.org/abs/1505.04597>
- [28] P. Henzler, V. Rasche, T. Ropinski, and T. Ritschel, "Single-image Tomography: 3D Volumes from 2D Cranial X-Rays," *Computer Graphics Forum*, vol. 37, no. 2, pp. 377-388, 22 May 2018 2018.
- [29] X. Ying, H. Guo, K. Ma, J. Wu, Z. Weng, and Y. Zheng, "X2CT-GAN: Reconstructing CT From Biplanar X-Rays With Generative Adversarial Networks," *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10611-10620, 15-20 June 2019 2019.
- [30] D. F. F. Rohit Girdhar, Mikel Rodriguez, Abhinav Gupta, "Learning a Predictable and Generative Vector Representation for Objects," presented at the European Conference on Computer Vision (ECCV), 2016.
- [31] B. Aubert, C. Vazquez, T. Cresson, S. Parent, and J. A. d. Guise, "Toward Automated 3D Spine Reconstruction from Biplanar Radiographs Using CNN for Statistical Spine Model Fitting," *IEEE Transactions on Medical Imaging*, vol. 38, no. 12, pp. 2796-2806, 2019.
- [32] G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2261-2269, 21-26 July 2017 2017.
- [33] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal Loss for Dense Object Detection," Available: <https://arxiv.org/abs/1708.02002>
- [34] J. L. B. Diederik P. Kingma, "Adam: A Method for Stochastic Optimization," *The 3rd International Conference for Learning Representations*, 2015.
- [35] scikit-image, "Marching cubes algorithm," ed. scikit-image, 2021.
- [36] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *Journal of Big Data*, vol. 6, no. 1, p. 60, 2019/07/06 2019.
- [37] M. M. Kanafi, "Radially averaged surface roughness/topography power spectrum (PSD)," ed. MATLAB Central File Exchange, 2021.



- [38] J. K. Handong Kim, Heekyung Yang, "A GAN-Based Face Rotation for Artistic Portraits," *Mathematics*, vol. 10, no. 3860, 2022.
- [39] S. Z. Rui Huang, Tianyu Li, Ran He, "Beyond Face Rotation: Global and Local Perception GAN for Photorealistic and Identity Preserving Frontal View Synthesis," *Computer Vision and Pattern Recognition*, vol. 2017, 2017.
- [40] W. J. Jinpyo Kim, Hyungmo Kim, Jaejin Lee, "CyCNN: A Rotation Invariant CNN using Polar Mapping and Cylindrical Convolution Layers," *Computer Vision and Pattern Recognition*, 2020.



บรรณานุกรม



จุฬาลงกรณ์มหาวิทยาลัย  
**CHULALONGKORN UNIVERSITY**



จุฬาลงกรณ์มหาวิทยาลัย  
**CHULALONGKORN UNIVERSITY**

## ประวัติผู้เขียน

ชื่อ-สกุล	นายดนุพงษ์ บุตรทองคำ
วัน เดือน ปี เกิด	18 กันยายน 2535
สถานที่เกิด	พระนครศรีอยุธยา ประเทศไทย
วุฒิการศึกษา	คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ที่อยู่ปัจจุบัน	131/7 ม.2 ต.บ้านเกาะ อ.พระนครศรีอยุธยา จ.พระนครศรีอยุธยา 13000



จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY