

การพัฒนาระบบวิเคราะห์รอยโรคผิวหนังจากภาพถ่ายทางการแพทย์โดยใช้ กระบวนการแบ่งส่วนขอบเขตรอยโรคและการจำแนกประเภทรอยโรคด้วย โครงข่ายประสาทเทียมเชิงลึก

Development of a Medical Image-Based Skin Lesion Analysis System Using Lesion Segmentation and Deep Learning Classification

สุรัชช ยุธธศิลป์ โชคดี คนอยู่ สุทธิภัทร ตั้งกุลมงคล
จิรายุส อรุณเดชาชัย¹

¹โรงเรียนวิทยาศาสตร์จุฬาภรณราชวิทยาลัย เลย์ อำเภอบางละมุง จังหวัดชลบุรี 42110

*E-mail: surachat04911@pcshsloei.ac.th

บทคัดย่อ

โครงงานนี้มีวัตถุประสงค์เพื่อพัฒนาระบบวิเคราะห์รอยโรคผิวหนังจากภาพถ่ายทางการแพทย์ โดยใช้ประยุกต์ใช้ปัญญาประดิษฐ์และการเรียนรู้เชิงลึก (Deep Learning) เพื่อช่วยระบุตำแหน่งและจำแนกประเภทของรอยโรคผิวหนัง ระบบประกอบด้วยสองกระบวนการหลัก ได้แก่ การแบ่งส่วนขอบเขตรอยโรค (Image Segmentation) ด้วยโครงข่าย U-Net และการจำแนกประเภทรอยโรค (Image Classification) ด้วยโครงข่าย ResNet โดยใช้ชุดข้อมูลมาตรฐาน ISIC 2018 ในกระบวนการแบ่งส่วนขอบเขตรอยโรค ประเมินประสิทธิภาพด้วยตัวชี้วัด Accuracy, Dice Coefficient, Dice Loss, Hausdorff Distance (HD95) และ Average Surface Distance (ASD) ผลการทดลองพบว่าโมเดล U-Net แยกขอบเขตรอยโรคออกจากผิวหนังได้อย่างมีประสิทธิภาพ โดยให้ค่า Dice Coefficient ในระดับสูง ขณะที่ค่า Dice Loss ลดลงอย่างต่อเนื่องเมื่อจำนวนรอบการฝึกเพิ่มขึ้น สะท้อนถึงการเรียนรู้ที่มีเสถียรภาพ ในกระบวนการจำแนกประเภทรอยโรค โมเดล ResNet จำแนกรอยโรคได้ 7 ประเภท ได้แก่ Melanoma (MEL), Melanocytic Nevus (NV), Basal Cell Carcinoma (BCC), Actinic Keratosis (AKIEC), Benign Keratosis-like Lesions (BKL), Dermatofibroma (DF) และ Vascular Lesions (VASC) ผลการวิเคราะห์ด้วย Confusion Matrix และ Heatmap แสดงให้เห็นว่าโมเดลสามารถโฟกัสไปยังบริเวณรอยโรคสำคัญได้อย่างเหมาะสมในหลายกรณี แม้ยังพบความสับสนระหว่างบางคลาสที่มีลักษณะใกล้เคียงกัน ผลการศึกษาโดยรวมชี้ว่าระบบที่พัฒนาขึ้นมีศักยภาพเป็นเครื่องมือช่วยคัดกรองรอยโรคผิวหนังเบื้องต้น และสามารถต่อยอดสู่ระบบสนับสนุนการตัดสินใจทางการแพทย์ (Clinical Decision Support System) หรือแอปพลิเคชันด้านสุขภาพดิจิทัลในอนาคต ทั้งนี้ระบบมีวัตถุประสงค์เพื่อสนับสนุนการคัดกรองเบื้องต้นเท่านั้น มิได้ใช้ทดแทนการวินิจฉัยของแพทย์ผู้เชี่ยวชาญ

คำสำคัญ: การวิเคราะห์รอยโรคผิวหนัง, การเรียนรู้เชิงลึก, การแบ่งส่วนภาพ, การจำแนกประเภทภาพ, U-Net, ResNet, ISIC 2018, ระบบสนับสนุนการตัดสินใจทางการแพทย์

1. บทนำ

โรคผิวหนังเป็นปัญหาสุขภาพที่พบได้บ่อยในประชากรทุกช่วงวัย โดยมีลักษณะอาการและระดับ ความรุนแรงที่แตกต่างกัน ตั้งแต่โรคผิวหนังที่ไม่รุนแรงไปจนถึงโรคร้ายแรง เช่น มะเร็งผิวหนัง ซึ่งการ วินิจฉัยที่ถูกต้องและรวดเร็วมีความสำคัญอย่างยิ่งต่อการรักษาและการลดความเสี่ยงต่อภาวะแทรกซ้อน

อย่างไรก็ตาม การวินิจฉัยโรคผิวหนังจ ว่าเป็นต้องอาศัยแพทย์ผู้เชี่ยวชาญ และยังมีข้อจ กัดในด้านจำนวน แพทย์ ระยะเวลา รอคิว และการเข้าถึงบริการทางการแพทย์ โดยเฉพาะในพื้นที่ห่างไกลหรือพื้นที่ที่มี ผู้ป่วยจำนวนมาก ในปัจจุบัน เทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligence: AI) โดยเฉพาะการเรียนรู้เชิงลึก (Deep Learning) ได้รับการ ามาใช้ในงานด้าน

การแพทย์อย่างแพร่หลาย หนึ่งในเทคนิคที่มีประสิทธิภาพสูงคือ Convolutional Neural Network (CNN) ซึ่งมีความสามารถในการเรียนรู้ลักษณะเชิงพื้นที่ของภาพ และสามารถนำมาประยุกต์ใช้ในการวิเคราะห์ภาพทางการแพทย์ได้อย่างมีประสิทธิภาพ ทั้งในด้านการแบ่ง ส่วนขอบเขตของอวัยวะหรือรอยโรค และการจำแนกประเภทของโรค โครงการนี้จึงมุ่งเน้นการพัฒนากระบวนการวิเคราะห์รอยโรคผิวหนังจากภาพถ่ายทางการแพทย์ โดยใช้ กระบวนการแบ่งส่วนขอบเขตของรอยโรค (Image Segmentation) เพื่อระบุตำแหน่งและขอบเขตของรอยโรคอย่างแม่นยำ และนำผลลัพธ์ที่ได้ไปใช้ในการจำแนกประเภทของรอยโรคผิวหนังด้วยโครงข่ายประสาทเทียมเชิงลึก (Deep Neural Networks) เพื่อช่วยเพิ่มความถูกต้องในการวิเคราะห์ข้อมูลภาพ ระบบที่พัฒนาขึ้นนี้มีเป้าหมายเพื่อเป็นเครื่องมือช่วยวิเคราะห์รอยโรคผิวหนังเบื้องต้น ซึ่งสามารถนำไปต่อยอดเป็นเว็บแอปพลิเคชันหรือระบบสนับสนุนการตัดสินใจทางการแพทย์ ช่วยลดภาระงานของแพทย์ เพิ่มโอกาสในการตรวจพบโรคในระยะเริ่มต้น และสนับสนุนการดูแลสุขภาพเชิงรุกของประชาชน

2. วิธีดำเนินการ

2.1 แหล่งข้อมูล

โครงการนี้ใช้ชุดข้อมูลภาพถ่ายรอยโรคผิวหนังและภาพมาสก์ (Segmentation Mask) สำหรับการฝึกและทดสอบโมเดลปัญญาประดิษฐ์ โดยข้อมูลประกอบด้วยภาพถ่ายทางการแพทย์ของรอยโรคผิวหนังจำนวนทั้งสิ้น 230 เคส ซึ่งแต่ละเคสประกอบด้วยภาพ ต้นฉบับ (Original Image) และภาพมาสก์ที่แสดงขอบเขตของรอยโรค (Ground Truth Mask) ชุดข้อมูลทั้งหมดถูกแบ่งออกเป็น 2 ส่วน ได้แก่ ชุดข้อมูลสำหรับการฝึกโมเดล (Training Set) ร้อยละ 80 ชุดข้อมูลสำหรับการทดสอบโมเดล (Test Set) ร้อยละ 20 การแบ่งข้อมูลในลักษณะนี้มีวัตถุประสงค์เพื่อให้โมเดลสามารถเรียนรู้จากข้อมูลได้อย่างเพียงพอ และสามารถประเมินประสิทธิภาพของโมเดลบนข้อมูลที่ไม่เคยเห็นมาก่อน

2.1.1 ที่มาของชุดข้อมูล (Dataset): ระบุว่าใช้ชุดข้อมูลอะไร มาจากแหล่งไหน

ชุดข้อมูลที่ใช้ในโครงการนี้ได้มาจาก ISIC Archive (International Skin Imaging Collaboration) ซึ่งเป็นฐานข้อมูลภาพถ่ายโรคผิวหนังขนาดใหญ่ระดับนานาชาติ ใช้สำหรับการวิจัยและการศึกษาโดยไม่เสียค่าใช้จ่ายชุดข้อมูลจาก ISIC Archive ประกอบด้วยภาพถ่ายรอยโรคผิวหนังหลายประเภทพร้อม ข้อมูลกำกับ (Label) และภาพมาสก์สำหรับงานแบ่งส่วนขอบเขตของรอยโรค (Segmentation) ซึ่งถูกจัดทำโดยผู้เชี่ยวชาญทางการแพทย์ ทำให้ข้อมูลมีความน่าเชื่อถือและเหมาะสมสำหรับการพัฒนาโมเดลทางการแพทย์

2.1.2 คำนวณขนาดตัวอย่างขั้นต่ำโดยใช้สูตร

การประมาณสัดส่วนของประชากร

การกำหนดขนาดตัวอย่างขั้นต่ำในโครงการนี้อ้างอิงจากสูตรการประมาณ สัดส่วนของประชากร โดยใช้ค่า Dice Similarity Coefficient (DSC) เป็นตัวแปรสำคัญ ซึ่งเป็นตัววัดความแม่นยำของการแบ่งส่วนขอบเขตภาพ

$$n = \frac{Z_{\alpha/2}^2 pq}{d^2}$$

ภาพที่ 1

สูตรการ

ประมาณสัดส่วนของประชากร

โดยกำหนดให้

p คือ สัดส่วนของลักษณะที่สนใจในประชากร ซึ่งในที่นี้หมายถึงค่า DSC

q = 1 - p

d คือ ค่าความคลาดเคลื่อนสัมบูรณ์ของสัดส่วนประชากร

โดยกำหนดค่าระดับนัยสำคัญ $\alpha = 0.05$ และค่าความคลาดเคลื่อนที่ยอมรับได้ (d) = 5%

อ้างอิงผลการศึกษาในงานวิจัยเรื่อง “Automatic prostate cancer boundary delineation” ขนาดตัวอย่างจึงถูกคำนวณตามรายละเอียดดังกล่าว

$$n = (1.96)^2 (0.87)(1-0.87) / (1.96)^2 (0.87)(1-0.87)$$

n = 230 images and masks

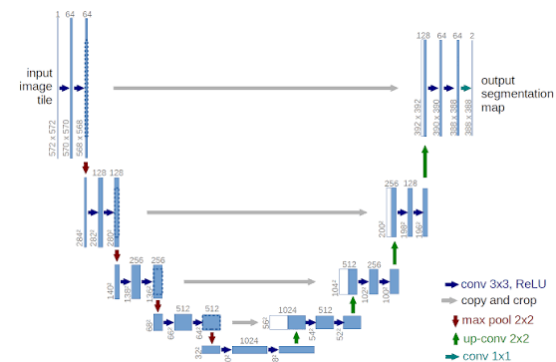
2.2 การสร้างและปรับปรุงข้อมูล (Preprocessing

& Augmentation)

ก่อนนำข้อมูลเข้าสู่กระบวนการฝึกโมเดล ภาพทั้งหมดถูกปรับปรุงคุณภาพเพื่อให้มีความเหมาะสมต่อการเรียนรู้ของโมเดล และลดผลกระทบจากความแตกต่างของแสง สี และสัญญาณรบกวนในภาพ

2.3 สถาปัตยกรรมโมเดล (Model Architecture)

รายละเอียด U-Net และโครงสร้างของ U-Net U-Net คือโมเดล ที่นิยมในการเอาไปใช้ในงาน Segmentation ทาง การแพทย์โดยการนำไป deep learning



ภาพที่ 2 โครงสร้าง U-Net

Encoder

ในสถาปัตยกรรม U-Net ที่ทำหน้าที่ดึงเอาลักษณะสำคัญ (features) ของภาพ ต้นฉบับออกมา โดยจะใช้ชั้นคอนโวลูชัน (Convolution) และการลดขนาดภาพ (Downsampling หรือ Max Pooling) ต่อเนื่องกันหลายครั้ง ผลลัพธ์ของ Encoder ในแต่ละชั้นคือ feature maps ที่บีบอัดขนาดลง แต่มีข้อมูลเชิงลึก (semantic information) มากขึ้น ทำให้โมเดลเข้าใจว่าในภาพมีวัตถุหรือขอบเขตแบบไหนบ้าง

Decoder

เป็นส่วนที่ทำหน้าที่ฟื้นฟูขนาดภาพกลับขึ้นมา จาก feature maps ที่ได้จาก Encoder โดยใช้การขยายภาพ (Upsampling หรือ Transposed Convolution) ที่ละชั้น และต่อ ชั้นคอนโวลูชัน เพื่อสร้างแผนที่ผลลัพธ์ (segmentation map) ในขั้นตอนที่โมเดลจะพยายามสร้างภาพใหม่ที่มีขนาดเท่ากับภาพต้นฉบับ แต่ในแต่ละตำแหน่ง ของพิกเซลจะมีค่า การคาดการณ์ว่าเป็นอวัยวะหรือไม่ (เช่น ค่าความน่าจะเป็นของต่อม ลูกหมาก)

Skip Connections

การเชื่อมต่อข้อมูลระหว่าง Encoder และ Decoder ในแต่ละระดับ โดยจะนำ feature maps จาก Encoder ชั้นนั้น มาต่อ (concatenate) เข้ากับ feature maps ของ Decoder ในระดับเดียวกัน การทำแบบนี้ช่วยให้ Decoder ไม่สูญเสียรายละเอียดเชิงพื้นที่ (spatial details) ระหว่างการ บีบอัดภาพ และทำให้ segmentation มีความละเอียดสูงขึ้น โดยเฉพาะขอบเขตของอวัยวะที่เล็ก หรือซับซ้อน

2.4 โมเดลการจำแนกประเภทรอยโรค

(Classification Model)

หลังจากได้ผลลัพธ์การแบ่งส่วนขอบเขตรอยโรคแล้ว ภาพบริเวณรอยโรคจะถูกนำเข้าสู่อัลกอริทึม ResNet เพื่อทำการจำแนกประเภทรอยโรคออกเป็น 7 ประเภท ได้แก่ MEL, NV, BCC, AKIEC, BKL, DF และ VASC โมเดล ResNet ใช้โครงสร้าง Residual Connections เพื่อลดปัญหา vanishing gradient และช่วยให้ โมเดลสามารถเรียนรู้ข้อมูลเชิงลึกได้อย่างมีประสิทธิภาพ

2.5 การเทรนโมเดล (Model Training)

ผัง Flowchart ของการนำ Model ไปเทรนนั้น การฝึกสอนโมเดลในโครงการนี้ใช้สถาปัตยกรรมแบบ U-Net ซึ่งเหมาะสำหรับงานจำแนกภาพ ทางแพทย์โดยเฉพาะ โดยเราได้นำกระบวนการฝึกสอนตามขั้นตอนต่อไปนี้

การเตรียมข้อมูล: ข้อมูล MRI ของต่อมลูกหมากถูกนำมาผ่านกระบวนการ preprocessing เช่น normalization และการปรับขนาด เพื่อให้เหมาะสมกับโมเดล

การแบ่งข้อมูล: แบ่งชุดข้อมูลออกเป็น 3 ส่วน ได้แก่ ชุดฝึกสอน (Training set), ชุดตรวจสอบ (Validation set) และชุดทดสอบ (Test set) เพื่อประเมินผลได้อย่างถูกต้อง

การฝึกสอนโมเดล: ใช้เทคนิคการปรับพารามิเตอร์ด้วย optimizer (เช่น Adam) และ loss function แบบ Dice Loss เพื่อให้โมเดลสามารถเรียนรู้การแบ่งส่วนภาพได้อย่างแม่นยำ

การปรับปรุงประสิทธิภาพ: มีการใช้เทคนิคต่าง ๆ เช่น early stopping และ data augmentation เพื่อป้องกัน overfitting และเพิ่มความสามารถในการ generalize ของโมเดล

3. ผลการทดลอง

3.1 ผลการทดลองของโมเดลแบ่งส่วนขอบเขตรอยโรคผิวหนัง

การทดลองในส่วนนี้มีวัตถุประสงค์เพื่อประเมินประสิทธิภาพของโมเดล ResUNet / U-Net ในการแยก บริเวณรอยโรคออกจากผิวหนังโดยรอบ โดยใช้ชุดข้อมูล ISIC 2018 สำหรับงาน Segmentation

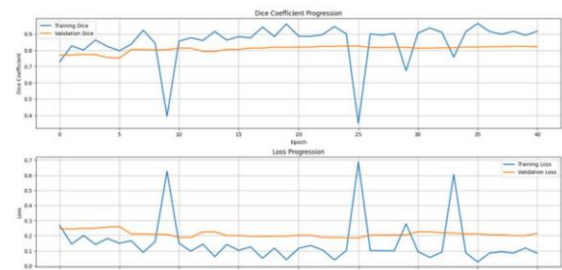
3.1.1 ตัวชี้วัดที่ใช้ในการประเมินผล

การประเมินผลของโมเดล Segmentation ใช้ตัวชี้วัดดังต่อไปนี้ Dice Similarity Coefficient (Dice Score), Intersection over Union (IoU), Accuracy

3.1.2 ผลการประเมินเชิงปริมาณ

จากการทดสอบกับชุดข้อมูลทดสอบ (Test Set) พบว่าโมเดลสามารถแยกขอบเขตรอยโรคได้อย่างมีประสิทธิภาพ โดยมีผลการประเมินดังนี้

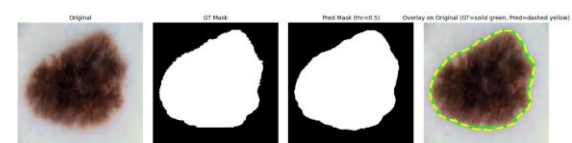
1) กราฟแสดงแนวโน้มของค่า Dice Coefficient Progression ต่อ Epoch และค่า Loss Progression ต่อ Epoch



ภาพที่ 3 แสดงกราฟแนวโน้มของค่า Dice Coefficient และค่า Loss ของโมเดล U-Net

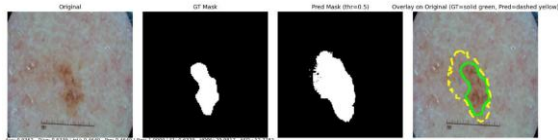
3.1.3 ผลลัพธ์เชิงคุณภาพ (Qualitative Results)

เมื่อพิจารณาผลลัพธ์เชิงภาพ พบว่าโมเดลสามารถแบ่งขอบเขตรอยโรคได้ใกล้เคียงกับ Mask จริงในหลายกรณี อย่างไรก็ตาม ในบางภาพที่รอยโรคมีขนาดเล็กหรือมีสีใกล้เคียงกับผิวหนังโดยรอบ โมเดลอาจไม่สามารถตรวจจับขอบเขตได้



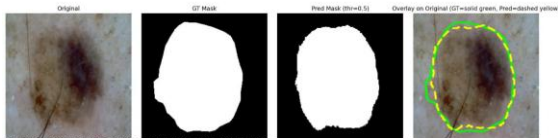
ภาพที่ 4 ตัวอย่างผลการแบ่งส่วนขอบเขตรอยโรคผิวหนังกรณีนี้ โมเดลสามารถตรวจจับ ขอบเขตได้อย่างแม่นยำ

ผลลัพธ์การแบ่งส่วนขอบเขตรอยโรคผิวหนังกรณีที่โมเดลสามารถตรวจจับขอบเขตได้อย่างแม่นยำโดยรอยโรคมีขนาดค่อนข้างใหญ่และมีความแตกต่างของสีและพื้นผิวจากผิวหนัง โดยรอบอย่างชัดเจน ทำให้โมเดล U-Net สามารถเรียนรู้ลักษณะเชิงพื้นที่และขอบเขตของรอยโรคได้ดี ส่งผลให้เส้นขอบของ Predicted Mask (เส้นประสีเหลือง) มีความสอดคล้องกับ Ground Truth (เส้นสีเขียว) ในระดับสูง



ภาพที่ 5 ตัวอย่างผลการแบ่งส่วนขอบเขตรอยโรคกรณีที่รอยโรคมีขนาดเล็กและมีสีใกล้เคียง กับผิวหนังโดยรอบ

ผลลัพธ์การแบ่งส่วนขอบเขตรอยโรคกรณีที่รอยโรคมีขนาดเล็กและมีสีใกล้เคียงกับผิวหนัง โดยรอบ ส่งผลให้ความแตกต่างเชิงคุณลักษณะ (feature contrast) ระหว่างรอยโรคและพื้นหลัง มีน้อย โมเดลจึงไม่สามารถตรวจจับขอบเขตของรอยโรคได้อย่างสมบูรณ์ ทำให้เกิดความคลาดเคลื่อนของขอบเขตระหว่าง Predicted Mask และ Ground Truth โดยเฉพาะบริเวณขอบ ด้านนอกของรอยโรค



ภาพที่ 6 ตัวอย่างผลการแบ่งส่วนขอบเขตรอยโรคกรณีที่ขอบเขตของรอยโรคไม่ชัดเจน

ผลลัพธ์การแบ่งส่วนขอบเขตรอยโรคกรณีที่ขอบเขตของรอยโรคไม่ชัดเจนและมีรูปร่างไม่ สม่าเสมอ ซึ่งอาจเกิดจากลักษณะพื้นผิวของผิวหนังหรือสัญญาณรบกวนในภาพ ส่งผลให้โมเดล เกิดความคลาดเคลื่อนในการกำหนดขอบเขตบางส่วน โดยเฉพาะบริเวณที่รอยโรคค่อย ๆ กลืน กับผิวหนังโดยรอบ

3.2 ผลการทดลองการจำแนกประเภทรอยโรค (Classification Results)

3.2.1 ตัวชี้วัดที่ใช้ในการประเมินผล

การประเมินผลการจำแนกใช้ตัวชี้วัด ได้แก่

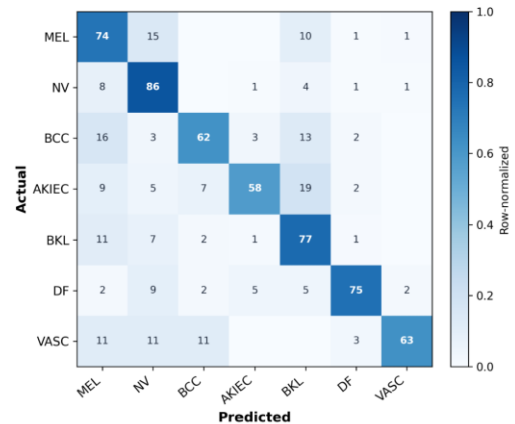
Accuracy, Confidence Score, Cross-Entropy Loss

3.2.2 ผลลัพธ์การจำแนก

จากผลการทดลอง พบว่าโมเดลสามารถจำแนกรอยโรคได้ในระดับที่น่าพอใจ โดยเฉพาะในกลุ่มรอยโรคที่มีลักษณะเด่นชัด เช่น NV และ BKL อย่างไรก็ตาม ในบางกรณีการแบ่งขอบเขตรอยโรคไม่ชัดเจน ผลการจำแนกอาจมีความคลาดเคลื่อน

DermaTrace — Confusion Matrix (ISIC 2018 test, n=1511)

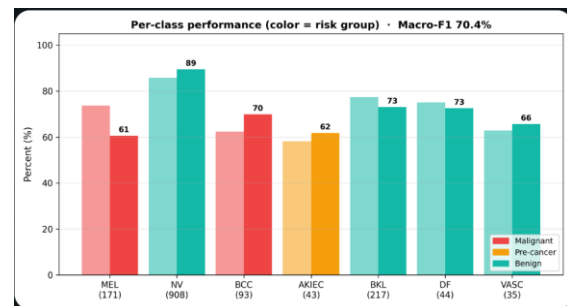
Accuracy 80.2% · Macro-F1 70.4%



ภาพที่ 7 Confusion Matrix แสดงผลการจำแนกรอยโรคผิวหนังของโมเดล Classification

จะเห็นว่าค่าบนแนวแยงหลักมีค่าสูงในหลายคลาส เช่น NV, BCC และ VASC แสดงให้เห็นว่าโมเดลสามารถจำแนกคลาสเหล่านี้ได้อย่างถูกต้อง อย่างไรก็ตาม พบความสัมพันธ์ระหว่างคลาสที่มีลักษณะทางภาพคล้ายกัน เช่น MEL กับ BKL และ AKIEC ซึ่งสะท้อนถึงความท้าทายของข้อมูลทางการแพทย์

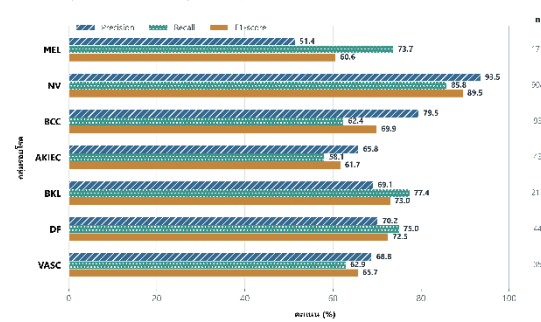
4.2.3 ผลลัพธ์เชิงคุณภาพ (Qualitative Results)



รูปที่ 8 ค่า Recall และ F1-score รายคลาสบนชุดทดสอบ ISIC 2018 โดยแบ่งสีตามกลุ่มความเสี่ยง ตัวเลขในวงเล็บแสดงจำนวนภาพของแต่ละคลาส และแบบจำลองมี Macro-F1 เท่ากับ 70.4%

ประสิทธิภาพการจำแนกรอยโรคผิวหนังรายคลาส

จากชุดทดสอบ ISIC 2018 test set (n = 1,511 ภาพ)

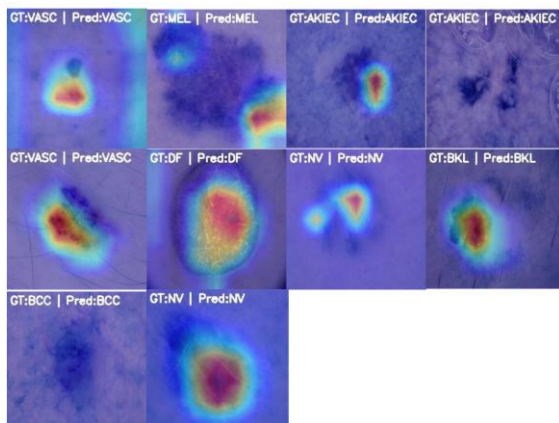


ภาพที่ 9 ประสิทธิภาพการจำแนกรอยโรคผิวหนังรายคลาสบนชุดทดสอบ ISIC 2018 จำนวน 1,511 ภาพ โดยเปรียบเทียบ Precision, Recall และ F1-score พร้อมแสดงจำนวนตัวอย่างของแต่ละคลาส (n)

ผลการทดลองพบว่า NV มี F1-score สูงที่สุดที่ 89% และ Recall 86% ซึ่งสัมพันธ์กับการมีจำนวนตัวอย่างมากที่สุด 908 ภาพ สำหรับกลุ่มมะเร็ง MEL มี Recall 74% แต่ F1-score 61% แสดงว่าแบบจำลองตรวจพบ melanoma ได้ค่อนข้างมาก แต่ยังมีผลบวกложหรือความแม่นยำเชิงบวกไม่สูง ส่วน BCC มี Recall 62% และ F1-score 70%

คลาส AKIEC ซึ่งเป็นกลุ่มก่อนมะเร็งให้ Recall 58% และ F1-score 62% ต่ำที่สุดด้าน Recall จึงยังมีโอกาสพลาดตัวอย่างบางส่วน ขณะที่ BKL และ DF ให้ F1-score ใกล้เคียงกันที่ 73% ส่วน VASC ให้ F1-score 66%

ความแตกต่างระหว่างคลาสสะท้อนผลของความไม่สมดุลของข้อมูล โดยคลาสที่มีจำนวนตัวอย่างน้อย เช่น VASC 35 ภาพ, AKIEC 43 ภาพ และ DF 44 ภาพ มีข้อมูลสำหรับเรียนรู้และประเมินผลจำกัดกว่า NV ดังนั้นค่า Macro-F1 70.4% จึงเหมาะสำหรับแสดงประสิทธิภาพโดยรวม เพราะให้น้ำหนักทุกคลาสเท่ากัน โดยไม่ปล่อยให้คลาส NV ซึ่งมีจำนวนมากครอบงำผลลัพธ์



ภาพที่ 19-28 ภาพรอยโรคผิวหนังที่โมเดลสามารถจำแนกประเภทได้ถูกต้อง จากภาพ Heatmap

พบว่าโมเดลให้ความสนใจกับบริเวณรอยโรคเป็นหลัก โดยเฉพาะบริเวณที่มีลักษณะสี ผื่นผิว และโครงสร้างของรอยโรคที่เด่นชัด ส่งผลให้ผลการทำนายสอดคล้องกับชนิดของรอยโรคจริง (Ground Truth)

4. สรุปและอภิปรายผล

4.1 สรุปผลการดำเนินงาน (Conclusion)

โครงการนี้มีวัตถุประสงค์เพื่อพัฒนาระบบวิเคราะห์รอยโรคผิวหนังจากภาพถ่ายทางการแพทย์ โดยประยุกต์ใช้เทคนิคปัญญาประดิษฐ์และการเรียนรู้เชิงลึก (Deep Learning) ซึ่งประกอบด้วยสอง กระบวนการหลัก ได้แก่ การแบ่งส่วนขอบเขตรอยโรค (Image Segmentation) ด้วยโมเดล U-Net และการจำแนกประเภทรอยโรคผิวหนัง (Image Classification) ด้วยโมเดล ResNet

ในส่วนของการแบ่งส่วนขอบเขตรอยโรค ผลการทดลองแสดงให้เห็นว่าโมเดล U-Net สามารถแยก บริเวณรอยโรคออกจากผิวหนังโดยรอบได้อย่างมีประสิทธิภาพ โดยมีค่า Dice Coefficient และ Accuracy อยู่ในระดับสูง และค่า Dice Loss มีแนวโน้มลดลงอย่างต่อเนื่องเมื่อจำนวนรอบการฝึก (Epoch) เพิ่มขึ้น ซึ่งสะท้อนให้เห็นว่าโมเดลสามารถเรียนรู้ลักษณะเชิงพื้นที่ของรอยโรคได้ดีขึ้น ตามลำดับ นอกจากนี้ ผลการประเมินเชิงคุณภาพจากภาพตัวอย่างแสดงให้เห็นว่าขอบเขตรอยโรคที่ โมเดลทำนายมีความสอดคล้องกับ Mask จริงในหลายกรณี โดยเฉพาะรอยโรคที่มีขอบเขตชัดเจนและมีความแตกต่างจากผิวหนังรอบข้างอย่างเห็นได้ชัด

ในส่วนของการจำแนกประเภทรอยโรค โมเดล ResNet สามารถจำแนกรอยโรคออกเป็น 7 ประเภท ได้แก่ MEL, NV, BCC, AKIEC, BKL, DF และ VASC ได้ในระดับที่น่าพอใจ โดยผลจาก Confusion Matrix แสดงให้เห็นว่าค่าบนแนวทแยงมีค่าสูงในหลายคลาส ซึ่งบ่งชี้ถึงความถูกต้องในการจำแนก อย่างไรก็ตาม ยังพบความสับสนระหว่างบางคลาสที่มีลักษณะทางภาพคล้ายคลึงกัน เช่น BKL และ AKIEC หรือ MEL และ DF

โดยภาพรวม ระบบที่พัฒนาขึ้นสามารถทำงานได้ครบทั้งกระบวนการตั้งแต่การระบุตำแหน่งรอยโรค ไปจนถึงการจำแนกประเภทของรอยโรค ซึ่งแสดงให้เห็นถึงศักยภาพของการประยุกต์ใช้ Deep Learning เพื่อเป็นเครื่องมือช่วยในการวิเคราะห์รอยโรคผิวหนังจากภาพถ่ายทางการแพทย์

4.2 อภิปรายผลการทดลอง(Discussion)

แสดงให้เห็นว่าสามารถรวมการแบ่งส่วนการจำแนกหลายคลาสและการอธิบายผลไว้ในระบบเดียวได้ accuracy บนชุดทดสอบ 80.15% สูงกว่าการเดาสุ่มอย่างชัดเจนแต่ macro F1 70.42% และความแตกต่างระหว่าง validation กับ test ชี้ว่าการวัด accuracy เพียงค่าเดียวไม่เพียงพอโดยเฉพาะเมื่อ NV มีสัดส่วนมากกว่าสองในสามของชุดฝึก MEL มี recall มากกว่า precision แสดงว่าโมเดลตรวจพบเมลาโนมาได้หลายกรณีแต่ยังให้ผลบวกเกินจริงขณะที่ความสับสนระหว่าง MEL,NV และ BKLรวมถึง BCC กับ BKL เป็นประเด็นที่ต้องปรับปรุงการใช้ Grad CAM ช่วยให้เห็นตำแหน่งที่โมเดลให้ความสนใจแต่ไม่ถือเป็นหลักฐานยืนยันความถูกต้องของคำทำนาย



ข้อจำกัดสำคัญคือ validation ถูกสุ่มในระดับภาพ ไม่ได้แยกตาม lesion หรือผู้ป่วย จึงอาจประเมินผลสูงกว่าการใช้งานจริง อีกทั้งข้อมูลส่วนใหญ่เป็นภาพเดอร์โมสโคปี จึงไม่ควรขยายผลไปยังภาพถ่ายโทรศัพท์ทั่วไปโดยไม่มีการทดสอบเพิ่มเติม งานต่อไปควรแบ่งข้อมูลตาม lesion identifier ประเมินหลายสถาบัน ทำ probability calibration รายงาน sensitivity และ specificity ตามวัตถุประสงค์การคัดกรอง และให้แพทย์ผิวหนังประเมินทั้งคำทำนายและ Grad-CAM ก่อนพิจารณาใช้ทางคลินิก

กิตติกรรมประกาศ

คณะผู้จัดทำขอขอบพระคุณโรงเรียนวิทยาศาสตร์จุฬาภรณราชวิทยาลัย เลย์ และอาจารย์ที่ปรึกษา ที่ให้คำแนะนำและสนับสนุนทรัพยากรในการพัฒนาโครงการ ตลอดจนผู้จัดทำชุดข้อมูล HAM10000 และ ISIC ที่เผยแพร่ข้อมูลเพื่อการศึกษาและวิจัย

เอกสารอ้างอิง

Ronneberger, O., Fischer, P., & Brox, T. (2015).

U-Net: Convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention, 9351, 234-241.

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017).

Grad-CAM: Visual explanations from deep networks via gradient-based localization. Proceedings of ICCV, 618-626.

Tan, M., & Le, Q. V. (2019).

EfficientNet: Rethinking model scaling for convolutional neural networks. Proceedings of ICML, 6105-6114.

Tschandl, P., Rosendahl, C., & Kittler, H. (2018).

The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Scientific Data, 5, 180161.

<https://doi.org/10.1038/sdata.2018.161>

Codella, N. C. F., Rotemberg, V., Tschandl, P., Celebi, M. E., Dusza, S., Gutman, D., et al. (2019).

Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the International Skin Imaging Collaboration. arXiv:1902.03368.