

A COMPARISON OF 3D CONVOLUTION NEURAL NETWORKS FOR BRAIN STROKE CLASSIFICATION WITH CT SCAN IMAGES

Chanon WORACHOTSUEPTRAKUN¹

1 Faculty of Commerce and Accountancy, Chulalongkorn university, Thailand;
6380073826@acc.chula.ac.th

ARTICLE HISTORY

Received: 21 April 2023

Revised: 10 May 2023

Published: 22 May 2023

ABSTRACT

Convolutional Neural Networks is widely used in image classification, detection, segmentation, etc, especially the application of medical applications that are very popular today. Most of the research in which convolution networks are applied in brain diseases. most commonly used image detection or segmentation but a few studies have used image classification methods for brain diseases. The researcher considers that medical image classification is an important method and can be further developed The image classification problem can be solved in a variety of ways. one of which is image classification through a deep learning network with a 3D convolution network that can help solve problems quickly and accurately. This research presents the application of 3D convolution networks together with image preprocessing. In order to classify stroke from CT-scans, it is divided into 2 types; Strokes and Normal. 3D convolution network is applied with 4 pre-trained models, consisting of Alexnet, VGG-16, Googlenet and Resnet. After that, details will be adjusted by bringing the strengths of each model to use and add a layer of depth that is a hallmark of finding key 3D features. Subsequently, this research applied K-Fold cross validation techniques to resolve the bias in each model before using the Confusion matrix to evaluate the performance of the model. Googlenet 3D found that performance produces the best results from the outcome evaluation. The results of the stroke classification test from CT-scan images gave an accuracy of 92.00%. This research shows that the use of a 3D convolution network with image classification approach should be further developed to benefit the medical field.

Keywords: Deep Learning, Convolutional Neural Network, Stroke Ct-Scan Images

CITATION INFORMATION: Worachotsueptrakun, C. (2023). A Comparison of 3D Convolution Neural Networks for Brain Stroke Classification with CT Scan Images. *Procedia of Multidisciplinary Research*, 1(5), 14.

การเปรียบเทียบสถาปัตยกรรมโครงข่ายประสาทคอนโวลูชัน 3 มิติ โดย การจำแนกโรคหลอดเลือดสมองจากภาพการฉายรังสีเอกซเรย์สมอง

ชานนท์ วรโชติสีปตระกูล¹

1 คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย; 6380073826@acc.chula.ac.th

บทคัดย่อ

แบบจำลองโครงข่ายคอนโวลูชัน หรือ ซีเอ็นเอ็น (Convolutional Neural Networks หรือ CNN) ได้รับการนำมาใช้กันอย่างแพร่หลายในการจำแนกภาพ, การตรวจจับ, การแบ่งส่วนภาพ และอื่นๆ โดยเฉพาะการนำมาประยุกต์ใช้ทางการแพทย์ในปัจจุบันที่ได้รับความนิยมอย่างมาก ซึ่งงานวิจัยส่วนใหญ่ที่นำโครงข่ายคอนโวลูชันมาประยุกต์ใช้กับโรคที่เกี่ยวข้องกับสมอง นิยมใช้เทคนิคการตรวจจับภาพ หรือการแบ่งส่วนรอยโรคของสมอง แต่มีงานวิจัยน้อยชิ้นที่ใช้วิธีการจำแนกภาพกับโรคที่เกี่ยวข้องกับสมอง ซึ่งผู้วิจัยเล็งเห็นว่าการจำแนกภาพทางการแพทย์เป็นวิธีการที่มีความสำคัญ และสามารถพัฒนาต่อยอดได้ ทั้งนี้ปัญหาในการจำแนกภาพสามารถทำได้หลากหลายวิธี หนึ่งในนั้นคือการจำแนกภาพผ่านโครงข่ายการเรียนรู้เชิงลึกด้วยโครงข่ายคอนโวลูชัน 3 มิติ ที่สามารถช่วยแก้ไขปัญหาดังกล่าวได้อย่างรวดเร็ว และแม่นยำ งานวิจัยนี้เสนอการประยุกต์ใช้โครงข่ายคอนโวลูชัน 3 มิติ ร่วมกับการประมวลผลภาพล่วงหน้า (Image Preprocessing) เพื่อจำแนกโรคหลอดเลือดสมองจากภาพการฉายรังสีเอกซเรย์สมอง (CT-Scan) แบ่งออกเป็น 2 ประเภท 1) ภาพของคนเป็นโรคหลอดเลือดสมอง 2) ภาพของคนไม่เป็นโรคหลอดเลือดสมอง โดยมีการนำโครงข่ายคอนโวลูชัน 3 มิติมาประยุกต์ร่วมกับแบบจำลองที่ถูกฝึกมาเรียบร้อยแล้ว 4 แบบจำลอง ประกอบไปด้วย อเล็กซ์เน็ต วีจีจี-16 กูเกิลเน็ต และเรสเน็ต หลังจากนั้นจะมีการปรับรายละเอียดโดยการนำจุดเด่นของแต่ละแบบจำลองมาใช้ และเพิ่มความลึกที่เป็นจุดเด่นของการค้นหาคุณลักษณะสำคัญของรูปแบบ 3 มิติ จากนั้นงานวิจัยชิ้นนี้ได้นำเทคนิค K-Fold Cross validation มาใช้เพื่อแก้ไขปัญหามือถือคติต่อแต่ละแบบจำลอง ก่อนที่จะใช้ Confusion matrix เป็นเครื่องมือในการประเมินประสิทธิภาพของแบบจำลอง ซึ่งพบว่าสมรรถนะแบบจำลองโครงข่ายคอนโวลูชันกูเกิลเน็ต 3 มิติให้ผลลัพธ์ที่ดีที่สุดจากการประเมินผลลัพธ์ โดยผลการทดสอบการจำแนกโรคหลอดเลือดสมองจากภาพการฉายรังสีเอกซเรย์ ให้ค่าความแม่นยำที่ 92.00% ซึ่งงานวิจัยนี้เห็นได้ว่าการนำโครงข่ายคอนโวลูชันรูปแบบ 3 มิติมาใช้ร่วมกับการจำแนกภาพที่นำเสนอมีความหวังที่สามารถนำไปพัฒนา และเป็นประโยชน์ต่อทางการแพทย์ได้

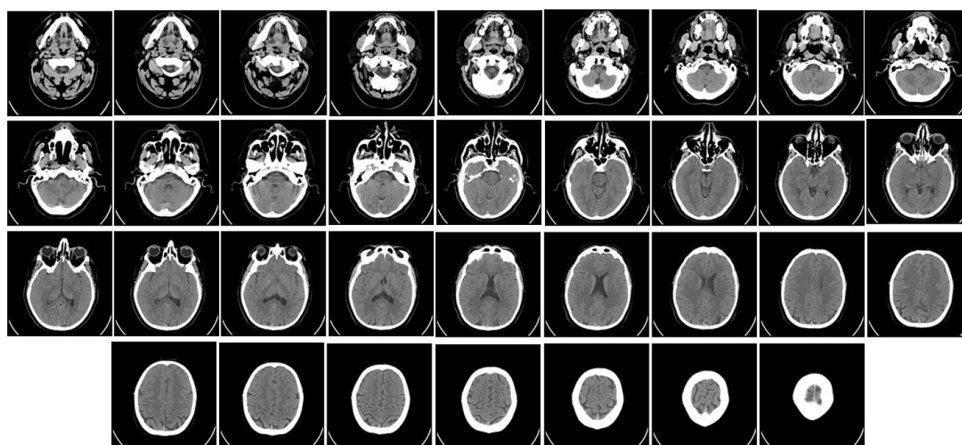
คำสำคัญ: การเรียนรู้เชิงลึก, โครงข่ายคอนโวลูชัน, ภาพการฉายรังสีเอกซเรย์สมอง

ข้อมูลการอ้างอิง: ชานนท์ วรโชติสีปตระกูล. (2566). การเปรียบเทียบสถาปัตยกรรมโครงข่ายประสาทคอนโวลูชัน 3 มิติ โดยการจำแนกโรคหลอดเลือดสมองจากภาพการฉายรังสีเอกซเรย์สมอง. *Procedia of Multidisciplinary Research*, 1(5), 14.

บทนำ

ในแต่ละปีจะมีผู้เสียชีวิตจากโรคหลอดเลือดสมองประมาณ 14 ล้านคนจากทั่วโลก และผู้รอดชีวิตกว่า 50 ล้านคนจะมีชีวิตอยู่ด้วยความทุพพลภาพถาวร จากข้อมูลขององค์การอนามัยโลก (World Stroke Organization. World Stroke Organization, 2021) โรคหลอดเลือดสมองเป็นสาเหตุการตายอันดับ 2 ของผู้ที่มีอายุมากกว่า 60 ปี และเป็นสาเหตุอันดับ 5 ในผู้ที่มีอายุระหว่าง 15-59 ปี อีกทั้งโรคหลอดเลือดสมองยังเป็นสาเหตุหลักของความทุพพลภาพในระยะยาว ส่งผลต่อการดำรงชีวิตด้วยความทุพพลภาพ (Fartakov, 2019) การเลือกวิธีการวินิจฉัยโรคนั้น ขึ้นอยู่กับแต่ละกรณี อาทิเช่น อาการของผู้ป่วย หรือเวลาที่เริ่มมีอาการ เป็นต้น ทั้งนี้ความล่าช้าใดๆ ในการรักษาพยาบาลส่งผลต่อการเพิ่มความเสียหายของผลกระทบที่รุนแรง และเสียชีวิต

ขั้นตอนในการวินิจฉัยโรคหลอดเลือดสมองสามารถทำได้หลากหลายวิธี หนึ่งในนั้นคือการทำ CT Scan ที่สามารถทำงานได้ซับซ้อนกว่าการเอกซเรย์แบบธรรมดา จะมีการแสดงภาพเป็น 3 มิติ และซอยภาพออกเป็นแผ่นบางๆ ในภาพลักษณะตัดขวางได้หลายสิบแผ่น ช่วยให้แพทย์วิเคราะห์ความผิดปกติของอวัยวะนั้นๆ ได้อย่างละเอียด



รูปที่ 1 ภาพการทำ CT Scan สมองของคนหนึ่งคน

สำหรับการจำแนกโรคของผู้ป่วยที่เป็นโรคหลอดเลือดสมองมีการนำเทคนิคการประมวลผลด้วยภาพมาช่วยในการวินิจฉัยโรคของผู้ป่วย หนึ่งในเทคนิคที่นิยมใช้กันคือใช้โครงข่ายคอนโวลูชัน (Convolutional Neural Networks) ซึ่งเป็นเทคนิคที่นิยมใช้สำหรับการจำแนกภาพทางการแพทย์ (Zhang, 2021) โดยมีการคิดค้นครั้งแรกในปี 1980 โดย Yann LeCun (Dickson, 2021) การทำงานของโครงข่ายคอนโวลูชัน มีด้วยกัน 3 รูปแบบ คือ 1 มิติ, 2 มิติ และ 3 มิติ โดยรูปแบบนำมาใช้ในการจำแนกภาพทั่วไปคือรูปแบบ 2 มิติ ในส่วนของรูปแบบ 3 มิติ จะใช้ในส่วนของการจำแนกภาพเคลื่อนไหว, ภาพถ่ายวิดีโอ, รูปภาพที่มีรูปแบบ 3 มิติ และอื่น

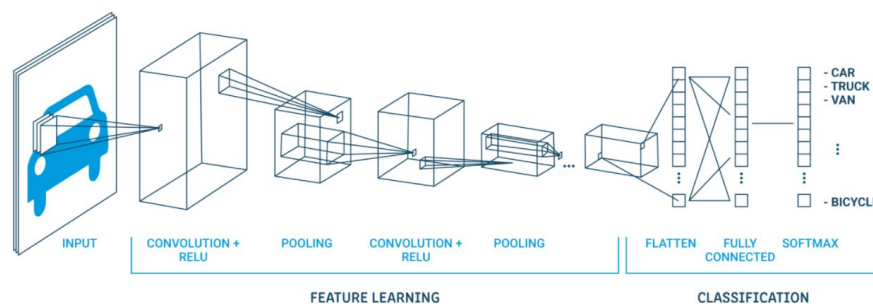
แบบจำลองด้วยโครงข่ายประสาทคอนโวลูชันรูปแบบ 3 มิติ (H. Fan, 2017) มีความสามารถในการคัดเลือกคุณลักษณะของข้อมูลได้อย่างมีประสิทธิภาพ เนื่องจากรูปแบบ 3 มิติ มีการประมวลผล หรือหน่วยความจำที่สูงมาก รวมถึงมีการค้นหารูปแบบความสัมพันธ์แบบ 3 มิติ คือ ความสูง ความกว้าง และความลึก ซึ่งมีประสิทธิภาพในการทำนายที่ดีกว่ารูปแบบ 2 มิติ ที่แยกรูปแบบเพียง 2 มิติเท่านั้นคือ ความสูง และความกว้าง ตัวอย่างเช่น การสแกนแบบ 2 มิติ สามารถแยกความแตกต่างระหว่าง วงกลม และ ทรงกลมได้ แต่หากเป็นการ สแกนแบบ 3 มิติจะสามารถจำแนกวัตถุที่เป็นทั้ง วงกลม ทรงกลม และทรงกระบอก

ผู้วิจัยมีความสนใจที่จะทำการศึกษการจำแนกภาพของคนเป็นโรคหลอดเลือดสมองจากภาพการฉายรังสีเอกซเรย์ (CT Scan) โดยเทคนิคการสร้างแบบจำลองโครงข่ายประสาทคอนโวลูชันรูปแบบ 3 มิติ โดยประยุกต์ใช้กับแบบจำลองที่เป็นที่นิยมในการนำมาใช้จำแนกภาพทางการแพทย์ (Liwei Tan, 2021) (Balasubramanian Raman, 2021)

ประกอบด้วย อเล็กซ์เน็ต วีจีจี-16 กูเกิลเน็ต และเรสเน็ต โดยจะมีการใช้เทคนิคการประมวลผลภาพล่วงหน้า เพื่อช่วยในการเพิ่มประสิทธิภาพของการทำโครงข่ายคอนโวลูชัน 3 มิติ ซึ่งงานวิจัยชิ้นนี้ มีการนำเทคนิค K-Fold Cross Validation เข้ามาใช้เพื่อแก้ปัญหาการมีอคติ (Bias) ในแต่ละแบบจำลอง โดยจะมีการใช้ Confusion Matrix เป็นเครื่องมือสำคัญในการประเมินผลลัพธ์ของการทำนายจากแบบจำลองที่สร้างขึ้น เพื่อนำผลจากการประเมินมาวิเคราะห์การเปรียบเทียบประสิทธิภาพของแต่ละแบบจำลองเพื่อให้ได้แบบจำลองที่เหมาะสมกับชุดข้อมูลภาพการฉายรังสีเอกซเรย์ของคนเป็นโรคหลอดเลือดสมองมากที่สุด

ทบทวนวรรณกรรม

โครงข่ายประสาทคอนโวลูชัน (Convolutional neural networks, CNN) (Lovekesh Vig, 2019) เป็นการใช้ประโยชน์จากความสัมพันธ์เชิงพื้นที่โดยบังคับใช้รูปแบบการเชื่อมต่อภายในระหว่างเซลล์ประสาทของเลเยอร์ที่อยู่ติดกัน โดยวิธีการทำงานของ CNN ที่ใช้ในการจัดประเภทรูปภาพจะดำเนินการค้นหาคุณลักษณะของรูปภาพ เช่น ขอบและส่วนโค้งต่างๆ และจากนั้นจะสร้างภาพที่เป็นนามธรรมมากขึ้นผ่านชุดเลเยอร์ที่ปิดบังไว้ ทั่วไปประกอบด้วยเลเยอร์ที่แตกต่างกันอย่างน้อย 4 ชั้น ได้แก่ Convolutional layer , Pooling/Subsampling layer, Fully connected layer และ Output layer



รูปที่ 2 โครงสร้างโครงข่ายประสาทคอนโวลูชัน (Hassan, 2018)

ในงานวิจัยเกี่ยวกับโครงข่ายประสาทคอนโวลูชันที่นำมาใช้เพื่อจำแนกโรคทางการแพทย์ Bhanumathi (Bhanumathi, 2019) มีการนำแบบจำลองโครงข่ายคอนโวลูชันที่ถูกฝึกมาเรียบร้อยแล้ว (Pre-trained Model) อาทิเช่น อเล็กซ์เน็ต , วีจีจี-16 และ กูเกิลเน็ต มาใช้ในการจำแนกคนเป็นโรคเนื้องอกในสมองจากผลสแกน MRI ซึ่งได้ผลลัพธ์ของแบบจำลองที่มีประสิทธิภาพ อย่างไรก็ตาม การนำเทคนิคโครงข่ายคอนโวลูชันแบบ 3 มิติ มาใช้ในการจำแนกโรคเกี่ยวกับภาพสมองยังมีจำนวนที่น้อย โดยงานวิจัยเกี่ยวกับสมองจำนวนมากจะนิยมใช้การแบ่งส่วนภาพ หรือการแบ่งส่วนรอยโรคเข้ามาเพื่อช่วยในการประกอบการตัดสินใจ (Ottesen, 2023) โดยใช้ในการแยกเนื้อเยื่อสมองส่วนที่เสียหาย และมีสุขภาพปกติ หลังจากนั้นจะนำมาตัดสินใจเลือก และรักษาต่อไป ซึ่งวิธีนี้สามารถช่วยในการคาดการณ์พื้นที่ และช่วยแพทย์ในการลดจำนวนเวลาการวินิจฉัยโรคจากผู้ป่วยที่มีจำนวนมาก ซึ่งสามารถช่วยในการประกอบการตัดสินใจในการเลือกดูแลผู้ป่วยฉุกเฉินก่อนเป็นอันดับแรก (A. V. Dobshik, 2023)

Anjali Gautam (Gautam, 2021) มีการนำเสนอการจำแนกโรคหลอดเลือดสมองจากภาพการฉายรังสีเอกซเรย์ (CT Scan) โดยการใช้โครงข่ายประสาทคอนโวลูชัน ซึ่งมีแนวทางการนำแบบจำลองที่ถูกฝึกมาเรียบร้อยแล้วอย่าง อเล็กซ์เน็ต และเรสเน็ต 50 เทียบกับแบบจำลองโครงข่ายคอนโวลูชันที่ผู้วิจัยนำเสนอ โดยมีการนำเทคนิคการประมวลผลภาพล่วงหน้าเพื่อปรับปรุงคุณภาพของข้อมูลภาพ ซึ่งเทคนิคการประมวลผลภาพล่วงหน้าสามารถช่วยดึงคุณลักษณะ (Feature) ที่มีความสำคัญจากชั้นต่างๆ ของ CNN ได้มากยิ่งขึ้น และสามารถจัดความไม่คมชัดของรูปภาพบางอย่าง รวมไปถึงมีการปรับความคมชัดเพื่อให้มองเห็นพื้นที่ที่ต้องการได้ดีขึ้น เพื่อให้การจัดหมวดหมู่มีความแม่นยำมากขึ้น

เทคนิคการประยุกต์ใช้โครงข่ายคอนโวลูชัน 3 มิติเริ่มได้รับความนิยมมากขึ้นในปัจจุบันในการนำมาใช้ร่วมกับข้อมูลทางการแพทย์เนื่องจากสามารถช่วยดึงคุณลักษณะสำคัญของข้อมูลได้ลึก และซับซ้อนมากยิ่งขึ้น A.V. Dobshik (A. V. Dobshik, 2023) มีการนำเสนอการใช้โครงข่ายคอนโวลูชัน 3 มิติ ในการแบ่งส่วน (Segmentation) ของรอยโรคหลอดเลือดสมองจากภาพ CT-Scan โดยมีการประยุกต์ใช้โครงข่ายคอนโวลูชันยูเน็ต 3 มิติ งานวิจัยชิ้นนี้ได้นำเทคนิคการประมวลผลภาพล่วงหน้า เพื่อช่วยเพิ่มประสิทธิภาพในการให้ข้อมูลสามารถแปลผลการแบ่งส่วนได้แม่นยำยิ่งขึ้น และมีการวัดประสิทธิภาพของแบบจำลองโดยการใช้เทคนิค K-fold Cross validation โดยใช้เครื่องมือ Confusion Matrix เป็นเครื่องมือสำคัญในการประเมินผลลัพธ์ ในการหาค่าความไว (Sensitivity), ความเฉพาะ (Specificity), ความถูกต้อง (Precision)

วิธีการวิจัย

งานวิจัยนี้ผู้วิจัยมีความสนใจที่จะสร้างแบบจำลองการเรียนรู้เชิงลึกโดยโครงข่ายประสาทคอนโวลูชันเพื่อเปรียบเทียบประสิทธิภาพของโมเดลแบบจำลองโครงข่ายคอนโวลูชันแบบอเล็กซ์เน็ต, กูเกิลเน็ต, วีจีจี-16, เรสเน็ต50 โดยมีชุดข้อมูลของภาพโรคหลอดเลือดสมอง และไม่เป็นโรคหลอดเลือดสมองวิธีวิจัยการสร้างแบบจำลองและเตรียมข้อมูลแบ่งออกเป็น 4 ขั้นตอนหลัก:

1) การเตรียมชุดข้อมูล

จัดเตรียมชุดข้อมูลภาพการฉายรังสีเอกซเรย์ของคนเป็นโรคหลอดเลือดสมอง และไม่เป็นโรคหลอดเลือดสมอง จากฐานข้อมูลออนไลน์ใน <https://www.kaggle.com/afriidrahman/brain-stroke-ct-image-dataset> โดยจากข้อมูลสามารถแบ่งข้อมูลได้เป็น 2 ประเภท คือ 1) ภาพโรคหลอดเลือดสมอง การฉายรังสีเอกซเรย์ (950 ภาพ จาก 40 คน) และ 2) ภาพไม่เป็นโรคหลอดเลือดสมองจากการฉายรังสีเอกซเรย์ (1551 ภาพ จาก 82 คน) หลังจากนั้นทำการแบ่งข้อมูล โดยการแบ่งข้อมูลออกเป็น ข้อมูลสอน ข้อมูลตรวจสอบ และข้อมูลทดสอบ

ตารางที่ 1 ในการฝึกสอนและทดสอบแบบจำลอง จะทำการแบ่งข้อมูลออกเป็น ข้อมูลสอน ข้อมูลตรวจสอบ และข้อมูลทดสอบ

ประเภทรูปภาพ	ข้อมูลสอน	ข้อมูลตรวจสอบ	ข้อมูลทดสอบ	จำนวนทั้งหมด
ภาพของผู้ที่เป็นโรคหลอดเลือดสมอง	570	190	190	950
ภาพของผู้ที่ไม่เป็นโรคหลอดเลือดสมอง	931	310	310	1,551
รวม	1,501	500	500	2,501

2) การสร้างแบบจำลองการทำนายโรคหลอดเลือดสมอง

ทำการทดสอบแบบจำลองการเรียนรู้เชิงลึกแบบคอนโวลูชัน 3 มิติทั้ง 4 แบบ คือ อเล็กซ์เน็ต, กูเกิลเน็ต, วีจีจี-16 และเรสเน็ต โดยจะมีการปรับแต่งโครงสร้างพื้นฐาน จากการเพิ่มหรือลดจำนวนตัวกรอง (Filter) ของชั้นคอนโวลูชัน (Convolution Layer) เพื่อให้โมเดลสามารถที่จะทำนายได้มีประสิทธิภาพมากยิ่งขึ้น หรือแม่นยำมากขึ้น

3) การเพิ่มประสิทธิภาพ และเปรียบเทียบประสิทธิภาพของแบบจำลอง

การเพิ่มประสิทธิภาพของแบบจำลองในงานวิจัยนี้ จะใช้เทคนิคการทำ K-Fold Cross Validation เพื่อหาชุดพารามิเตอร์ที่ดีที่สุดของแต่ละแบบจำลอง เพื่อใช้ในการเปรียบเทียบ และหาโมเดลที่ดีที่สุดจากทั้ง 4 แบบจำลอง

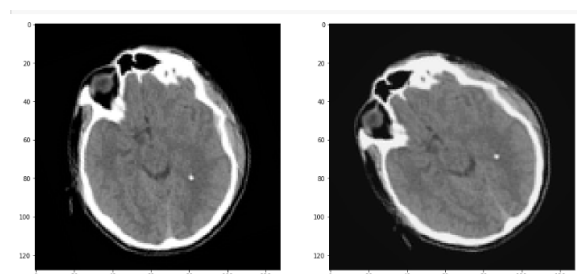
4) การประเมินแบบจำลองผลการทำนายโรคหลอดเลือดสมองจากภาพการฉายรังสีเอกซเรย์

ตัววัดที่ใช้ในการประเมินแบบจำลองโครงข่ายประสาทเทียมคอนโวลูชันทั้ง 4 แบบ ประกอบด้วย อเล็กซ์เน็ต, กูเกิลเน็ต, วีจีจี-16 และเรสเน็ต โดยจะตรวจสอบผลลัพธ์การทำนายด้วยแบบจำลอง คอนฟิวชันเมทริกซ์ (Confusion Matrix) และ

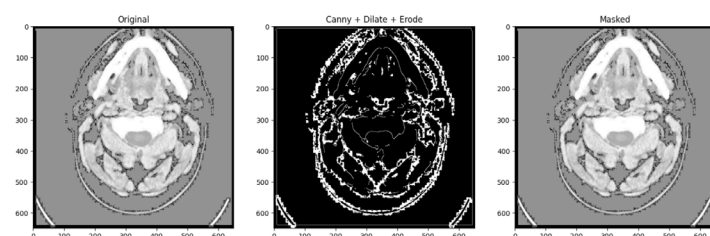
ใช้ตัวสถิติที่ใช้วัดความถูกต้องของการพยากรณ์เพิ่มเติม ได้แก่ ค่าความแม่นยำ (Accuracy), ค่าความครบถ้วน (Recall), ค่าความถูกต้อง (Precision) และ F1-Score

การทดสอบสมมติฐานของการวิจัย

การวิจัยการเปรียบเทียบโครงข่ายคอนโวลูชันทั้ง 4 รูปแบบ การสอนแบบจำลองเป็นขั้นตอนที่จะสร้างแบบจำลองสำหรับการทำนายคนเป็นโรคหลอดเลือดสมอง และไม่เป็นโรคหลอดเลือดสมอง เนื่องจากข้อมูลของงานวิจัยชิ้นนี้เป็นรูปแบบประเภทภาพขนาด 650x650 ซึ่งมีจำนวนพิกเซลที่ต้องใช้คำนวณจำนวนมาก และในแต่ละแบบจำลองมีโครงสร้างของการนำเข้าข้อมูล (input) ที่แตกต่างกัน งานวิจัยชิ้นนี้จึงทำการตัดสินใจเพื่อที่จะลดปัญหาการ Bias ของข้อมูลโดยการปรับขนาดของข้อมูลนำเข้าของแต่ละโมเดล ด้วยรูปภาพขนาด 128 x 128 หลังจากนั้นทำการปรับค่าน้ำหนักของนิวรอนเน็ตเวิร์ก และตัวกรอง รวมไปถึงคำนวณหาค่าน้ำหนักฟังก์ชันกระตุ้นและปรับค่าตาม optimizer ที่ใช้ในแต่ละโครงสร้าง เนื่องจากชุดข้อมูลกลุ่มตัวอย่างมีจำนวนไม่มาก จึงต้องมีการเพิ่มข้อมูลภาพโดยใช้เทคนิคการทำ Data Augmentation เพื่อเพิ่มจำนวนตัวอย่างของรูปภาพให้มีความหลากหลายเพิ่มขึ้น โดยทำการใช้เทคนิค การหมุนภาพ (Rotate) ,การปรับขนาดภาพ (Resize) อีกทั้งยังนำเทคนิคการทำ Image preprocessing เพื่อปรับจูนข้อมูลภาพให้เป็นรูปแบบ 3 มิติเพื่อให้สอดคล้องกับงานวิจัยชิ้นนี้

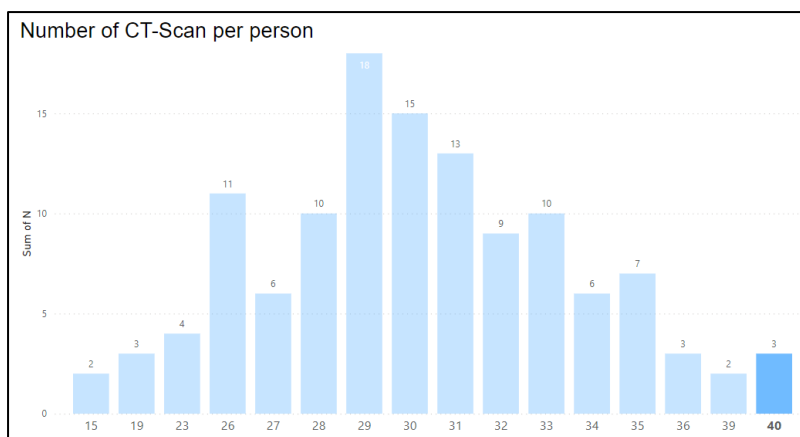


รูปที่ 3 ตัวอย่างการทำ Data Augmentation โดยการหมุนภาพ



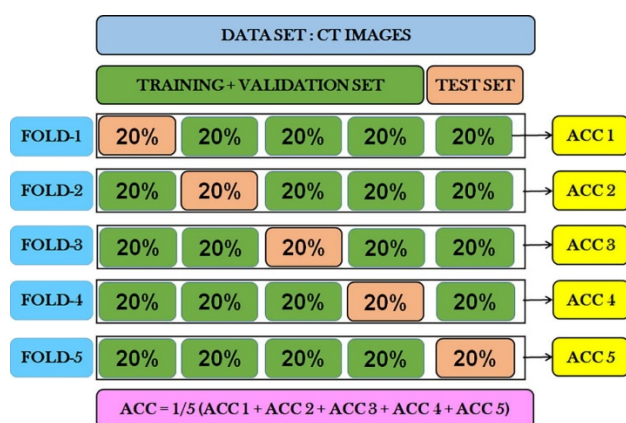
รูปที่ 4 ตัวอย่างการทำการประมวลผลภาพล่วงหน้ากับชุดข้อมูล

งานวิจัยชิ้นนี้มีการนำเทคนิคการทำ Image preprocessing เพื่อปรับจูนข้อมูลภาพให้เป็นรูปแบบ 3 มิติ แต่มีข้อจำกัดในเรื่องของข้อมูลภาพที่สามารถกำหนดชั้นความลึกได้เพียง 40 ชั้น ซึ่งเป็นค่าที่มากที่สุดจากภาพการฉายรังสีเอกซเรย์ต่อคนในชุดข้อมูลของงานวิจัยชิ้นนี้ ดังรูปตัวอย่างที่ 5



รูปที่ 5 ตัวอย่างภาพการหาค่าสูงสุดของภาพถ่ายรังสีเอกซเรย์ของแต่ละคน

โครงข่ายประสาทเทียมคอนโวลูชัน 3 มิติทั้ง 4 แบบจำลอง ประกอบด้วย อเล็กซ์เน็ต, วีจีจี-16, กูเกิลเน็ต และ เรสเน็ต จะมีการนำเทคนิค K-Fold Cross Validation เข้ามาเพื่อใช้จัดการกับปัญหาการ Bias ของแบบจำลอง รวมถึงเพื่อปรับปรุงสมรรถนะของแบบจำลองให้ดียิ่งขึ้น



รูปที่ 6 การแบ่งข้อมูลแบบ K-Fold Cross Validation (Pushpendra Singh, 2021)

โดยงานวิจัยชิ้นนี้กำหนดให้ K = 5 หลังจากที้นำข้อมูลแบ่งส่วนเป็นที่เรียบร้อย จึงทำการ Training Data เพื่อทดสอบว่าแบบจำลองชุดไหนทำงานได้ดีมากที่สุด โดยทำการวัดประสิทธิภาพจากค่า Confusion Matrix

ผลการวิจัย

จากการสร้างแบบจำลองทั้ง 4 โครงสร้าง และการใช้เทคนิค K-Fold Cross Validation กับโมเดลทั้ง 4 รูปแบบ ต่อมา เป็นขั้นตอนของการทดสอบแบบจำลองด้วยชุดข้อมูลทดสอบใช้ค่า prediction ค่าทำนายเท่ากับ 0 หมายถึงทำนายว่าเป็นภาพของคนที่เป็นโรคหลอดเลือดสมอง และเมื่อค่าทำนายเท่ากับ 1 หมายถึงทำนายว่าเป็นภาพของคนไม่เป็นโรคหลอดเลือดสมอง จากนั้นทำการสรุปผลเป็น Confusion matrix และแสดงผลการทำนายเป็นแต่ละค่าดังนี้ precision, recall, f1 score และaccuracy ดังตารางที่ 2

ตารางที่ 2 การวัดประสิทธิภาพของแบบจำลองทั้ง 4 รูปแบบ Alexnet, VGG-16, Googlenet และ Resnet

ตัวชี้วัด Model	Confusion matrix				Macro average			Accuracy
	TP	FN	FP	TN	Precision	Recall	F1-Score	
3D - Alexnet	289	15	46	150	90.91%	76.53%	83.10%	87.80%
3D - VGG16	283	32	47	138	81.18%	74.59%	77.75%	84.20%
3D - Googlenet	303	10	30	157	94.01%	83.96%	88.70%	92.00%
3D - ResNet	273	49	46	132	72.93%	74.16%	73.54%	81.00%

โดยสมรรถภาพของแบบจำลองทั้ง 4 โครงสร้างคือ อเล็กซ์เน็ต 3 มิติ, วีจีจี 3 มิติ, ภูเก็ตเน็ต 3 มิติ และ เรสเน็ต 3 มิติ จากตารางที่ 1 พบว่าแบบจำลองที่ให้ค่าความแม่นยำที่สูงที่สุดเมื่อเทียบกับโมเดลอื่นคือ Googlenet และมีแบบจำลอง Alexnet เป็นจำลองที่ให้ค่าความแม่นยำที่ต่ำกว่าเพียงเล็กน้อยรองลงมา ที่มีค่าความแม่นยำที่ 90.91% และ 87.50% ตามลำดับโดยแบบจำลอง Resnet ให้ค่าความแม่นยำที่ต่ำที่สุดเมื่อมีการเปรียบเทียบกับแบบจำลองอื่น ซึ่งสันนิษฐานได้ว่า เมื่อแบบจำลองมีข้อจำกัดในการกำหนดชั้นด้วยความลึก (Depth) เท่ากับ 40 ตามข้อจำกัดของภาพฉายรังสีเอกซเรย์ อาจทำให้แบบจำลองบางโครงสร้างไม่สามารถดึงคุณลักษณะเด่นของแบบจำลองออกมาได้ ซึ่งส่งผลต่อการทดสอบสมรรถนะของแบบจำลองในการแยกคุณลักษณะของรูปภาพชุดใหม่ได้ไม่ดีมากนัก จากที่กล่าวมาข้างต้น เนื่องมาจากแบบจำลองโครงข่ายคอนโวลูชันเรสเน็ตซึ่งมีความสามารถเด่นในการจัดการกับโครงข่ายที่มีความลึกที่ค่อนข้างมาก แต่เนื่องด้วยข้อจำกัดของความลึกของภาพที่มีเพียง 40 ชั้น ส่งผลให้โครงข่ายคอนโวลูชันเรสเน็ตไม่สามารถแสดงประสิทธิภาพของแบบจำลองได้อย่างเต็มที่

สรุปและอภิปรายผลการวิจัย

สำหรับการทดลองเพื่อจำแนกภาพถ่ายรังสีเอกซเรย์ของผู้ป่วยที่เป็นโรคหลอดเลือดสมอง และผู้ป่วยที่ไม่เป็นโรคหลอดเลือดสมอง โดยใช้ชุดข้อมูลที่ถูกรวบรวมจากออนไลน์ที่น่าเชื่อถือ และมีขนาด 650 x 650 จากข้อจำกัดของข้อมูลที่มีจำนวนน้อยผู้วิจัยจึงใช้เทคนิคการเพิ่มข้อมูลภาพจากการใช้เทคนิค Data Augmentation หลังจากนั้นทำการแปลงภาพให้กลายเป็นรูปแบบ 3 มิติ ที่กำหนดให้มีความลึก 40 ชั้น เพื่อให้สามารถเข้าสู่กระบวนการของแบบจำลองโครงข่ายคอนโวลูชัน 3 มิติได้ในภายหลัง ในขั้นตอนถัดไปผู้วิจัยได้ทำการสร้างแบบจำลองโครงข่ายคอนโวลูชัน 3 มิติซึ่งประกอบไปด้วย โครงข่ายประสาทแบบ อเล็กซ์เน็ต 3 มิติ วีจีจี 3 มิติ ภูเก็ตเน็ต 3 มิติ และ เรสเน็ต 3 มิติ

ในขั้นตอนถัดไปการใช้เทคนิค K-Fold Cross Validation เข้ามาเพื่อช่วยในการลดปัญหาของการ Bias ในแต่ละแบบจำลอง ซึ่งมีการนำทุกแบบจำลองใช้เทคนิค K-Fold Cross Validation ก่อน และหลังจากนั้นจึงนำมาทำการเปรียบเทียบประสิทธิภาพของแบบจำลองโดยใช้เครื่องมือ Confusion Matrix ในการหาค่า precision, recall, f1 score และ accuracy โดยผลการทดสอบที่ได้ผลลัพธ์ที่ให้ค่าความแม่นยำมากที่สุดคือแบบจำลองคอนโวลูชันภูเก็ตเน็ต ที่ให้ค่า 92.00% และให้ค่า precision, recall, f1 score มีค่าเป็น 94.01% ,83.96% ,88.70% ตามลำดับ

การวิจัยชิ้นนี้มีการเรียนรู้การเปรียบเทียบกันระหว่างแบบจำลองโครงข่ายคอนโวลูชัน ภูเก็ตเน็ต ที่มีผลการทดลองที่ดีที่สุด ผู้วิจัยเชื่อว่า เนื่องมาจากแบบจำลองโครงข่ายคอนโวลูชันภูเก็ตเน็ต เป็นแบบจำลองที่มีแนวคิดที่จะพัฒนาการดึงคุณลักษณะสำคัญของตัวข้อมูลไม่เพียงแค่นิวรัล แต่ยังสนใจในแนวทางด้วย จึงสามารถที่จะ Capture ความสัมพันธ์ที่หลากหลายมากขึ้น อีกทั้งยังมีการใช้จุดเด่นของ Kernel size 1x1 ซึ่งทำให้สามารถช่วย Capture ความสัมพันธ์นิวรัลได้เป็นอย่างดี จึงส่งผลให้งานวิจัยชิ้นนี้มีข้อจำกัดของความลึกของข้อมูลภาพ ส่งผลให้แบบจำลองโครงข่ายคอนโวลูชันภูเก็ตเน็ตมีผลการทดลองที่มีประสิทธิภาพมากที่สุดในงานวิจัยชิ้นนี้ โดยที่อีกทั้ง 2 แบบจำลองประกอบด้วยแบบจำลองคอนโวลูชันอเล็กซ์เน็ต และ วีจีจี ที่มีผลการทดสอบรองลงมาตามลำดับจากแบบจำลองภูเก็ตเน็ต ผู้วิจัยเชื่อว่าความลึกของชั้นที่ 40 ชั้น ไม่ได้ส่งผลต่อแบบจำลองทั้ง 2 แบบมากนัก เนื่องจากเป็นแบบจำลองที่ไม่ได้มีชั้นของคอน

โวลูชันที่มากขึ้น ต่างจากแบบจำลองคอนโวลูชันเรสเน็ต แต่ผู้วิจัยมีความเชื่อว่าตัวอย่างของข้อมูลที่มีอยู่อย่างจำกัด เป็นตัวแปรสำคัญของแบบจำลองทั้ง 2 แบบนี้ เนื่องจากแบบจำลองคอนโวลูชันนอเล็กซ์เน็ต และวีจีจี เป็นแบบจำลองที่มีความเหมาะสมกับชุดข้อมูลที่มีปริมาณที่มาก และข้อมูลที่ครอบคลุมลักษณะสำคัญในชุดข้อมูล เพราะทั้ง 2 แบบจำลองมีจุดเด่นที่ความสามารถในการจดจำลักษณะของข้อมูลได้ดี ซึ่งงานวิจัยชิ้นนี้มีข้อจำกัดในเรื่องของชุดข้อมูลที่มีอยู่อย่างจำกัด จึงอาจส่งผลให้ประสิทธิภาพของทั้ง 2 แบบจำลองแสดงประสิทธิภาพออกมาได้อย่างไม่เต็มที่

จากที่กล่าวข้างต้น เนื่องจากข้อจำกัดของจำนวนความลึกของชั้นในชุดข้อมูลภาพที่มีเพียง 40 ชั้น อาจส่งผลกระทบต่อประสิทธิภาพของแบบจำลองเรสเน็ตที่มีความสามารถในการจัดการกับโครงข่ายที่มีความลึก อาทิเช่น Resnet-50 หรือ แม้แต่ Resnet -101 แต่เนื่องด้วยข้อจำกัดของความลึกของภาพที่มีเพียง 40 ชั้น เต็มที่ ซึ่งส่งผลต่อการทดสอบสมรรถนะของแบบจำลองในการแยกคุณลักษณะของรูปภาพชุดใหม่ได้ไม่ดีมากนัก ผู้วิจัยเชื่อว่าเป็นเหตุผลสำคัญที่ทำให้โครงข่ายคอนโวลูชันเรสเน็ตไม่สามารถแสดงประสิทธิภาพของแบบจำลองได้อย่างเต็มที่

ท้ายที่สุดนี้ จากการศึกษาในงานวิจัยชิ้นนี้หากต้องการให้ประสิทธิภาพของแบบจำลองดียิ่งขึ้น อาจมีการนำเทคนิคอื่น ๆ มาใช้ร่วมกับโครงข่ายคอนโวลูชัน อาทิเช่น การใช้เทคนิคประมวลผลภาพล่วงหน้าโดยจัดชั้นส่วนกระดูกในสมองออก (A. V. Dobshik, 2023) เพื่อให้แบบจำลองเรียนรู้เฉพาะบริเวณที่สำคัญ อีกทั้งการใช้เทคนิคแบ่งส่วนรอยโรค การตรวจจับ และวิเคราะห์หรรอย ประยุกต์กับโครงข่ายประสาทคอนโวลูชัน เพื่อให้โมเดลสามารถจำแนกภาพได้ดี และมีประสิทธิภาพยิ่งขึ้น

เอกสารอ้างอิง

- A. V. Dobshik, S. I. (2023). Acute ischemic stroke lesion segmentation in non-contrast CT images using 3D convolutional neural networks.
- Balasubramanian Raman, a. A. (2021). Towards effective classification of brain hemorrhagic and ischemic stroke using CNN. *ScienceDirect*.
- Bhanumathi, V. a. (2019). CNN Based Training and Classification of MRI Brain Images. *Institute of Electrical and Electronics Engineers (IEEE)*.
- Dickson, B. (2021). What are convolutional neural networks (CNN).
- Fartakov, E. I. (2019). Possibilities and prospects of hypothermia in the treatment of ischemic stroke. *ResearchGate*.
- Gautam, B. R. (2021). Towards effective classification of brain hemorrhagic and ischemic stroke using CNN. *ScienceDirect*.
- H. Fan, X. N. (2017). FPGA-based 3-dimensional convolutional neural network. 2017: in: International Conference on Field Programmable Logic and Applications (FPL) . ,. *Institute of Electrical and Electronics Engineers (IEEE)*.
- Hassan, M. (2018). VGG16 - Convolutional Network for Classification and Detection. 2018;. Retrieved from <https://neurohive.io/en/popular-networks/vgg16/>
- Liwei Tan, H. W. (2021). Stroke Lesion Detection and Analysis in MRI Images Based on Deep Learning. *Hindawi*.
- Lovekesh Vig, M. D. (2019). A Comparison of Shallow and Deep Learning Methods for Predicting Cognitive Performance of Stroke Patients From MRI Lesion Images .
- Ottesen, J. A. (2023). 2.5D and 3D segmentation of brain metastases with deep learning on multinational MRI data., . *Frontiers in Neuroinformatics*.

Pushpendra Singh, S. D. (2021). Automated detection of COVID-19 from CT scan using convolutional neural network. 2021: ScienceDirect. *ScienceDirect*.

World Stroke Organization. *World Stroke Organization*. (2021). Retrieved from <https://ncdalliance.org/world-stroke-organization-0>.

Zhang, H. Y. (2021). identification and Diagnosis of Cerebral Stroke through Deep Convolutional Neural Network-Based Multimodal MRI Images., . *Hindawi XML Corpus*.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.



Copyright: © 2023 by the authors. This is a fully open-access article distributed under the terms of the Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).